Shaping University Teaching Towards Measurement for Accountability

Problems of the
Graduate Skills Assessment Test

Kate Chanock, Rosemary Clerehan, Tim Moore and Anne Prince

The testing of graduate attributes is the new panacea in the search for more efficient forms of university teaching. Yet the new test designed to measure such attributes makes a range of questionable assumptions about who students are, what they've learned, and the cultural equipment they bring with them. Kate Chanock, Rosemary Clerehan, Tom Moore and Anne Prince sharpen their pencils and take the test.

Over the last decade, in both Australia and Britain, universities have been under increasing pressure to make themselves accountable for the extent to which they cultivate in their students transferable skills and 'attributes' that will prepare them for the rapidly-changing world of employment. The Government's *Striving for Quality* document (2002) begins by pointing out that 'Australia is actively positioning itself within an international 'knowledge-based economy', which has placed new demands on higher education.' Therefore, it goes on to argue,

> Higher education institutions should produce graduates with skills, knowledge and learning outcomes that promote individual development and that the nation requires for continued economic, social and cultural development. The new century is generating a need for 'emerging' skills and knowledge that have not been previously a focus of higher education. These include initiative and enterprise skills; information literacy and management skills… (p. ix)

The Government has questioned the effectiveness of higher education in this regard, in view of the Nielsen survey on employer satisfaction with graduates conducted in 2000. Here, employers reported that graduates often had weaknesses in the areas of 'creativity and flair'; oral (and, to a lesser extent, written) communication; 'interpersonal skills'; and 'understanding of business practice' (ACNielsen, 2000).

Much consideration has been given to the question of what role universities should play in the preparation for work; and, concomitantly, to ways of measuring the outcome of their efforts. This paper looks at a testing instrument commissioned by the Department of Education, Science and Training (DEST) and developed by the Australian Council for Educational Research (ACER). This 'Graduate Skills Assessment' (GSA) Test is still at a pilot stage, but the Government has shown some interest in it as a method of quality assurance, and appears to be leaving it open as a policy option in the future (DEST, 2003, p. 41). It is timely therefore to look closely at what this test promises and how it works.

WHAT ARE 'GRADUATE ATTRIBUTES'?

The focus and design of the GSA Test have evolved in the context of the larger movement towards adopting and implementing 'graduate attributes' in universities around the country. As the issue of educational accountability has been conceived mainly around the notion of graduate skills and attributes (the distinction between these terms is rarely clarified in the literature, see Clanchy & Ballard, 1995), a number of different bodies have put forward lists of these. What they encompass, generally, is what the Australian Technology Network (comprising five universities in different states) defines as:

> The qualities, skills and understandings a university community agrees its students should develop during their time with the institution. These attributes include, but go beyond, the disciplinary expertise or technical knowledge that has traditionally formed the core of most university courses. They are qualities that also

prepare graduates as agents for social good in an unknown future. (Bowden et al., 2002: Executive Summary)

The list proposed in the West Review (reported in Bowden et al., 2002), for example, includes

● the capacity for critical, conceptual and reflective thinking in all aspects of intellectual and practical activity;

● technical competence and an understanding of the broad conceptual and theoretical elements of his or her fields of specialisation;

● intellectual openness and curiosity, and an appreciation of the interconnectedness, and areas of uncertainty, in current human knowledge;

● effective communication skills in all domains (reading, writing, speaking and listening);

● research, discovery, and information retrieval skills and a general capacity to use information;

● multifaceted problem solving skills and the capacity for team work; and high ethical standards in personal and professional life, underpinned by a capacity for self-directed activity.

By now, most Australian universities have developed statements setting out the graduate attributes they value (Hager, Holland & Beckett, 2002). These commonly reflect, but often go beyond, the list above. Additions (from the lists available on each university's website) encompass a range of liberal attitudes whose origins go back to Newman's *The Idea of a University* (Jones, 2002). These include aims to develop 'social responsibility' (University of New England); to 'value social justice, tolerance and responsible action' (Victoria University); to 'work toward improvement in society' and 'act in environmentally sustainable ways' (University of Canberra); and 'a desire to continually seek improved solutions and to initiate, and participate in, organisation and social change' (University of Wollongong). They also include attributes considered advantageous in the world of work, such as 'be[ing] flexible in approach and adaptable to change' (Victoria University); 'hav[ing] the ability to initiate new ideas, implement decisions and cope with uncertainty' (University of Canberra); 'be[ing] able to work effectively within culturally diverse settings' (Victoria University); and 'demonstrate[ing] international perspectives as a professional and as a citizen' (University of South Australia). These give us a picture of the kinds of qualities and skills which universities have nominated as desirable for graduates (although how they could be expected to demonstrate some of these, at the moment of graduation, remains puzzling).

ACER explains that both universities and employers were consulted in developing the GSA. This process produced a list of seventeen skills including (among others) a skeptical habit of mind; toleration of uncertainty; and global and cross-cultural perspectives (ACER, 2002a). The developers then narrowed the list to a set of skills that could, in their view, be readily and reliably tested by a psychometric test composed

of mainly multiple-choice questions. The categories chosen to be tested were 'interpersonal skills', 'critical thinking', 'problem solving', and 'written communication'.

On its website, ACER has made available a sample set of questions representing the types of questions that might be asked in any iteration of the GSA (ACER, 2002b). It has also posted an example of the assessment report that candidates receive, showing them their score, how it compares to 'the distribution of the results for the middle 60% of students from fields of study similar to [theirs]', and what the scores at each level indicate about the candidate's ('Sam Sample's') skills in each category. In this paper, we will quote from the sample questions, and from the descriptions of what ACER says can be inferred about a candidate who scores at Level 3 (the highest level). We will consider to what extent particular sample questions seem able to test what they purport to test, at the level at which they claim to test it. It is not wholly satisfactory to base this examination on such a partial sample – it may be that questions we have not seen would be more suitable than those we have – but we can only comment on the ones that are publicly available.

CHALLENGES IN DESIGNING THE TEST

The task which ACER has set itself is bound to be a difficult one, because the test is to be administered to students in all fields, both entering and leaving their degree courses. The content, therefore, has to be material that any secondary school leaver can understand; but, at the same time, show what 'value' has been 'added' to the candidates' skills in the course of their degree. Moreover, while universities invite their students to deal with complexity and ambiguity, to identify problems and frame questions, and to gather and consider a wide array of evidence, the test has to present its questions and problems de-contextualised, has space for very little evidence, and uses a format that asks candidates mainly to eliminate wrong answers from a set of choices. The lack of fit with common objectives of learning in higher education has raised concerns, even on the part of the body that commissioned it.

The Government's *Striving for Quality* discussion paper (DEST, 2002) provides an overview of the arguments for and against the national testing of graduates generally, and of the GSA Test (pp. 20–21). On the one hand, it is argued that such a test provides 'an impartial measure' of student performance that is 'not coloured by differences in academic standards in particular courses or institutions'. But on the other, the authors of *Striving for Quality* acknowledge that such tests can end up just measuring 'test-taking ability rather than generic skills and that students can be 'taught to the test', which can skew curriculum and the integrity of learning experiences'. Additional concerns are expressed in the paper about whether a single high-stakes exit test can provide a complete

picture of graduate capacity and whether such a regime might discriminate against already disadvantaged groups.

Weighing up the need it perceives for an easily understood, nationally standardised measure of universities' success in inculcating graduate skills against the objections noted above, the Government then asks,

● Should the Commonwealth mandate that the Graduate Skills Assessment Test becomes a requirement of entry to and exit from higher education, providing students and employers with an effective indicator of student generic attributes?
● Should the Commonwealth, directly or indirectly, co-ordinate the development of an integrated, accessible publication of outcomes data to inform the community on the relative quality of universities? (Striving for Quality, p. 33).

An examination of the test itself should help provide us with an answer to these questions.

SAMPLE QUESTIONS FROM THE GSA TEST

Section 1: Interpersonal Understandings

Perhaps the most problematic section of the sample questions is the category of 'Interpersonal Understandings', aimed at testing a candidate's sensitivity to the feelings, attitudes and behaviours of others, particularly as they relate to work contexts. The claims about what this section reveals about a candidate's interpersonal skills (if s/he scores highly) are both broad and deep, encompassing skills that employers would certainly like new applicants to have:

LEVEL 3 *(The highest level)*
● Demonstrates sophisticated insight into, and makes subtle inferences about, roles, relationships, behaviours, feelings, attitudes and motives.
● Demonstrates subtle insight into aspects of effective teamwork, leadership, negotiation and communication.
● Recognises potentially appropriate actions or responses to delicate or complex interpersonal problems.

How, then, does the marker know that the candidate is capable of such insights? Some of the questions on which this judgement is based refer to a passage of dialogue that shows a group of students organising themselves for teamwork:

Raelene, Carlos, Betty and Andy have a tertiary studies project to complete within a very tight deadline. They are all accountable for the final product, and all must demonstrate their individual contribution.

Betty: There's so much to do and so little time! Let's meet together socially first so that we can really get to know each other.

Or should we brainstorm some ideas first?

Raelene: Well, I think we should first work out what each of us is going to take responsibility for. What would you rather do, Andy?

Andy: As little as possible! *(Group laughter.)*

Raelene: Yes, but what contribution can you make? Are you better at writing/editing, or research, or talking to people?

Andy: Oh, whatever. I'm comfortable with any of those. I'll just do what I'm directed to do.

Raelene: Carlos? What about you?

Carlos: Well, I don't really like talking to people on the phone so I'd prefer to do the research.

Betty: Oh, I thought I could do that. Perhaps Carlos and I could both do the research and Andy could do the phone contacts. Raelene, you could do all the writing up and presentation.

Raelene: Well, we probably only need one person to do the research so it might be better if Carlos does that and if you and Andy do the phone contacts.

Betty: Well, I'd really prefer to do the research. That's something I'm good at. Carlos would be OK with Andy helping him.

Andy: I don't care what I do or who I work with. Just tell me, someone, so I can go and have my lunch!

Raelene: Look, to get the project done in time we have to make sure everyone has a specific task. Carlos, you're probably better off doing the research and Betty, we need someone confident for the telephone communication side of the project. Andy doesn't mind what he does so he could help you with the phone calls. How does that sound?

In this passage we get a glimpse of the group's dynamics, with Raelene taking charge and Betty resisting Raelene's efforts to assign her a task she does not think best suited to her abilities. The associated question that candidates must complete interrogates their understandings of the needs of the group, as opposed to the interests of any individual. It is apparent that the project will go ahead faster if the option in 'C' is taken:

18. Which one of the following responses from Betty would best meet the group's needs now?

A Betty challenges the leadership role that Raelene seems to have assumed on the grounds that she, Betty, would be a better leader.

Shaping University Teaching Towards Measurement for Accountability

B   Betty rejects Raelene's suggestion on the grounds that Raelene is the best person in the group to do the telephoning.

C   Betty accepts the suggestion and begins thinking about the people she and Andy need to contact.

D   Betty repeats her suggestion of the group meeting socially to talk through the issue.

E   Betty reluctantly gives way and then sits back waiting for instructions.

If we share the GSA developers' values, and their assumptions about how people should behave in an unstructured social situation centred on a task, the question seems straight-forward enough; but still we may question whether the choice of 'C' really shows that a candidate is able to form 'sophisticated insights' and make 'subtle inferences'? Is it possible, indeed, to do this on the basis of a few lines of uncontextualised dialogue?

The format of the test must raise some concerns, in this regard, about the possibilities of both gender bias and cultural bias. Males have been found to be more successful in multiple choice tests than females (Gipps & Murphy, 1994), an effect that has been attributed to their willingness to 'abstract the problem from the context, while girls attend to the totality of the problem context, focussing on more clues than do the boys' (Gipps & Stobart, 1993, p. 59). By presenting a social interaction, the test invites students to attend to 'clues'; but by requiring them to choose one answer from an array, it works against their ability to do so. For question 19 ('Which one of the following assessments is most accurate in relation to Betty?'), the 'right' answer is the following option:

> **E.** *Strengths:* confident, assertive, willing worker.
> *Weaknesses:* inclination to be forceful, persistent

We might wonder how 'persistence' gets to be construed as a weakness. And we are trying to recall instances where male leaders in our society get criticised for an 'inclination to be forceful'.

Turning from gender to culture, if we were to bring a different set of cultural assumptions to the test we might choose one of the other options. In a culture that values reaching a consensus through discussion until all parties feel their views have been properly considered, 'D' could be considered a better choice for the project to be conducted effectively (Cortazzi & Jin, 1997). Alternatively, a more hierarchically organised cultural frame might seek to challenge Raelene's assumption that she can simply step into the leadership role unbidden. A sense of how the test-writer's own culture and values influence identification of the 'correct' response is inescapable. In one question, there appears to be a tension between gender values and management values. On the whole, however, high scores in this section will reflect an ability to recognise (though not nec-

essarily, of course, a commitment to apply) Western assumptions about personality, sociability, social roles and management values. From an examination of the sample questions, it seems a large claim that this will require subtle inferences; at the same time, the framing of the questions and desired responses appears to conflict with the ethos of cultural inclusiveness which most universities explicitly endorse.

In every sample question we have seen, in fact, 'interpersonal relations' are treated as if they were conducted in the same way everywhere. But what is considered appropriate and effective social behaviour in any situation is defined by culture (including in this case a broad range of workplace cultures that exist in our own society). Praise is lavished in one culture, withheld in another. Directness is mandated in one, avoided in another. The role that power, authority, experience, politeness, aggression or deference ought to play in a negotiation varies widely. In line with this thinking, the best way of establishing a team could well encompass 'C' or 'E' below, as well as the 'correct' answer, 'A':

> **Unit 10, Question 22:** You have been asked to set up a new project team. Which one of the following is likely to be most appropriate and important at the first formal meeting of a new project team?
>
> **A**   Establish a common sense of direction and expectations.
> **B**   Agree on a set of rewards and punishments for behaviours during the project.
> **C**   Start by having some drinks and food to relax everyone and keep things sociable.
> **D**   Make lists of strengths and weaknesses of people and share these prior to delegating tasks.
> **E**   Appoint the most dominating person as leader and chairperson so they can direct the project.

(Indeed, it is perhaps surprising, in view of Raelene's triumph in the previous scenario, that 'E' is not the 'right' answer here.)

Section 2: Critical Thinking

For the testing of 'Critical Thinking', the challenge for the designers of the GSA Test is to find a suitable generalist content on which the assessment of these abilities might be hung. Again, we turn to the report on 'Sam Sample' to see what these questions purport to show about high-scoring candidates. A person who scores at the highest level, according to this report, is one who:

● Comprehends complex and implicit meanings and relationships in text, and makes subtle and cogent inferences about these.

● Analyses text and uses inference to identify subtle or complex evidence, lines of reasoning, logical flaws, arguments, assumptions, consequences, rhetorical devices, analogies etc.

● Evaluates credibility and validity of complex or subtle evidence, reasoning and argument implicit in text, generating appropriate criteria for evaluation if required.

To develop a test instrument that can accurately measure students' abilities to engage in such a complex range of thinking skills is not an easy task. This is largely for the reason pointed out by McPeck (1981), that critical thinking of its nature must always be directed towards some specific content. 'Thinking, by definition, is always thinking about something, and that something can never be 'everything in general' but must always be something in particular' (p. 4). That 'something in particular' for students is the diverse range of subjects and disciplines they choose to study. And significantly, as McPeck (1992) and others (eg. Clanchy & Ballard, 1995) have suggested, the standards for critical thinking that apply in each of these disciplinary areas may not be exactly identical.

An additional challenge for the test designers is to come up with items which have unequivocally correct and incorrect answers. One such attempt is the item below which requires candidates to assess the relationship between:

i) a proposition (Our society will benefit from less government intervention and regulation) and

ii) related statements (The future offers great opportunities and great challenges for our society to deal with/Less government intervention should involve less expense on parliament and politicians)

The following proposition is the basis of questions 16 & 17.

'Our society will benefit from less government intervention and regulation'.

Each of the questions below contains a statement and a set of alternatives, A–E. For each question, you are to indicate which alternative most appropriately describes the relationship between the statement and the proposition.

16. Statement: 'The future offers great opportunities and great challenges for our society to deal with'.
In relation to the proposition, the statement:

A offers significant support.

B can be used as a significant counter.

C merely repeats or offers insignificant support.

D merely contradicts or offers an insignificant counter.

E is irrelevant and cannot be used as support or to counter.

17. Statement: 'Less government intervention should involve less expense on parliament and politicians'.
In relation to the proposition, the statement:

A offers significant support.

B can be used as a significant counter.

C merely repeats or offers insignificant support.

D merely contradicts or offers an insignificant counter.

E is irrelevant and cannot be used as support or to counter.

In item 16, the correct response is given as E (a relationship of 'irrelevance') which on first pass may appear straightforward enough. But there can always be a problem if one thinks about these issues a little too much. For example, if one came to the proposition with concerns about a relinquishing of control to non-government forces (eg. market forces) – as some students do – then the assertion that our society faces 'great challenges', might be seen as a 'significant counter' (option B) to the laissez faire position being proposed. That is to say that, a student might believe that society can only deal effectively with the 'challenges' it faces, through some collective governmental action. Indeed 'interventionism' as a political philosophy is often premised on such a belief. Thus, the student of politics may find herself disadvantaged by knowing just a bit too much about the issue in question – and by subjecting it to just a little too much critical analysis.

Similarly in item 17, the correct answer, 'A' as given (a relationship of 'significant support'), may also be contentious. Again a more critical and informed student might query whether some potential cost saving on parliaments and politicians really qualifies as a 'significant support' (our emphasis) for a deregulated, laissez faire system. The student might indeed conclude that, as support, it is a bit on the 'trivial' side (option C) – at least trivial when compared to other larger putative benefits of a market-oriented system he has read about in his economics degree. Indeed, on this point, we note that both items 16 and 17 rely very much on the notion that distinguishing between so-called 'significant' and 'insignificant' support (or counter) will be entirely unproblematic for students. But these are surely matters of individual judgment – and we cannot be at all certain that a test-taker's sense of this distinction will correspond exactly with that of the test designer.

The main problem in this section stems from the test designers' belief that a complex range of skills can be effectively tested within such a narrow and rigid format. McPeck (1981) has suggested that critical thinking is best characterised as 'an appropriate use of reflective scepticism'. In relation to the items we have considered, an appropriate use of this scepticism might be to question whether the relationship between such propositions and statements is really as unproblematic as is being suggested. Certainly students in the intellectual tasks they engage in in their university degrees – in the writing of essays, in discussions in tutorials etc. – will have got into the habit of thinking that these matters are never as simple as they might first appear.

Section 3: Problem Solving

The section on 'Problem Solving' tests numeracy and the ability to understand information presented in graphic or tabular form. These are certainly desirable skills for a university graduate to have; and some of the problems posed are ones that can be expected to arise in the workplace. The one below, for example, has to do with arranging a timetable to avoid clashes:

Shaping University Teaching Towards Measurement for Accountability

Question 1. Simon is responsible for scheduling the weekly meetings of four sporting clubs at his campus. The meetings are to be held at lunchtimes, Monday to Friday. Each club has one meeting each week. All the clubs are able to meet in any lunchtime when they do not have training. The clubs train at lunchtime(s) on the following days:

Club	Training Days
Netball	Mondays and Wednesdays
Mountaineering	Tuesdays
Triathlon	Mondays and Thursdays
Orienteering	Fridays

People who join a club must attend all training sessions for that club.

1   Simon draws up four possible schedules for the meetings, shown below. Which one of these schedules does not involve a clash with a club's training?

A

	M	Tu	W	Th	F
N					*
M	*				
T				*	
O			*		

B

	M	Tu	W	Th	F
N					*
M	*				
T		*			
O			*		

C

	M	Tu	W	Th	F
N					*
M		*			
T			*		
O	*				

D

	M	Tu	W	Th	F
N		*			
M			*		
T		*			
O					*

E   **A**ll four schedules involve a clash.

Any workplace would welcome an employee who can solve timetabling problems. For this question however, the candidate is not asked to solve the problem, as this has already been done by the test-writer. All the candidate is asked to do is to choose among the ready-made alternative solutions. The Sample Report says that a candidate who scores at the highest level in this kind of problem 'applies strategies to solve problems involving complex and abstract relationships and multiple steps and conditions, and evaluates solutions to such problems'. Yet the most that a candidate can demonstrate, by getting the correct answer here, is 'evaluat[ing] solutions to such problems'.

Again, because the questions are set at a level that high school leavers could tackle, they are not likely to be able to show how much more a university graduate can do. To ask the question which we have just examined is not without merit: the correct response simply does not match the claims that are being made for it in the ACER list of 'skills displayed'. And in a real workplace, the correct answer might be 'do it on a spreadsheet'.

Section 4: Written Communication

It is in the section on 'Written Communication', however, that the problem of testing at a suitable level is most apparent.

According to the test report, a student who performs at the highest level:

● Demonstrates insightful and critical understanding and analysis of ideas and issues

● Organises shapes and develops material effectively and coherently for the required purpose

● Uses language precisely and fluently, with effective command of vocabulary, syntax and other linguistic conventions

These are reasonable requirements of a piece of text. But let us examine a sample task:

> **Written Communication. Argument Task**
> Consider the following comments and develop a piece of writing presenting your point of view on one or more of the issues.
>
> Your response will be judged on:
> ● The quality of your ideas and opinions, regardless of the position you take;
> ● How well your argument is organised and structured; and
> ● How clearly and fluently your views are expressed.
>
> **A**   The media is the watchdog of society, bringing into the open things that many people would prefer to remain secret.
> **B**   News is more often than not created by the media rather than simply reported by it.
> **C**   The media is just the modern form of the age-old need to inquire after, listen to and pass on information.
> **D**   The pervasive influence of the media has meant that people are no longer able to genuinely think for themselves; they just repeat whatever they've read in the paper or heard on television or radio.

This requires of candidates the kind of writing that university students are taught not to do: that is, to air their opinions without recourse to any reliable sources of evidence. Leaving aside the point that no writing topic can be 'discipline free' and Journalism and Communications students will start from a different knowledge base from the rest, the 'content' presented here is a set of assertions about which students are expected to adopt opinions and argue.

Without a knowledge base or access to anything that might count as evidence, students are reduced to doing what their lecturers would fail them for: mounting an argument based on personal experience and individual opinion. This limitation of the task format places considerable stress on the adequacy of the first descriptor to reflect the kind of learning conferred by a university education. The focus of the second on organising 'material' – which asks students to muster whatever they have in their head – renders the 'required purpose' a fairly limited one. And so, to assess whether the candidate uses language 'precisely and fluently' will be difficult in these circumstances as the assessment will not be able to access the kind of writ-

ing which students can produce when they are engaging cognitively with a body of knowledge.

SUMMARY: LIMITATIONS OF THE GSA TEST

Among the problems with the GSA, then, is first and foremost that it neither tests what it purports to test, nor pitches its tasks at the level to which we would hope these skills would develop in the course of a university education. This is particularly worrying because the Department of Education, Training and Youth Affairs has said, in its *Higher Education Report* (2004), 'This tool provides an objective measure of the generic skills of graduates'. The intention is to promote the test to employers, whereby options to be followed up include 'use of the GSA as a standard recruitment tool due to the transferable and employer relevant nature of the four skill domains' (pp. 25–26).

In addition to the question of the test's validity, as we have suggested, it seems likely to disadvantage some candidates through the cultural assumptions underlying it. And, while we have not focused on the question of language here, this could further constitute a particular difficulty for non-English speaking background students (eg. question 6, 'By how much does the total number of lessons that the five teachers can teach exceed the number of lessons required to teach the five subjects to the Year 7 class?') It is not surprising that it is noted in the ACER GSA Summary Report that 'Language spoken at home' seems to be a significant factor in the performance of candidates (ACER, 2001, p. vi).

A further group that may be disadvantaged by the test is students with learning disabilities or mental illness. Gosden and Hampton (2001) point out that 'Generic skills testing specifically discriminates against students with learning disabilities and some mental disorders because it is the inability to master particular types of generic skills that defines these types of disabilities … skills in reading, writing and the application of focused attention' (pp. 20, 26).

Perhaps the most serious concern is the negative effect that the test, as a quality assurance instrument, could have ultimately on the quality of university programs. If the test is mandated in the system, universities, for reasons of their survival, will want to ensure that their students do well on it. In such an arrangement, it is not hard to envisage valuable time being given over in already overcrowded curricula to training students in the ultimately trivial skills of test preparation.

And in the light of our analysis above, if students are successful finally on the test, we cannot be confident that they do in fact possess those skills the nation requires for 'continued economic social and cultural development' (DEST, 2002). The only thing we can really be confident about is that they have the skills to pass a psychometric test whose validity we question.

WHAT ALTERNATIVE IS THERE?

Fortunately, a test such as the GSA is not the only way in which community concerns about the development of desirable skills and attributes in higher education can be addressed.

In their response to the Government's higher education review, titled *Forward from the Crossroads: Pathways to Effective and Diverse Australian Universities* (AVCC, 2002), the Australian Vice-Chancellors' Committee rejects the GSA as an unsuitable instrument for assessing what a university education adds to the generic skills of graduates from different fields. It does not, they feel, justify its cost of approximately $12 million a year to test 200,000 students (p. 28). A number, in fact, are opting – or thinking of opting – for an alternative which is not without problems but is in many ways more suited to the task.

It is not a psychometric test but a thorough audit of the ways that universities teach and document what their students are learning. Hager, Holland and Beckett (2002) believe that this will offer benefits to universities in stimulating course development, and to their students in encouraging more reflective learning and enhancing their employability. It will, however, involve time and effort.

As the five ATN universities describe it in their report (Bowden et al, 2002), universities pursuing this sort of audit are engaged in a process of:

● Identifying suitable graduate attributes for themselves.

● Identifying where, in the curriculum, efforts are made to inculcate these.

● Where these learning opportunities are implicit, making it explicit what is being learned and how (on the need for this, see also Kemp & Seagraves, 1995, who found that students often did not recognise integrated skills instruction unless it was quite explicit).

● Where these opportunities are lacking, finding ways of integrating the development of the desired attribute into the curriculum.

In this process, the audit, the curriculum development, and the documentation are all important elements (the ATN report on the web links to a number of case studies where these issues are being addressed). If this can prove feasible and affordable, and if developments are carried out by the universities with appropriate top-down and bottom-up consultation, this seems to offer a more productive method of monitoring skills development than does the GSA Test. The approach is consistent with the recommendation of Hyland and Johnson (1988), that

> Since there is no common or agreed definition of (key) skills – with proponents differing about whether core/key skills apply to specialised areas of knowledge, outcomes of behaviour, ways of

Shaping University Teaching Towards Measurement for Accountability

thinking, or the attitudes, values and personality traits of students – such generalised skill-talk is fatuous and redundant, and should be replaced by descriptions of common curriculum experiences. (p. 164)

CONCLUSION

In this paper, we have discussed some of the problems with the GSA that are suggested by the sample questions ACER has so far made available. The use of this instrument is at an early stage, as is the alternative being attempted by the ATN and other universities across the country. In deciding which methods to adopt to teach and test graduate skills, university communities will be wise to consider the detail as well as the intentions of the alternatives and their implications for teaching and learning in their institution. *a*

Kate Chanock heads the Humanities Academic Skills Unit at La Trobe University. Her background is in Anthropology, African History, and teaching English as a Second Language. Her main research interests are the cultures and discourses of academic study and the challenges facing students with a learning disability.

Rosemary Clerehan is Head of the Language and Learning Services Unit at Monash University. In additional to graduate skills assessment, her research interests include evaluation of information for patients; web-based support for student writing; and cross-cultural issues in teaching and learning.

Tim Moore works in Language and Learning Services Unit, Monash University based in the Faculty of Arts. He is currently doing PhD research into critical thinking and the disciplines.

Anne Prince has a background in ESL/EFL teaching and Higher Degrees by Research administration. She currently works in the Education Faculty at Monash University.

REFERENCES

ACNielsen Research Services (2000), *Employer satisfaction with graduate skills.* Evaluations and Investigations Programme, Department of Education, Training and Youth Affairs, Canberra. Retrieved November 11, 2004, from http://www.dest.gov.au/archive/highered/eippubs/eip99-7/eip99_7pdf.pdf

Australian Council for Educational Research (2001), *Summary Report.* Retrieved November 11, 2004, from http://www.acer.edu.au/tests/university/gsa/summary_test_reports.html

Australian Council for Educational Research (2002a), *University Testing.* Retrieved November 11, 2004, from http://www.acer.edu.au/tests/university/gsa/development.html

Australian Council for Educational Research (2002b), *GSA Sample Questions.* Retrieved November 11, 2004, from http://www.acer.edu.au/tests/university/gsa/samplequestions.html

Australian Vice-Chancellors' Committee (2002), *Forward from the Crossroads: Pathways to Effective and Diverse Australian Universities.* The AVCC submission to the Higher Education Review. Retrieved November 11, 2004, from http://www.avcc.edu.au/documents/policies_programs/statements/Response_26Sept02.pdf

Bowden, J., Hart, G., King, B., Trigwell, K., & Watts, O. (2002), *Generic capabilities of ATN university graduates.* Teaching and Learning Committee, Australian Technology Network (RMIT, QUT, UTS, Curtin & UniSA). Retrieved November 11, 2004, from http://www.clt.uts.edu.au/ATN.grad.cap.project.index.html]

Clanchy, J. & Ballard, B. (1995), 'Generic skills in the context of higher education,' *Higher Education Research and Development,* 14 (2), 155-166.

Cortazzi, M. & Jin, L. (1997), 'Communication for learning across cultures'. In D. McNamara, & R. Harris (Eds.), *Overseas Students in Higher Education,* London & New York, Routledge, 76-90.

Department of Education, Science and Training (2002), *Striving for Quality. Learning, Teaching and Scholarship,* Commonwealth of Australia. Retrieved November 11, 2004, from http://www.dest.gov.au/crossroads/pubs.htm#quality

Department of Education, Science and Training (2003), *Our Universities: Backing Australia's Future,* Canberra, Commonwealth of Australia.

Department of Education, Training, and Youth Affairs (2004), *Higher Education Report for the 2004 to 2006 Triennium,* Commonwealth of Australia. Retrieved November 11, 2004, from http://www.dest.gov.au/highered/he_report/2004_2006/default.htm

Gipps, C. & P. Murphy (1994), *A Fair Test: Assessment, Achievement, and Equity.* Buckingham, Open University Press.

Gipps, C. & Stobart, G. (1993), *Assessment: A Teacher's Guide to the Issues* (2nd edition). London, Hodder & Stoughton.

Gosden, R. & Hampton, G. (2001), Generic skills assessment: A new problem for tertiary students with learning disabilities. *Australian Journal of Learning Disabilities* 6 (1), 20-27.

Hager, P., Holland, S., & Beckett, D. (2002), *Enhancing the learning and employability of graduates: The role of generic skills.* Business/Higher Education Round Table. Available by ordering from http://www.bhert.com/publications_PolicyStatements.htm

Hyland, T. & Johnson, S. (1988), 'Of cabbages and key skills: Exploding the mythology of core transferable skills in post-school education'. *Journal of Further and Higher Education,* 22 (2), 163-172.

Jones, J. (2002), 'Generic attributes: An agenda for reform or control?' In B. James, A. Percy, J. Skillen & N. Trivett (eds.) *Changing identities: Proceedings of the Language and Academic Skills Conference held at the University of Wollongong 29-30 Nov. 2001.* Retrieved November 11, 2004, from http://learning.uow.edu.au/LAS2001/index.htm

Kemp, I. J. & Seagraves, L. (1995), 'Transferable skills – can higher education deliver?' *Studies in Higher Education,* 20 (3), 315-328.

McPeck, J. (1981), *Critical thinking and education.* New York, St. Martin's Press.

McPeck, J. (1992), 'Thoughts on subject specificity'. In S. Norris (ed.) *The generalizability of critical thinking: Multiple perspectives on an educational ideal* (pp. 198-205). New York, Teachers College Press.

Whitston, K. (1988), 'Key skills and curriculum reform'. *Studies in Higher Education,* 23 (3), 307-319.