This paper is part of the symposium on "Gender and Assessment of Physics in Context: Getting It Right!" It examines ways in which current practices privilege the "masculine" over the "feminine" and presents an agenda for gender inclusive assessment practices. It is argued that physics like other domains of knowledge, is a constructed entity, and the way it has been interpreted by school curriculum and assessment practices has ensured it has strong links through hegemonic masculinity to power in our society. A case study of assessment practices from Victoria, Australia, indicates how some changes in a state-wide end-of-schooling credential, the Victorian Certificate of Education (VCE), dramatically changed the achievement of girls relative to boys in physics. A twenty year bias in assessment, in favor of boys, was turned around in physics when a wider variety of learning activities, skills and tasks were valued by becoming part of the work requirement and common assessment task format for the VCE. This section of the symposium also demonstrates that whilst this sudden shift in performance challenges what we define as "excellence" and "competence" in physics, it is not true for all girls or all boys. Socioeconomic factors intersect with gender performance profiles when data of girls in the VCE is collated by geographical region. Contains 69 references. (Author)
Interrupting Gendered Assessment Practices

by

Gaell M. Hildebrand

University of Melbourne
Interrupting Gendered Assessment Practices.

ABSTRACT:

It is through our assessment practices that we give clear messages to students about what it is we most value. This section of the symposium examines the ways in which current practices privilege the "masculine" over the "feminine" and presents an agenda for gender inclusive assessment practices. It is argued that physics, like other domains of knowledge, is a constructed entity, and the way it has been interpreted by school curriculum and assessment practices has ensured it has strong links through hegemonic masculinity to power in our society. A case study of assessment practices from Victoria, Australia indicates how some changes in a State-wide end-of-schooling credential, the Victorian Certificate of Education (VCE), dramatically changed the achievement of girls relative to boys in physics. A twenty year bias in assessment, in favour of boys, was turned around in physics when a wider variety of learning activities, skills and tasks were valued by becoming part of the "work requirement" and "common assessment task" format for the VCE. No longer is achievement based only on examinations made up of decontextualized assessment tasks. Girls now excel at physics in Victoria. This section of the symposium will also demonstrate that whilst this sudden shift in performance challenges what we define as "excellence" and "competence" in physics, it is not true for "all girls" or "all boys". Socioeconomic factors intersect with gender performance profiles when data of girls in the VCE is collated by geographical region.
The image of science is strongly gendered and aligned with hegemonic masculinity (e.g., Harding, 1986; Tuana, 1989; Thomas, 1990; Kirkup and Keller, 1992) but this mystique is a distortion of the concealed reality which frequently accommodates concepts from the right column of dualisms.

Linda Shepherd (1993) reveals the existing, but heavily veiled, feminine face of science that includes:

- **knower/known interactivity**: Heisenberg’s uncertainty principle in physics (you cannot measure both the momentum and the position of an electron because in measuring one you interfere with the other) along with Chaos theory reveal the interdependence of the observer and the observed;
- **subjectivity**: feelings are significant when research is motivated by love and desire and where hunches come before hypotheses;
- **multiplicity**: a web of interactivity exists between and among phenomena;
- **cooperation**: the importance of care and empathy in sustaining an harmoniously working research team;
- **intuition**: another way of knowing which is valued in highly esteemed, speculative scientists; and
- **holistic**: seeing the relatedness of ideas through interdisciplinary studies which show larger patterns, challenges underlying values like simplicity, abstraction and reductionism in science.

She argues that only when science integrates the feminine with the masculine, and replaces either/or conceptualisations with both/and thinking, will there be an acceptance of the complexities of reality.

Yet, as Jay Lemke shows school science further distorts the field of science by:

- generating a catalogue of ‘facts’ for students to recall and presenting science as if it is possible to produce absolutely objective truths;
- pretending that a scientific method exists - even when we know that real scientists, funded through politically-driven sources, seek evidence through using the research techniques that will most likely provide what they desire;
- teaching with the expectation that only a ‘super-intelligent elite’ can ever understand science’s concepts; and thus teaching most students to trust powerful technocrats and politicians who make decisions based on scientific, and hence unchallengeable, evidence. (Lemke, 1990)

**Gendered dualisms and assessment**

For each pair of gendered dualisms, looking through the post-structural feminist lens enables us to challenge the ‘either/or’ assumptions that value the dominant paradigm, allowing us to see that ‘both/and’ notions can lead to a more equitable view of assessment.

---

**Reward holistic learning**

(both abstract and holistic)

Assessment has largely valued abstraction and analysis over holism, relatedness and synthesis, particularly in science where the real world is often seen as too ‘messy’ and complex to illustrate with neat mathematical models. If our assessment procedures or ly examine students’ ability to suspend what they know about their world, while they blindly manipulate formulae or regurgitate information, then we ought not be disappointed when students fail to apply concepts to the real world.

Ensuring science and technology are considered in their social context with assessment of their benefits for the environment and human beings may be the most important change that can be made in science teaching for all people, both male and female. (Rosser, 1990, p. 72)

Anecdotal evidence suggests that many girls do prefer to learn concepts situated in their social context rather than abstract, fragment and compartmentalise their understandings. To value holistic learning, assessment tasks ought to be set within a social context and reward synthesis of ideas where theory and practice are clearly interconnected.

**Encourage qualitative understanding**

(both quantitative and qualitative)

Testing the authenticity of a proposed assessment task by checking whether it seeks evidence of qualitative understanding, rather than simplistic manipulation of quantitative data, is one way to recognise that many girls strive for this. For example:

My curiosity simply did not extend to the quantitative solution. I just didn’t care to figure out how much. I was more concerned with the ‘why’ and the ‘how’. I wanted verbal explanations with formulae and computations only as a secondary aid. Becoming capable at problem solving was not a major goal of mine. But it was a major goal of the course. (Michelle in Tobias, 1990, p. 40)

There is considerable anecdotal evidence from teachers which suggests that girls are more troubled by a feeling that they ‘don’t really understand’, an important factor in their withdrawal from subjects/courses. Boys appear to be less concerned by this and will continue a subject when their grades indicate that they ‘know enough’. If we value deep understanding then we should be build it into our assessment processes. Also, many students, frequently girls, want feedback on their work that goes beyond a quantitative grade. As assessors of student achievement we need to provide extended oral and
credibility and applicability than validity; about dependability and authenticity than reliability (Guba and Lincoln, 1989; Cambourne and Turbill, 1994; Gipps, 1994). In this model of assessment the teacher, as a human being, is seen as a responsive instrument, able to detect many nuances of performance from multiple sources, which no external, objective test can ever perceive. Compare this with the dominant paradigm which values external assessors over the teacher, testing over work requirements, written over oral, quantitative over qualitative, print over other text forms, and so on.

The knower is not distanced from the known (both knower/mind and knowable/Nature)

Within both science and assessment, the dominant view is that the knower is distanced from the known, the relationship between the two ‘is that between a subject and an object, radically divided, which is to say, no worldly relation.’ (Keller, 1985, p.79) Hence, a corollary of the objectivity/subjectivity dualism is the desired separateness of the learner from the material they learn, and from the observer judging their learning. Yet we know that learners are not distanced from, but are from the material they learn, and from the observer judging their learning. Yet we know that learners are not distanced from, but are formed by, their learnings, and that observers/assessors make decisions based on their understandings of what learners ought to be able to know and do.

Yet, students arrive in our classrooms with prior understandings and conceptions that have been constructed over time through their unique interactions with their world. This is the basis of constructivism (Fensham, Gunstone and White, 1994) which is currently producing a revolutionary paradigm shift (Kuhn, 1970) in the teaching and learning of science. Girls and boys, as groups, generally have had very different out-of-school experiences which result in their school learning beginning at different starting points. Sometimes these prior experiences are assumed for all students and assessed accordingly, even though one group such as girls, may have had little opportunity to learn the skills, or about the phenomena, outside the classroom.

Because of prior experiences and constructions, assessment practices have commonly given an unfair advantage to particular groups of boys. For example it has been shown (e.g. R. Murphy, 1982; Harding, 1981, 1991; Blum and Azencot, 1986; P. Murphy, 1989; Gipps and Murphy, 1994) that many girls tend to do better in assessment tasks composed of structured and extended response questions, whereas boys, as a group, will do better if questions are posed in a multiple choice format. Thus, designing a test exclusively using a single question format would advantage one sex over the other, simply by the question format.

Additionally, students bring gendered interpretations of their own success or failure. For example, causal attribution studies (e.g. Ames, 1984) show that when girls do well they often attribute their success to external factors such as luck, a good teacher or easy assessment tasks, whereas boys tend to attribute their success to internal factors such as innate ability.

Constructions of gender also interact with self-perceptions of ability when boys tend to over-estimate strongly in mathematics, less in English, whilst girls tend to be closer in estimation of their actual performance, but to underestimate more in mathematics than boys (Bornbolt, Goodnow and Cooney, 1994). Reliance upon any one assessment device, such as testing or self-assessment, will thus provide the possibility of systematic discrimination.

Teachers also come to the classroom with prior experiences, assumptions and values with which they are constructing understandings about what they see as ‘acceptable’ feminine or masculine behaviours, for themselves and their students, and bring gendered perspectives on the way knowledge itself is organized. For example, when Spear investigated teacher blind marking of school work, the same pieces of work that were arbitrarily labelled as done by either sex, she found that ‘...work attributed to a boy received higher mean ratings than the same work attributed to a girl.’ (1984, p. 373.)

Valerie Walkerdine goes further when she finds that ‘girls are considered lacking when they perform well, and boys are still possess something when they perform poorly’ (Walkerdine, 1984). Walkerdine found that the girls often performed at least as well as boys, but the teacher’s interpretation of their work was very different and unintentionally influenced by the student’s sex. Teachers’ gendered assumptions can be displayed in many ways such as capacity judgments are based on future promise for boys/men and performance for girls/women.

Thus three ways that assessment processes could privilege those masculine are: assuming equivalent out of school experiences of sexes; assuming assessment techniques are gender-neutral for knowledge demonstration; being blind to gendered expectations teachers bring to school. Recognising these interactions between knower and the known would suggest that more equitable assessment would use: negotiation of starting points for learning; multiple collection techniques; and a variety of assessors.

Multiplicity provides higher quality information (both hierarchical and multiplicity)

Introducing multiplicity and variety into assessment practices would break down some past hierarchical patterns where particular assessors of data are considered more important than others. For a variety of assessors, including the student, a peer, the teacher, or external authority, can provide moderated authenticity, where...
into science, and equity within its current construction, are considered worth pursuing. These women are looking through a liberal feminist lens, and are trying to skill up girls so they can do science, and assessment tasks, that are abstract, analytical, objective, rational, quantitative, competitive, and focus student effort on the end result, usually access into tertiary science courses.

Yet many other McClintock members, probably the majority, are actively working towards creating a new construction of science, at least in schools, in which

the values traditionally ascribed to women are given a positive and central place ... and there is a ... belief that the quality of life has priority over economics or efficiency or 'rational' planning ... [where] scientific activities ... reflect a balance with and not an exploitation of nature ... [and there is] an alteration of world view... from the analytical fragmentation of modern science to a holistic view in which social, ethical and moral considerations are unquestionably involved ... [where the] scientific community [is] based on co-operation, social accountability and accessibility and ... a respect for and equal valuations of different forms of knowledge, including the 'irrational' and the 'subjective'. (Manthorpe, 1982, p. 75)

Looking through this lens, the development of 'McClintock' approaches can be seen as radical feminist: celebrating the feminine and women's ways of relating to the world. The concept of a 'pedagogy for girls in science', has become the major focus of the work of the McClintock Collective, although the term pedagogy is not used. Whilst it is still unclear what feminist pedagogies might be, speculation on them is based on co-operation, social accountability and accessibility and ... a respect for and equal valuations of different forms of knowledge, including the 'irrational' and the 'subjective'.

(P.163) rather than a constant pursuit towards an ordering of the world based on dualisms - an ordering typically excluding or diminishing one side of the pair.

Respect for difference and multiplicity has become a guiding principle in the work of many McClintockers (a short-hand term used by members of the Collective). Multiple sites of political action; multiple forms of feminism, depending on the context and audience of the action; multiple goals for changing science and teaching, not always directed at increasing girls' participation in science; multiple pedagogical practices that may meet the needs of some girls and teachers some of the time; and multiple starting points for teachers interested in addressing gender issues in their classrooms. Thus McClintockers work on multiple fronts to deconstruct the restrictive and dysfunctional dualities of gender and of science. In that sense, a glimpse through the post-structural feminist lens, is available within the Collective's work.

However, more McClintock Collective work has centred on the 'development and dissemination' of alternative forms of non-discriminatory and empowering pedagogy, which may challenge schooling's complicity in reproducing gendered inequality' (Kenway and Modra, 1992, p. 141). The manifest ways of doing this, and the frames of reference of the viewers/actors are in a constant state of flux, and are characterised by their very fluidity and flexibility. Whilst very little direct analysis on the impact of this work has been undertaken, the publicly available outcomes of the new VCE are, at least to some degree, a measure of the success of the Collective's work.

Lessons from physics

Whilst the pattern of assessment practices outlined here are common to all the VCE subjects, physics is chosen as a barometer because of its extreme position as the science subject most closely aligned with most 'hegemonic masculinity'.

The opportunity that the introduction of the VCE created, provided two catalysts to speed up the implementation of McClintock pedagogies. Firstly, there was much professional development time for VCE teachers that the Collective organised. Secondly, the study designs incorporated assessment approaches, advocated by McClintockers, which necessitated changed classroom practice. Thus we were able to build in integrated, formative work requirements and assessment tasks, with explicit guidelines and criteria, which valued qualitative understanding, assessed all course goals, were set in a real world context, required a variety of types of data sources, and using different assessors. An example of new data sources illustrates: teachers who had never considered the 'visual' as a legitimate way of summarising learning were required to ask students to submit a poster on their research (in physics) or produce a concept map of ideas (in chemistry).
New teaching approaches

An example of one new teaching strategy used in the Nuclear Energy topic is 'a role play of people at a public meeting held to decide if a nuclear power plant should be built in the area.' (VCAB, 1991b, p. 15). Teachers now also use brainstorming, creative writing, jigsaw techniques for co-operative learning and other McClintock Collective teaching strategies (see Gianello, 1988; Lewis and Davies, 1988; Hildebrand, 1989). In the Year 11 topic, Nuclear Energy, those students who have chosen the context, 'Development of the bomb', can complete one work requirement with a briefing paper that responds to a focus statement such as: 'A group of students from your school is visiting Japan on a study tour, and one of the places they will go to is Hiroshima. They ask your physics class to prepare some information for them on the physics of the atomic bomb and its effect on people.' (VCAB, 1991b, p. 24). Many other focus statements for this briefing could also be negotiated with students. The quite radical shift in the conceptualisation of learning built by the work requirement system means it is no longer possible to teach physics in the old ways.

Assessment matching the course aims

Three of the ten aims in the new physics course are:

* become aware of physics as a particular way of knowing about the world which interacts with the setting, both social and personal, within which it is pursued;
* understand some of the practical applications of physics in present and past technologies, examining the social usefulness of such technologies as well as problems associated with them;
* develop the capacity and confidence to communicate their knowledge of physics effectively (VCAB, 1991a, p. 1-2)

A genuine attempt has been made to ensure that these are built into the assessment program: the first is highlighted in the work requirement where students have to build a file of changing ideas about light and matter; the second in the research project (CAT 3) and the third in both the extended investigation (CAT 1) and in CAT 3.

How has the assessment transformed the curriculum?

The actual physics curriculum statement is not markedly different from that in other similar level courses. It includes: light, heat, sound, electricity, electronics, motion, forces, structures, radioactivity and nuclear energy, etc. The difference here is that the integrated assessment processes created by the formative work requirement system necessitated new approaches to curriculum implementation. For example, in order to enable students to satisfactorily complete the prescribed work requirements, teachers were required to base all the central ideas in real world contexts. Also, as the work requirements demand a range of learning tasks, a broader repertoire of teaching strategies must be employed. The guidelines and criteria are explicit for each assessment task which now value qualitative understanding as well as quantitative ability; and, together, the assessment tasks cover all the course goals. These changes have meant that girls experience a physics curriculum which is very different from that of the old course.

Has this interruption altered outcomes for girls?

Looking through the first frame of reference, access and equity or liberal feminist, an analysis of girls' participation and achievement in physics can now be undertaken.

Participation of girls in physics

In 1970, 16 per cent of all girls and 49 per cent of all boys studied physics in Year 12. (See Note 1). By 1985 this had declined to 8.9 per cent of girls and 34.8 per cent of boys. This pattern continued for the final years of the old course. But for 1992, the first year of the full new course, the proportion of girls increased with 9.4 per cent of all girls choosing the new physics course while the corresponding figure for boys continued to decline to 28 per cent of all boys. By 1994, physics educators were concerned about the overall drop in physics participation, down to only 7.8 per cent of all girls and 26.9 per cent of boys; but this occurred at a time when the overall retention rate had increased, suggesting that the extra students who now stayed on to complete their VCE were not taking up physics. Many more boys than girls are studying physics, but the new course has made a small shift in the participation of girls given that 33 per cent of the cohort are now girls.

Achievement
VCE girls consistently outperformed boys. Looking through the liberal and Nash, 1994); and over the full range of forty-four studies within the physics curriculum. The data suggests that the assessment processes are now more equitable: girls are a very select group and it ought to be expected that they would outperform boys. Girls' mean scores were greater than those for boys in all but two Common Assessment Tasks - the two test CATs in chemistry - across all five science subjects (Cox and Nash, 1994); and over the full range of forty-four studies within the VCE girls consistently outperformed boys. Looking through the liberal feminist lens, it would seem that the changes to the VCE have been good for girls.

Why do girls achieve better grades with the VCE?

I recorded the following reactions, which were given by unidentified science teachers at the workshop on 'CATs and Sex' run by Peter Cox and Mary Nash (Cox and Nash, 1994), in response to: 'Why do girls do better, especially on the teacher-assessed CATs?'

- Girls choose physics, but boys get told to do it.
- Girls are a more select group, ability-wise.
- Girls are more careful and they read the assessment criteria.
- Girls ask the teacher what the criteria mean.
- Girls put in a draft and get feedback.
- Boys tend to be 'slap-dash' in their submissions.

Some of these answers reflect a view through the radical feminist lens, extolling the qualities which young women bring to their study of physics. One of the physics curriculum specialists (Firkin, 1993) places much of the credit for the shift in achievement profiles to the style of assessment which 'does seem to allow girls to show what they know and what they can do much better than the old system'. She also suggests that the first physics CAT, the internally assessed, extended investigation, has a confidence boosting effect for girls which continues into the test CATs. Firkin stresses that the course itself is no less rigorous or difficult.

- Girls apparently do '... respond better to science if more co-operative and interactive modes of learning ... are part of the pedagogy' (Tobias, 1990, p. 70). While there are grounds for celebration, a number of concerns remain.

Good for all girls?

Using the liberal feminist and the radical feminist lenses these changed assessment practices could be said to have been good for girls. But looking through the post-structural feminist lens we see the need to recognise the many differences within the broad category 'girls' and see that the interactions based on socio-economic factors and ethnicity, at least, challenge claims that all is well.

Some research, (e.g. Jones, 1989; Wyn, 1990) suggests that girls from lower socio-economic backgrounds actively resist pedagogical and assessment practices based on 'interpretation, exposition and independent work' (Jones, 1989, p. 29), precisely those practices which are now being given greater status. Teese et al (1994) clearly show that not all girls are doing well in the new VCE. Their data shows that both in enrolment patterns and achievement levels, girls from low socio-economic areas are well behind girls from areas with higher proportions of the population having tertiary qualifications and higher status occupations. For example, the 1992 participation rate of girls in physics for the working class north-west region of Melbourne was less than 7 per cent as compared to 12 per cent of VCE girls in the more affluent inner-eastern region. The differences are even more dramatic when achievement is considered: in 1992 less than 14 per cent of girls from the north-west region gained high grades (defined as in the top 20 per cent) whereas 39 per cent of girls who lived in the inner-eastern region did so (Teese et al, 1994).

Other evidence also suggests that some girls suffer pressure and trauma caused by teachers who expect girls to live up to the prejudices that their teachers have of them as belonging to particular ethnic categories. For example, Chinese-Australian girls who do not like, and are not highly competent at mathematics and science, are treated negatively by many teachers who believe they should fit an unrealistic stereotypical pattern (Fan, 1994).

The current curriculum and assessment changes are obviously not broad enough in scope to ensure effective pedagogies for all girls and all boys. A way forward might be to build in more negotiation of curriculum content and assessment processes, within broad parameters which maintain 'public credibility' (Gipps, 1994, p. 173) in the entire credentialling process, whilst allowing for differences in interest, learning styles and preferred assessment modes to cater for distinct groups of students.

Two steps forward, one step back

During 1994, a General Achievement Test (GAT), basically a conventional 'aptitude' test, was introduced and all teacher-assessed CATs are now moderated by it. The GAT creates several educational concerns: the presumption that external ('content-free') tests are more valid than teacher assessments; the validity of assuming that performance on qualitatively different types of tasks can be compared; the degree of inclusivity of the GAT for ethnicity, race and socio-economic factors as well as for gender; the performance differences on the GAT which are driven by the format chosen (predominantly multiple choice); and the removal of the powerful professional development experiences which occurred while teachers shared ideas.