THE TEACHING OF THINKING
John Edwards
Associate Professor
School of Education
James Cook University

Abstract
An increasing body of research suggests that schooling is not particularly effective. Despite this, there has been little change in educational practice. The direct teaching of thinking offers a promising addition to current mainstream practice.
This paper reports on a series of studies investigating the direct teaching of thinking. Such research raises a number of serious measurement issues which will be outlined and discussed. Application of this research in business and industrial settings is a recent development. Techniques for exploring and improving professional thinking will be presented.

Introduction
In the mid-1980s public and professional focus on teaching thinking skills became central in the United States. Reports such as 'A Nation at Risk' (1983), calls from employers (e.g., Chatham, 1982), predictions by futurists (e.g., Naisbitt, 1982), and results of polls and surveys, all pointed to the critical importance of thinking skills training for the citizens of the twenty-first century. Carolyn Hughes, in her foreword to the ASCD (U.S. Association for Supervision and Curriculum Development) publication Developing Minds (Costa, 1985), reported:
the most recent Gallup Poll asked teachers and the public to rate goals of education as to their importance. At the top of teacher ratings and near the top of public ratings was developing the ability to think creatively, objectively and analytically. When John Goodlad surveyed teachers, parents and students, he found that intellectual development of students was consistently identified as the most important goal of schooling. (p.ix)
This renewed public focus was set in a context of international concern for declining standards. 'Back to basics' movements are underway in many countries, including the United States, Britain and Australia. At the same time, there are those such as Costa (1985) who suggest a 'Forward to basics' movement is needed, with a set of new basics relevant to the twenty-first century rather than the Industrial Revolution. Chipman and Segal (1985) offer a more pragmatic view in this context:
Because explicit instruction in thinking and learning skills has received little attention in the schools, it is likely that large improvements are possible. It is much easier to improve instructional outcomes in a new or neglected area than to achieve significant improvements in instructional methods that have undergone decades or centuries of evolutionary improvement by trial and error. (p.3)
While there appears to be an acceptance of the central place of teaching thinking skills in the curriculum, arguments on how to go about it range widely. The two major camps at the extremes of the argument are those who advocate embedding thinking skills in the traditional disciplines and curriculum (e.g., Glaser, 1985) and those who advocate teaching thinking as
a new discipline in the curriculum (e.g., de Bono, 1976). At the same time, there is still little evidence of any widespread implementation of thinking skills programs at an educational system-wide level. The short-lived but large-scale attempt by Machado in Venezuela to make thinking a central curriculum component (Dominguez, 1985) is a rare exception.

Teaching thinking
The range of approaches taken to teaching thinking is broad, as can be seen from Figure 1.

FIGURE 1. Selected approaches to teaching thinking (Brandt, 1985, p.245).

<table>
<thead>
<tr>
<th>Authority</th>
<th>Approach</th>
<th>Materials</th>
<th>Type of thinking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authority</td>
<td>Approach</td>
<td>Materials</td>
<td>Type of thinking</td>
</tr>
<tr>
<td>Whimbey</td>
<td>'precise processing' (several) problem solving</td>
<td>Feuerstein 'mediated learning experiences'</td>
<td></td>
</tr>
<tr>
<td>Instrumental Enrichment</td>
<td>mental operations</td>
<td>Meeker (Guilford) assessment, practice</td>
<td></td>
</tr>
<tr>
<td>SOI</td>
<td>intelligence</td>
<td>Glade (Upton) teaching</td>
<td></td>
</tr>
</tbody>
</table>
| Strategic Reasoning | thinking skills | Beyer 'direct'
| -thinking skills | Paul (Ennis) 'dialectical'
| -critical thinking | Lipman discussion |
| Philosophy for Children | reasoning | de Bono heuristics |
| CoRT Thinking | perception | Perkins 'design' Odysseyc |
| thinking | |

Some approaches involve the use of specially published teaching thinking materials, others involve strategies to be used with existing materials. Some approaches involve the direct teaching of thinking strategies, others involve stimulating students to develop their thinking skills over time through engaging in activities such as dialectical discussions. Some approaches involve the teaching of thinking as a discrete activity, others teach thinking through the academic disciplines. It is this last dichotomy that Polson and Jeffries (1985) use when reviewing programs for instruction in general problem-solving and other cognitive skills. They suggest that there are two major alternative frameworks: information processing and divergent production. While the validity of this categorisation is problematic, Polson and Jeffries (1985) do identify one of the central questions in examining approaches to teaching thinking: What is the function of domain-specific knowledge in perception and comprehension? All information-processing analyses of these processes conclude that both require extensive domain-specific information ... Writers in the divergent-production paradigm implicitly or explicitly assume that domain-independent perceptual and comprehension processes can be taught, and it is possible to teach students general, domain-independent problem solving skills. (Polson and Jeffries, 1985, p.428)
The neo-Piagetian research of Case (1985) supports the learning of discipline-specific schemata that reduce cognitive load. This approach, based on the concept of a fixed mental processing space, suggests that general thinking skills of great power do not exist. Furthermore, research on skilled memory (Chase and Ericsson, 1981) and on problem-solving in physics (Larkin, 1985) emphasises the central role of domain-specific knowledge. Others who support the domain-specific perspective include Glaser (1985), who sees learning and thinking skills being acquired through the teaching of the content and concepts of familiar knowledge domains, and Greeno (1980), who argues that recent problem-solving research has allowed clear identification of the specific knowledge required to solve specific problems in a way that overshadows the value of supposed general problem-solving strategies.

Simon (1980), however, sees a role for both approaches. The evidence from close examination of AI (artificial intelligence) programs that perform professional level tasks, and the psychological evidence from human transfer experiments, indicate both that powerful general methods do exist and that they can be taught in such a way that they can be used in new domains where they are relevant. (p.86) On the other hand, Edward de Bono (1976) and Feuerstein et al (1985), producers of the two thinking skills programs most widely used around the world (CoRT Thinking and Instrumental Enrichment), argue strongly for the teaching of thinking processes free from traditional content for psychological, sociological and pedagogical reasons.

The Instrumental Enrichment (IE) program is designed to provide mediated learning experiences for the learner. Feuerstein et al (1985) see these as critical to normal human cognitive development and criticise Piaget's ignoring of their role. They are seen as providing the learner with strategies and sets that enable maximum benefit to be gained from exposure to stimuli. Feuerstein rejects the notion of immutable inborn abilities as primary determiners of intellectual performance. He believes in cognitive modifiability whereby 'structural changes' in the intellectual state of the organism can be achieved through the provision of appropriate mediated learning experiences. These can change an individual's cognitive structure and capacity to learn. Exercises in the IE program are designed to nurture proper learning sets and systematic data gathering behaviour, to induce skills in comparative analysis thereby improving relational insights, and to remove attitudinal inhibitions.

Feuerstein attaches considerable importance to the relatively "content-free" nature of the materials used in his program. The intent in the design of these materials was to make them as free of curricular content as possible. The rationale for this approach alludes to several resistances to the use of content material, resistances arising from the learner, the teacher, the material itself, and the student's history of failure with such material. Moreover, the use of content-free material is assumed to help keep the student's attention focussed on the objective of correcting specific deficient functions rather than on the content itself, a focus
that is presumed to facilitate realization of the program's general goal. (Nickerson et al, 1985, p.153)

The CoRT (Cognitive Research Trust) program is designed to teach children a set of thinking tools, or executive thinking strategies, which they can use to overcome the limitations of their existing and relatively fixed thinking patterns. De Bono (1985) also argues for a 'content-free’ approach: Attending to content distracts from attending to the thinking tools being used. Theory predicts this outcome: you cannot build metapatterns at one level and experience patterns on another level at the same time. Experience backs up this expectation. Wherever there has been an attempt to teach thinking skills and content together, the training in thinking seems to be weaker than when these skills are taught in isolation. (p.206)

De Bono acknowledges that subject idioms exist. Nevertheless he claims that it is possible to establish both habits of mind and specific thinking techniques that can be applied in any subject area. For example, he cites the willingness to look for alternatives as a generalisable thinking habit, and deliberate provocation as a technique that can be applied to generate ideas in any situation.

This problem of domain-specific versus generalisable thinking skills is an issue of great practical significance when looking at where to focus in teaching thinking. It is fundamental therefore to ask whether generalisable thinking skills exist and whether they can be taught directly to students, independent of existing academic discipline domains. This question was the focus for the series of studies reported in this paper, which explore the implementation of de Bono's CoRT Thinking program as a generalisable thinking skills program. The CoRT thinking skills program was selected to fill this role for a number of reasons. Firstly, it has been widely used for almost twenty years, and there is great enthusiasm for the program from many users. Secondly, this researcher had practical experience with the program (Edwards and Baldauf, 1983), and preliminary data, both quantitative and anecdotal, suggested that the program did have measurable positive effects on students’ thinking skills. Thirdly, the program is easy to implement. The materials are relatively low-cost, no special teacher-training is required, and the program has an appeal to teachers. Finally, the program is clearly recognised as an heuristics-based program which claims to teach generalisable thinking skills.

Evaluation of thinking skills programs

There have been three thorough reviews of evaluations of thinking skills programs, by Nickerson, Perkins and Smith (1985), Segal, Chipman and Glaser (1985), and Mansfield, Busse and Krepelka (1978). Evaluative studies of thinking skills programs have not been given a high priority by researchers, as evidenced by the following comments: As our discussion of educational programs shows, current understanding is generally insufficient to permit convincing evaluation of programs with respect to their claim to develop complex cognitive skills. (Chipman and Segal, 1985, p.6)
Solid evaluative data on the effectiveness of programs for teaching thinking are scarce. (Nickerson et al, 1985, p.315) Everyone thinks, therefore a course of thinking must be expected to show that it has improved thinking. Otherwise what is the point of devoting time to it? This is self-evident. But the trouble arises when we find that testing thinking is extraordinarily difficult and beset with pitfalls. (de Bono, 1976, p.201)

Mansfield, Busse and Krepelka (1978), following their review of the effectiveness of creativity training programs, concluded that most evaluation studies seemed to support the view that creativity can be taught. Just as importantly, they identified serious conceptual and methodological problems with most evaluation studies. A number of the problems they raised are also mentioned by Nickerson et al. (1985), and by a range of authors in Segal, Chipman and Glaser (1985). Many of the issues are well-known, such as failure to control for Hawthorne effect, failure to assign students or classes randomly to treatment conditions, and a range of obstacles to controlled scientific experimentation in educational contexts. The list below is an amalgamation of the issues raised by these reviews which highlights the major evaluation issues which need to be considered when researching the teaching of thinking.

1. There is a major problem with selecting appropriate performance measures in this area. Nickerson et al. (1985) report little agreement on what is important. A wide range of measures is used across different studies, so comparisons are difficult. They wonder to what extent attempts to teach thinking are constrained by the available tests. Not only may the test take a narrow view of a particular aspect of thinking, but the test may be inaccurate. Nickerson et al. (1985) see a challenge to select or develop tests more finely tuned to the objectives of programs, and they recommend that this should be an integral part of program development and presentation. Also, some of the skills acquired by students may be very different to those set out in program objectives, so evaluations should include some general skill component. Sometimes these indirect effects may turn out to be the more important. The testing program should also include an assessment of the negative effects in other areas which may result from the gains or emphasis in one narrow area.

2. Ideally one wants information that focuses on generalisability and how lasting any effects are, therefore there is a need to assess outside classrooms and over substantial time periods. The difference between long-term effects and short-term effects which may be due to changes in motivation and response-set is important.

3. With small sample sizes, teacher effects are hopelessly confounded with treatment effects. With larger samples, monitoring of treatment effects is very difficult. Detailed descriptions of differences between teachers and between treatments need to be supplied.

4. The state of theory relating to intellectual competence is still too primitive, in general, to warrant giving great weight to the underlying theory in assessing programs.
5. While there may be difficulties with formal evaluation attempts, it is a fact reaffirmed in numerous contexts that wishful thinking distorts casual observations (Nickerson et al., 1985).

6. The inclusion in programs of tasks very similar to the criterion tests is often a problem. Similarly there can be an influence due to student awareness of the types of responses required or simply due to greater persistence.

7. There is a need to control for the quality of teaching. Too often the treatment teachers are the volunteers. There can be similar problems with providing appropriate control group treatments.

8. There are often a range of problems with statistical analysis, such as: the use of individual scores rather than classroom means as the unit of analysis; the use of several univariate analyses rather than a single multivariate analysis; and focussing on statistical significance to the exclusion of consideration of practical significance through studying also the magnitude of differences.

9. A close analysis of the goals of programs is critical, to see whether they have any real worth in broader contexts. That a program does what it claims to do, does not make the program a worthwhile program.

10. There is sometimes a tendency towards premature closure. Nickerson et al. (1985) suggest that failure to find a difference between experimental and control groups may be due to noisiness of data or inappropriateness of measurement techniques. It does not necessarily demonstrate that no difference exists. On the other hand there is sometimes the tendency to assume an unwarranted link between improved performance on tests and real-life abilities or accomplishments.

These general problems cover the major weaknesses commonly found in evaluation studies in the thinking skills area.

The CoRT Program and its evaluation

The CoRT program was designed by Dr Edward de Bono and was first published in 1973. The name CoRT comes from the Cognitive Research Trust which de Bono established at Cambridge, England. CoRT consists of sixty lessons divided into six sections, CoRT-1 to CoRT-VI, of ten lessons each. The program offers instruction in a selection of specific thinking skills. De Bono (1983) summarised four levels of objectives for CoRT:

1. That there is an area in the curriculum where thinking is treated directly in its own right.

2. That pupils should come to regard thinking as a skill that can be improved by attention, learning practice.

3. That pupils should come to regard themselves as thinkers.

4. That pupils should acquire a set of transferable thinking tools which they can carry to other situations. (p.117)

It is recommended that CoRT lessons are taught with a mixture of direct instruction by the teacher, student group work, class discussion, individual work and homework projects.

Over the last six years a series of detailed investigations of the effects of CoRT in classroom practice have been completed by Edwards and Baldauf (1987) and Edwards (1991 and 1992). These studies have involved teaching
the CoRT-1 program to groups of Grade 7 (approximately 12 years old) children. The CoRT-1 program is a set of ten lessons aimed at broadening student thinking through the teaching of seven specific thinking skills. Each lesson takes about forty-five minutes, resulting in a total of approximately eight hours of instruction.

In the first study (Edwards & Baldauf, 1987), sixty-seven Grade 7 students completed the CoRT-1 program over a four-week period. They were tested using a broad range of measures before starting CoRT-1 and again eleven weeks after completing the program. As in previous studies, the anecdotal evidence gathered strongly supported the value of the CoRT-1 materials for teaching thinking. Interviews with the children eleven weeks after the program revealed positive attitudes to the program and its effect on their thinking. The view was supported in general by feedback from parents. The students also showed statistically significant gains in: scholastic aptitude, as measured by the Otis-Lennon (1982) School Ability Test (OLSAT) - Intermediate Form R; self concept as a learner as measured by Waetjen's (1967) Self Concept as a Learner (SCAL) Scale; and flexibility and originality of thinking as measured by the Torrance (1984) Test of Creative Thinking, Verbal Form A.

This study revealed serious measurement problems, particularly in the area of self concept. The Self Concept as a Learner Scale was shown to be unreliable (Baldauf, Edwards & Matthews, 1985). With respect to essays it was found that the form of essay written by the student had a dominating effect on student scores. Therefore, unless one can ensure that each student is writing in the same genre, scores from essays are likely to be misleading and invalid.

As a result of the problems with the SCAL Scale, four new scales were developed for use in this area of research. These are described in detail in Edwards (1988). The four scales are: The Self Concept as a Thinker (SCAT) Scale, a broad self-report measure of student perceptions of their own thinking skills; the Teacher Assessment of Student Thinking (TAST) Scale, a teacher's version of SCAT, providing a cross-validation of the students' self-ratings on SCAT; and two parallel forms of the Thinking Approaches Questionnaire (TAQ) Scale where students rate themselves on the thinking approaches covered by the CoRT-1 program.

In the second study (Edwards, 1992), 202 Grade 7 students in two schools were used. Two classes in each school formed the treatment group of 115 students, and two classes in one school and one in the other formed a control group of 87 students. The CoRT-1 treatment took five weeks and the students were pre-tested, post-tested and given a delayed post-test four weeks later.

When compared with the control group, the students who had done CoRT-1 showed improved performance on: The OLSAT (Deviation IQ); the fluency, flexibility and originality scales of the Torrance Test; the TAQ Scale; and TAST Scale. Of these gains the improvements in scholastic aptitude (Deviation IQ), flexibility and originality of thinking, and thinking approaches (TAQ) were statistically significant. Of these improvements, only the improvement on TAQ was maintained from post test to delayed post
test at a statistically significant level. There was no relative improvement in self concept as a thinker as measured by SCAT. When student performance on the normal school-based mid-year and end-of-year examinations were compared, interesting trends emerged. The CoRT-trained students showed a statistically significant overall improvement in academic performance compared to the control group. There were statistically significant improvements in social science performance, and particularly language arts performance. There were no changes in performance in mathematics or science. These results support de Bono's (1976, p.141) assertion that CoRT improves performance in language arts but that it is difficult to produce gains in knowledge-bound subjects. Both of the studies reported here reveal impressive short-term gains associated with an eight-hour one-shot treatment with the CoRT-1 program. A longitudinal study has been commenced to explore the effects of the CoRT program in more detail. A group of twenty-six Grade 7 students were taught the full sixty lessons of the CoRT program in one year. The CoRT thinking skills were also infused throughout the curriculum. Two control groups were used in the study, a no-treatment control group and a group that received only the CoRT-1 program in the second half of the year. Data analysis is in progress at present. Independent testing by the school, using standardised tests on the CoRT trained students, revealed large gains in learning abilities, study skills, mathematics comprehension, reading comprehension, and vocabulary skills. These students are being followed through their junior high school years and their performance is being compared to that of their classmates and of the original control comparison groups. As part of this study, detailed observations (Clayton & Edwards, 1989) of teacher behaviours while teaching CoRT were made using Ryan's (1960) Classroom Observation Record. Teaching the full CoRT program in one year was an extremely demanding exercise for the teacher. By the end of the year the teacher's pedagogy was characterised by being more democratic, optimistic, adaptable, integrated, responsive, understanding and broad than it had been earlier in the year. She also became more fair, attractive, steady, poised, systematic and confident (Clayton & Edwards, 1989).

Melchior and Edwards (1989) reported on the implementation of the CoRT-1 program in a junior high school in New York. Again anecdotal data for the success of the CoRT-1 implementation were strong. Over three years students spoke positively of the value of the CoRT-1 training, as did teachers and parents. Through application of the SCAT Scale a complex pattern of results emerged. In general, there were improvements in self concept as a thinker, particularly at Grade 7 and 8 level. Students saw themselves as being more interested in ideas, thinking more broadly, being able to tell which ideas are important, having more important ideas, using their thinking in real life, and being better problem solvers. At the same time, the students acknowledged a lack of concentration, an unwillingness to ask questions when they do not understand, a tendency not to think about where their own thinking is heading, and trouble with making decisions. These shifts in both directions raise the interesting issue of what shifts
one would expect when students are initially made more aware of their own thinking. I could be that initial drops are to be expected when weakly-based self perceptions are subjected to close self examination.

Measurement issues currently under investigation
Of the many complex data gathering, data analysis, and data interpretation issues that we are aware of in this field, many of which have been outlined earlier in this paper, the most critical for us is that many of the tests in common use do not seem to cognitively engage children. If this is the case, then what do their test scores mean?

In an attempt to develop measures of student thinking which cognitively engage children a research team at James Cook University began by exploring what it is that children commonly think about. Two hundred and fifty children were interviewed in small groups with follow-up individual interviews of representative children. This led to the development of a 148 item questionnaire, the "What I Think About" Scale which was administered to 1200 children. The questionnaire survey confirmed that the Scale did contain topics that occupied the thinking of late primary and early secondary school students. Factor analyses revealed major areas about which the children thought: environmental issues, social issues, entertainment, sport, and national and personal finances. (Alloway, Bond and Edwards, 1992)

Specific topics with the highest ratings were: what's on the news, family relationships, and destruction of the environment. Strong response rates were also elicited by the following topics: rock music, clothes, TV, friends, holidays, food, and doing well at school. Response patterns were stable across subsamples based on age, sex, and level and type of school attended.

On the basis of the interview data and the survey results a set of measures to allow the profiling of students’ thinking was developed. Sixteen measures were trialled with ten grade 8 children in a local high school. The measures were based on the ideas of a group of practitioners and theorists in the areas of creativity, thinking, and the teaching of thinking: De Bono, Feuerstein, Piaget, Torrance, and information processing theorists such as Simon and Sternberg.

Trialling of the measures was done with individual children, using the stimulated recall interview technique (Marland and Edwards, 1982). While working on any one measure the child was video-taped with two cameras - one focussed on the child from the front to reveal any talking, gesturing, facial expressions or other actions, the other focussed from above the child on the writing pad used by the child to record his or her response. These two images were then mixed to produce a split-screen image, and the recording formed the basis for a stimulated recall interview with the child. These interviews were tape recorded and transcribed. The focus of the interviews was on revealing the thinking of the child while working on the thinking measure and on revealing the sources of ideas generated in response to the thinking measure.
It became immediately clear to the research team that while our measures did, in most cases, cognitively engage the children, the original aim of developing measures to be administered to groups of children and then taken away and ‘scored’ warranted close reassessment. Comparisons between our interpretations of students’ written responses and the students’ reports on their thinking processes and the sources of their ideas revealed that interviews amplified the written test results in ways that could not be incorporated into a standardized written test format. As a result we are now working on the development of approaches to revealing students' thinking based on stimulated recall interviews. This has obvious implications both for the use of standardized tests, and moderating tests such as the new Queensland core skills test, in schools. Procedures for developing such measures, and procedures for incorporating them into educational practices in schools are currently being explored. This work is likely to continue for a significant number of years.

Professional thinking
A number of studies on exploring and improving professional thinking are currently in progress at James Cook University. These studies use the stimulated recall interview technique to reveal professional thinking, and a range of approaches to improve professional thinking are being trialled. These are based on Butler's (1992) model of human action and change. The model proposes a central process of reflection, defined as an active personal process that is a dialogue between the inner self and the outside social context that enriches the self and enhances human action (Butler, 1992). Within the context of the inner self are the domains of personal practical knowledge, a store of personal knowledge that accrues over time and is enriched and renewed by a constant checking of the consequences and efficacy of the practice (Butler, 1992); and world views, the individual's subjective, historical, contextual way of viewing the world. Within the outside social context are the domains of public knowledge, all that the individual "is offered in the form of theories, formal knowledge, policy directives, research results, hints and folk-lore, community expectations" (Butler, 1992) and so on; and professional practice, with its ramifications for the individual, and for the people with whom the individual interacts.

Conclusion
The teaching of thinking is a major current focus in education in the U.S., and increasingly in U.S. business. More recently, the education sector in Australia has also paid increasing attention to this area. However the effectiveness of most programs or approaches designed for teaching thinking has not been carefully evaluated by independent researchers. Any teacher or researcher interested in research in this area is faced with a complex task for which there are few valid and/or reliable performance measures. This paper has looked at the major issues raised by reviews of research on teaching thinking, outlined the potential of thinking skills programs for use in schools, and described some of the measurement and industrial
application issues currently under investigation. These are areas which have been largely ignored by the educational research community in Australia.

References
Council for Educational Research, 87-106.


The direct teaching of thinking skills: CoRT-1, an evaluative case-study. ng, facial expressions or other


low-cost, no special teacher t, and Edwards (1992). The early from post-test to delayed post- (Edwards, 1991)t at grade 6, 7 and 8 levels ces human action (Butler, 1992)

A common strategy used in exploring the personal practical knowledge of professional people is to have them reflect on, and clarify, their understanding of their own thinking, and the ways they deal with change. Sharing such understandings is generally a powerful learning experience. The public knowledge sought by professionals in an attempt to improve their performance commonly includes work on developing thinking skills. Self-improvement action plans emerge from the blending of work on public knowledge, personal practical knowledge, world views, and current practice - revealed principally through reflection, journal keeping, discussion and action. While considerable success has been experienced with these approaches in consulting situations, detailed research has only recently begun. Once again a number of complex research design and measurement issues have emerged. These will be reported in future publications.


Waetjen, W.B. (1967) Self-Concept as a Learner Scale. Cleveland State University:

W.B. Waetjen.