

# EVALUATING THE FACTORS THAT FACILITATE A DEEP UNDERSTANDING OF DATA ANALYSIS

Oliver K. Burmeister  
Centre for Information Systems Research  
Swinburne University of Technology

## ABSTRACT

Ideally the product of tertiary informatic study is more than a qualification, it is a rewarding experience of learning in a discipline area. It should build a desire for a deeper understanding and lead to fruitful research both personally and for the benefit of the wider community. This paper asks: 'What are the factors that lead to this type of quality (deep) learning in data analysis?' In the study reported in this paper, students whose general approach to learning was achieving or surface oriented adopted a deep approach when the context encouraged it. An overseas study found a decline in deep learning at this stage of a tertiary program; the contention of this paper is that the opposite of this expected outcome was achieved due to the enhanced learning environment. Though only 15.1% of students involved in this study were deep learners, the data analysis instructional context resulted in 38.8% of students achieving deep learning outcomes. Other factors discovered that contributed to deep learning outcomes were an increase in the intrinsic motivation of students to study the domain area; their prior knowledge of informatics; assessment that sought an integrated, developed yet comprehensive understanding of analytical concepts and processes; and, their learning preferences. The preferences of deep learning students are analyzed in comparison to another such study of professionals in informatics, examining commonalities and differences between this and the wider professional study.

## INTRODUCTION

The contribution of this paper is its evaluation of those factors most likely to lead to deep learning in the data analysis domain. The paper demonstrates the very important place of data analysis in the process of producing business information systems and the inherent need for a critical understanding of the design task. This need is not solely for computing professionals, but extends to all business professionals who are involved in the analysis and design of business information systems. It is therefore necessary to students majoring and minoring in business computing.

The main question driving this study has been: 'What are the factors that contribute to deep learning in data analysis students?'

This paper begins with a section defining the terms used. This is done in recognition of the fact that not all readers will be familiar with both data analysis and learning theory terminology. It then develops the argument that data analysis teaching is of primary importance to both tertiary students majoring in computing and those business students who take minors in computing. The paper then examines the design and methodology used for the case study. Finally the results are detailed and analyzed.

## DEFINITION OF TERMS

**Action learning** is 'the development of self by the mutual support of equals' (Revans, 1982, p 633). It is a social process by which those partaking in the study learn with and from one another. The intention is to cause one another to reexamine afresh ideas they might otherwise have continued to take for granted. Action learning is useful in helping people learn more effectively from their experience. However it is more than learning by doing, it involves 'review, reflection, rethinking and reinterpretation of taken-for-granted knowledge and experience' (Ballantyne, Bruce & Packer, Nov 1993, p 2).

**Data** "A representation of facts or ideas in a formalized manner capable of being communicated by some process." (IFIP, 1966). Data is distinguished from **information**, which is data that has been *processed* in some manner to make it meaningful to one or more end-users.

**Data analysis** A disciplined approach to analyzing the meaning and properties of the **data elements** in existing clerical forms and computer files, independently from the **systems** that produce and use this **data**. (Bold as in original, Martin, 1990, p 456)

**Data model** "A logical map of **data** which represents the inherent properties of the data independently of software, hardware, or machine performance considerations." (Bold as in original, Martin, 1990, p 457)

**Deep learning** This concept grows out of an understanding that not all forms of learning are the same. Some forms of learning are better than others. Deep learning is considered the best. Whilst the student's approach to learning is important, regardless of that approach all learners can "deep" learn if the teaching context provides for this. Biggs describes deep learning as involving a deliberate search for meaning, an analytical approach; it is an approach to learning that leads to an understanding of the structural complexity of a task. Other types of learning include **surface learning** and **achieving learning**. Surface learning reflects a student's decision to get by with minimal effort and involvement. Such students learn facts without necessarily understanding their meaning or how to apply them. Achieving learning reflects students who maximize their chances of status by playing the competitive game (Biggs, 1991). Deep learning is generally regarded as synonymous with intrinsic motivation (Davis & Bostrom, 1994), anchored learning or conditionalising learning (Brown, 1993), contextual learning (Lawson, 1991; Kirby 1991) and the highest form of learning (Collis & Biggs, 1986a & b).

**Design Task** This is the modelling of the new system that is required by the user or enterprise. It usually involves preparing a logical design first, followed by a physical design of the new system. (McLeod, 1994)

Some distinguish these instead as two different design tasks, the first is Systems Design and the second is Construction Design (Olle, 1993).

**Enterprise Data Model Approach** Essentially similar to data modelling, it shifts emphasis away from the user's data resource, to that of the business organization. The underlying premise is that if "all the firm's data is stored in a database, there is no limit to the information support the firm's computer-based systems can provide." (McLeod, 1994, p 469)

**Information engineering** A "data-driven, but process sensitive technique that is applied to the organization as a whole...Although the technique suggests a balance between data- and process-oriented methods, it is clearly data driven; data models are built first and process models are built later." (Whitten, Bentley & Barlow, 1994, p 154)

**Process modelling** This models the flow of data through business processes in the new system being designed. A common means of documenting the process model is the *Data Flow Diagram*.

**SOLO taxonomy** The Structure Of Learning Outcomes (SOLO) taxonomy is a system of classifying learning outcomes according to five levels, from incompetence (level 1) to expertise (level 5). The top two levels are associated with deep learning (Biggs, 1991).

### THE PRIMACY OF DATA ANALYSIS

Though numerous changes have been proposed, the design task remains essentially a two step process (Olle, 1993). It involves analyzing the data requirements that take place within the business system being investigated. It involves analyzing the procedures or functions of that system. Even with the trend toward information engineering, data modelling is still the first and primary step (Whitten, Bentley & Barlow, 1994). Similarly, the emerging Enterprise Data Model Approach (McLeod, 1994), though it is a development of the traditional data model, still recognizes the singular need for a detailed understanding of the data requirements of a business enterprise. In the words of James Martin: "Systems have proven to be much easier to build and much cheaper to maintain when thorough data modeling has been completed" (Martin, 1989, p 23). Understanding of the primacy of data analysis over other aspects of the design task was explored in an action learning case study of the way data analysis is taught at the School of Information Systems, Swinburne University of Technology.

The design task is part of the System Development Life Cycle (SDLC). Traditionally this "cycle" has been seen to consist of numerous stages, which are variously described but generally synonymous with first analysis, then design, through to implementation and support. Recently the cycle stages have undergone a change. With the advent of information engineering and enterprise data modelling, a further stage has been added to the start of the SDLC. This is the planning stage, which seeks to establish strategic priorities over which applications get developed within an organization (Whitten, Bentley & Barlow, 1994). Nonetheless, the design task remains and 'if anything' has grown in importance, because increasingly the data requirements of the whole organization require modelling.

The design task consists of two main phases - analysis of (and modelling) the data and analysis of (and modelling) the processes of the system to be developed. One might argue that process modelling and data modelling are both necessary and therefore can be carried out in either order with equal success. Yet whilst it is certainly true that both are necessary, until the trend toward integrating the two into a single methodology succeeds, modern approaches such as in information engineering, suggest the accepted primacy of the two is with data modelling (Whitten, Bentley & Barlow, 1994; McLeod, Jr., 1994). The clearer one's understanding of the data the better will be the resultant process model.

This approach to system development is reflected in the structure of business computing subjects within the Bachelor of Business, Swinburne University of Technology. The pivotal subject is Data Analysis and Design (DAD) which is the basis of the case study the rest of this paper is built around. It is the first post core subject students must take for all 3 majors, and 2 of the recommended 3 minors in business computing; it is one of 3 suggested electives for the third minor (Swinburne University of Technology, 1995). In this subject students are taught how to analyze and model data from different sources, forms, questionnaires, interviews and written descriptions. The emphasis is on gaining an understanding of the problem in its business context and on the management of the organization's data to ensure that the information produced by the database system is relevant and accurate. Students are taught to use data analysis methods to produce a conceptual data schema and then to test their models by implementing them using Oracle's SQLPLUS. Because of its pivotal position, it is imperative that students learn the content and processes they are taught in a way that can be built upon by future subjects in business computing.

Herein lies a potential dilemma for informatic teaching in this area. Having established the primacy of data analysis and structured the course accordingly, how does one effect learning that lasts? Many lecturers intuitively recognize that whilst students will study hard and pass a subject, some of them remember little of value a few months later. On the other hand, a few are able to competently work with the subject matter taught many months and even years later. How does one encourage the latter type of learning?

### THE CHALLENGE OF TEACHING CRITICAL THINKING

Ideally the product of tertiary study is more than a qualification. Such study should be a rewarding experience of learning in a discipline area. It should build a desire for a deeper understanding and lead to fruitful research both personally and for the benefit of the wider community. In secondary school teaching can be deterministic, with right and wrong answers. However, tertiary study goes beyond this. Certainly in the study of data analysis, students must become aware of the fact that there is not necessarily one right answer, and that one must design

alternative solutions, from which to choose the best design for the given situation (Moody & Shanks, 1994). What then are the factors that lead to this type of quality learning in data analysis? Wouldn't it be interesting to measure the extent to which tertiary study actually produces the above as an end result? What is the actual "bottom line" in terms of students' ability to think analytically? Do our existing assessment criteria measure this type of thinking ability adequately? To what extent is this an innate approach to study or one developed in the learning environment? The remainder of this paper is an action learning approach, that seeks to address these questions using Data Analysis and Design, semester 1, 1994, as a case study.

## DESIGN: MEASUREMENT AGAINST A STANDARD

The primary means of evaluating the extent to which deep learning was facilitated was the measurement of student learning against the SOLO taxonomy, an internationally accepted standard for grading learning (Collis & Biggs, 1986a & b; Biggs, 1979 & 1991). It operates in two distinct modes, open and closed, both of which were used in this study.

The closed mode of the SOLO taxonomy was used to develop the final exam (worth 80% of the assessment). This mode guides the development of leveled questions each successively requiring more critical responses from the student.

The open mode of the SOLO taxonomy was used to grade student responses according to 5 levels of thinking; levels 4 and 5 generally require student responses of such structural complexity as could only be achieved if they have engaged in deep learning strategies (Kirby, 1991).

Though the SOLO taxonomy was the primary standard of evaluation, the Study Process Questionnaire (SPQ) (Australian Council for Educational Research, 1985) and the Myers-Briggs Personality Type Indicator questionnaire (MBTI) (Myers & Myers, 1987) were used to seek out possible correlation that would explain the facilitation or otherwise of deep learning outcomes.

## METHODOLOGY

An action learning methodology was adopted for this study. In the context of this study, the teaching panel were challenged to rethink the overall purpose and teaching strategy of the subject Data Analysis and Design (DAD). Several means of inquiry were employed to examine the factors contributing to deep learning in DAD students. These were:

1. *Teaching processes and assessment design.* Two factors that have been shown to influence deep learning are the teaching environment and the type of assessment demands placed upon students. The author convened a teaching panel of 5 other full time lecturing staff meeting frequently at first, less often as the semester progressed. We met for two main purposes that affected this study. Firstly to examine the assessment - how could the closed mode of the SOLO taxonomy be used to redesign the way we assessed our students, thereby better grading their assessment responses according to their level of critical thinking?

Secondly to review the teaching processes in order to redesign them to enhance critical thinking in our students. Through reading about research into deep learning in other contexts, the author gained an insight into a number of things the DAD teaching staff could do to enable students to engage in this type of learning.

2. *Final exam assessment.* All 85 final exam papers were scored according to the open mode of the SOLO taxonomy. The exam covered four main areas, mastery of SQL, normalization by decomposition, functional dependency diagrams and entity relationship modelling. Of these only the last three proved suitable for an assessment of deep learning. The consensus of the teaching panel had been that a multistructural understanding (SOLO level 3) was a sufficient mastery level for students to reach in SQL. Given it is possible for good surface learners to reach that stage of understanding, the results of that section of the exam paper are not included in the results reported below.

3. *Approach to learning.* Some students are by nature/background more inclined to deep learn than others. So the quest here was three-fold. Firstly, how many students adopted a deep learning approach to begin with. Secondly, Could the learning approach of data analysis students be deepened, and if so how? Lastly, assuming that their learning approach can be deepened, how many students could be shifted toward adopting a deep learning approach in the semester under study? The second was encouraged by devoting a small proportion of lecture and tutorial time to teach students deeper ways to go about their learning, and by inviting an educational psychologist along to 2 lectures to talk to students about differing learning motivations, preferences and strategies. The first and last were measured according to the Study Process Questionnaire (SPQ). The SPQ was used before the semester study began and after the final exam had been completed.

Considerable research into tertiary student learning approaches has been conducted (for example, Marton & Saljo, 1976; Biggs, 1978); however a literature review reveals that nothing has been done specifically with data analysis. The SPQ is an adaptation for the tertiary sector of a survey that has been successfully used for this purpose in secondary schools across 5 continents. However, some difficulties were experienced in measuring the approach to learning. The Study Process Questionnaire asks students to respond to the questions that are subject dependent as for their "favorite" subject - this obviously need not be Data Analysis and Design. Moreover, the after measure was conducted immediately following their final morning exam in this subject. This proved unwise because many students had an afternoon exam the same day and chose not to stay after the exam to complete the questionnaire; only 43 of the 85 who sat the exam filled out the questionnaire, compared with 79 who took the before measure. In addition, 13 of the students who took the after measure did not give their names and 2 did not take the before measure, whilst some who took the before measure had since dropped out of the subject. All this made comparisons between the before and after measures complicated, let alone comparisons to the SOLO assessed students responses. Nonetheless in comparison to overseas trials of this method (Gow, Kember & Cooper, 1994), the results reported here are encouraging for informatics educators.

4. *Learning preferences.* Could it be that certain personality types are more suited to data analysis study than others? If so, might there be a correlation between personality type and learning outcome? In order to

build a profile of the personality types that predominated in our students the Myers-Briggs Personality Type Indicator questionnaire (MBTI) was used as a measurement instrument. The MBTI contains four separate indices each reflecting one of four preferences concerning a person's perception and judgments. The tool enables one to describe sixteen MBTI types of 'preferred or habitual patterns of intellectual functioning, information processing and the formation of ideas and judgments' (Jensen, 1987, p 182).

## TEACHING PROCESSES & ASSESSMENT DESIGN

Working on the premise that deep learning depends in large part on the manner and environment in which a subject is taught (Ramsden, 1984), it was necessary to review the teaching processes in Data Analysis and Design. As a result a number of changes were adopted. These included introducing the concepts of learning maps and chunking to lectures. They also included trialing the 4Mat system (McCarthy, 1987) and direct education of students about different methods of learning, with encouragement given for them to become deep learners. Also trialed was the idea of communicating with students using their preferred modalities (mainly visual, auditory and kinesthetic). Experimenting with these various teaching methods was a lot of fun for both students and staff.

Amongst the problems experienced was the fact that not all these methods were consistently applied by all 6 members of the teaching panel. Another was the level of feedback to students on questionnaires taken up. Ethically feedback was restricted to those students who individually sought to find out their result from the subject convenor (the author); relatively few students took up this opportunity. Furthermore, the 4Mat system and the preferred modality teaching were not used very successfully, simply because too much was being changed at once. However the use of learning maps, chunking and teaching about metalearning gained general acceptance among staff and students.

Research into learning (Evans, 1991; Lawson, 1991) shows that deep learners are able to structure their learning, integrating concepts and thus are able to facilitate their recall of pertinent information and apply it in contextually relevant situations. Yates & Chandler (1994) who use synonymous terms such as "mental models" and "mental schemas" see the use of learning maps as an integral aspect of using prior knowledge to best advantage. The larger the knowledge base of the learner, the easier it becomes for them to acquire new knowledge in the same area. Use of learning maps in DAD could therefore build on the common knowledge base students had developed during their first year Bachelor of Business core Information Technology subject. This core subject precedes DAD, recognizing that coming into tertiary study students have hugely varying degrees of exposure to computing, from none at all to several years of computing studies at their school. The Information Technology subject is needed to ensure that all students intending to continue with some form of computing study have reached a common denominator of understanding that will then be built on in future subjects. This assumption of a common prior knowledge influences DAD teaching. Learning maps were employed in this study to give DAD students an insight into the purpose of each topic and how each topic related to the whole subject domain.

As teaching staff we were able to initially structure the subject content in a way that would facilitate an overall understanding of the content and processes being taught. Later in tutorials students were encouraged to develop their own learning maps to structure their learning and build bridges between new and old material they had learnt. Yates and Chandler see learning maps as "instructional scaffolding" that facilitates construction of new knowledge. They say: "Contemporary cognitive theories of learning see the development of knowledge as a process of active construction" (Yates & Chandler, 1994, p 5). One of the learning maps used during lectures is given below as an example 'Fig. 1'. It gives an overview of the conceptual data modelling techniques taught during the semester. Parts were successively shaded to indicate the topic of the current lecture, thus aiding students to appropriately place their current learning into the overall structure of the subject.

Fig. 1 shows that the approach taken in this subject is to analyze the data requirements of a business situation. This sets the scene and encourages students to see where the design task fits into the overall scheme of the System Development Life Cycle. It also helps students to learn the principles in an appropriate context setting; research into teaching generic informatic principles shows this is best accomplished by anchoring the learning in two or more contextual settings (Burmeister, 1994; Davis & Naumann, 1994; Davis & Bostrom, 1994; Brown, 1993).

After discussing the business situation and the process of systems analysis (which is covered in greater detail in other subjects), two broad approaches to data analysis are taught. First the traditional *analytic* approach of normalization by decomposition (Courtney & Paradise, 1992), followed by normalization by *synthesis* using functional dependency diagrams (Eden, 1994; Courtney & Paradise, 1992) and entity relationship modelling (Chen, 1976). In the latter two the final detailed design is the synthesis of a series of less detailed intermediate designs (Date, 1986). The schemas that students produce need to be tested and hence they are taught to create Oracle tables using SQLPLUS (Oracle's implementation of SQL) and to query those tables to the satisfaction of themselves and their tutors. The result is either a satisfactory relational DBMS solution or the need for further analysis and redesign. The synthesis and analytic approaches are not taught as being exclusive of each other. Instead students are taught that in database design one is likely to first synthesize the initial design and then analyze it looking for anomalies. As Kroenke and Dolan put it, "both approaches are necessary in the logical database design process" (Kroenke & Dolan, 1994, p 132). In teaching the process of normalization, assessment only goes so far as third normal form (3NF), though class examples show situations where it is also necessary to go to Boyce-Codd normal form (BCNF) and to fourth normal form (4NF) schemas. Students are told about the possibility of needing to go to fifth normal form, but no examples are given. The reason for teaching beyond 3NF is again to broaden student awareness of the complexities that can be involved in the design process and to encourage deep interaction with the material being studied.

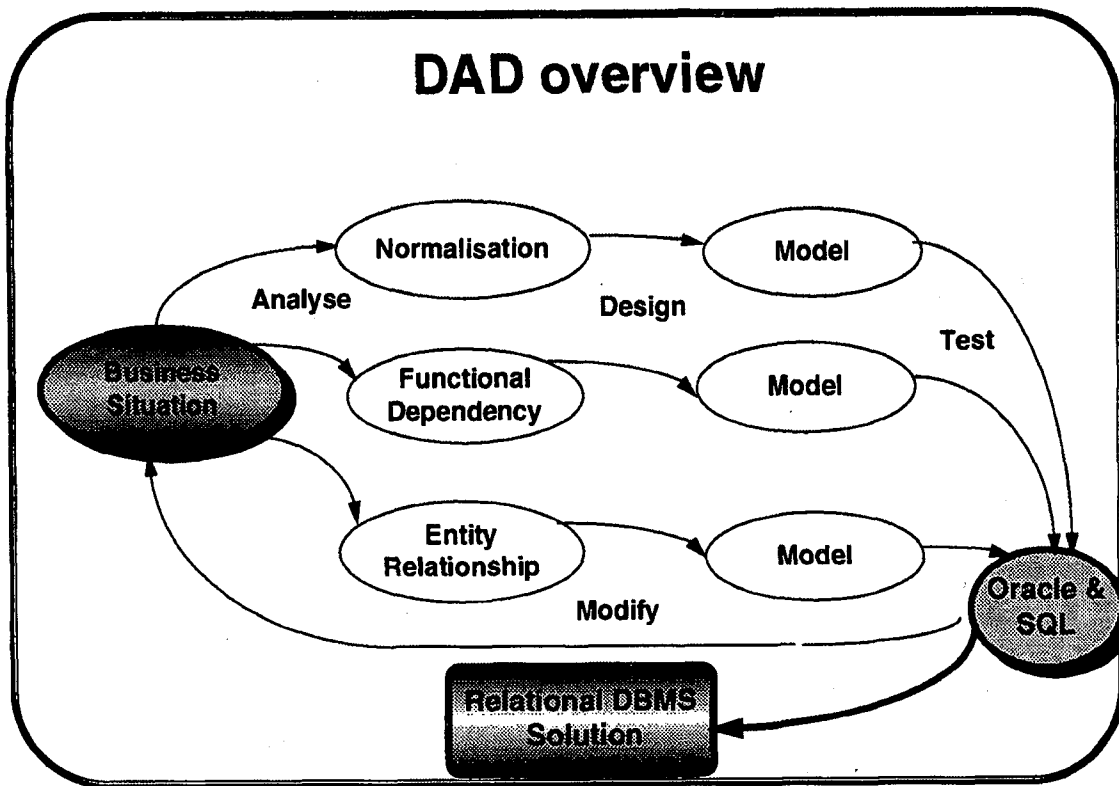


Fig. 1. Data Analysis and Design Learning Map.

Chunking is not so much a strategy for deep learning, as a memorization technique. The idea being that people remember material in "chunks" of 7 plus or minus 2 (McLeod, 1993). It is useful in teaching because the lecturer can isolate the important chunks that students should concentrate on in each lecture; this not only helps them at the time of that lecture, but also facilitates their review of the material covered at a later date. Unfortunately this can promote better surface learning as well as enhance deep learning - a two edged sword!

Metalearning was useful to the mature aged students in particular. At least they were the most vocal during and after the semester in their appreciation of the few minutes taken at tutorial times to help them facilitate their own learning. As well as helping students engage in metalearning during tutorial time, Jan Hastings (Swinburne's Coordinator of Learning and Educational Development) attended 2 of the lectures to speak with the students about better ways of engaging in the learning process.

Assessment design proved difficult. As a teaching panel it was found that our definitions of deep learning for undergraduate (second year) students taking DAD differed. Reaching unanimity took several meetings, and numerous reworkings of sample exam questions and likely sample student responses. As we engaged in this our understanding of the closed mode of the SOLO taxonomy grew, as did our understanding that "deep" can mean different things at different levels of a tertiary program of study. For example, some of our initial attempts at formulating questions in entity relationship modelling could not have been satisfactorily answered at SOLO level 5 except by masters students. The whole process of redesigning the assessment helped us all sharpen our understanding of the aims of the subject and its place within the structure of business computing in the Bachelor of Business. It should also be noted that the students were informed as the semester began, during the semester and at the end of teaching (before exams) of the manner in which the method of exam assessment had changed and the degree to which past papers would be helpful to them.

#### FINAL EXAM ASSESSMENT

The results presented in Table 1 pertain to all 85 students who sat for the final end of semester examination, worth 80% of their assessment in Data Analysis and Design (DAD). Deep learning is here equated with those students whose responses reflected either SOLO levels 4 (relational) or 5 (cross-contextual). As stated above, mastery of SQL is not included in these results; it accounted for 20% of the assessment value of the exam paper. It should also be noted that assignment results are not reported in Table 1. This is in part because not all students who completed the assignments attempted the final exam. It is also because the focus of the assessment redesign was the final exam. Nonetheless the three assignments were used to facilitate deep learning. Briefly, the first two were 'hurdles'; simple cases that students could attempt repeatedly. Their purpose was to provide accurate feedback to students on their progress and to facilitate their acquisition of reasonable proficiency in both data analysis and SQL. The last assignment was a more difficult case, that required they compare alternative solutions and implement the best possible relational database solution in Oracle.

(N = 85)	Number of deep learners
Only normalization by decomposition (NORM)	7
Only functional dependency diagrams (FDD)	0
Only entity relationship modelling (ERM)	13
Both NORM and FDD	0
Both NORM and ERM	7
Both FDD and ERM	4
All three areas	2
Total number of deep learners:	33

**Table 1 Deep learners by data analysis topic**

From Table 1 it is apparent that whilst a sizable minority (38.8%) of DAD students engaged in deep learning, the majority of students did not, and even those who did, did not do so consistently. The latter is to be expected, given that by definition deep learning reflects an intrinsic interest in a topic, a desire to study it for its own sake. Thus it is entirely reasonable to expect students to follow their own study interests and to see those interests reflected in their learning outcomes. Unless a deep learner applies achievement strategies to his or her learning it is even possible that such a student may fail; this could be the case if they were to develop an interest in what is only a very minor aspect of study in a subject. Thus the 2 students whose responses reflected deep learning in all three areas may have been students who had a deep interest in all aspects of data analysis, or they may have applied deep learning strategies to their achieving learning motivations - this may be determined by examining their responses to the Study Process Questionnaire.

The assessment outcomes intended with this study were achieved! Though no High Distinctions were awarded at all, in comparing SOLO outcomes and overall grades obtained in the subject, it was found that the students who were awarded a Distinction also reflected deep learning in one or more areas. However the reverse was not true, that is, not everyone who deep learned in one or more areas was awarded a Distinction; interestingly, none of these students failed the subject.

The above table also indicates that only 6 students (7.1%) deep learned in the area of functional dependency diagrams; only those students who also deep learned in one or both of the other two areas. This may reflect a difficulty with the question itself, or the manner in which this topic is taught. This is an area of future investigation by the teaching panel.

#### **The instructional environment factor**

The 38.8% result is encouraging for informatics staff at Swinburne University. A Department of Employment, Education and Training (DEET) report prepared during 1992 (McLeod & Burgess, 1993) reported that only 15.1% of students enrolling in first year business subjects were deep learners in their approach to learning; the Study Process Questionnaire was used to determine this. This study therefore seems to lend support to the notion that the context of the learning environment is a key factor in achieving deep learning outcomes (Kirby, 1991). It may well be that only 15.1% of these students were inclined to deep learn as a natural tendency, however many more actually engaged in deep learning.

One could hypothesize then that if contextual understanding of data analysis is a major determining factor of deep learning outcomes, individual differences among students are not. This hypothesis can be subjected to analysis in terms of gender, age and student origin. These three areas are identified by the author as the most likely to disprove the hypothesis, for the following reasons.

#### **Gender**

A national study of Australian high schools for Years 8 through 11 was conducted examining learning approaches. This study reported that boys' though not girls' use of achieving and deep approaches declined (Biggs, 1991). This suggests that gender and/or age differences might be found in the above data (Table 1). The next table repeats the information presented above, but breaks it up on the basis of gender.

Gender	Deep learner	Other	Total
Male	16	34	50
Female	17	18	35
Total	33	52	85

**Table 2 Gender and learning**

Given the discrete nature of the data in Table 2, the chi-square goodness-of-fit test is used to determine if a statistically significant factor exists. The hypothesis ( $H_1$ ) being tested is that there is a gender influence on deep learning of data analysis. The value for  $X^2$  is 2.378, which given 1 degree of freedom is not even significant at 90%, let alone above 95%. This leads one to conclude the null hypothesis ( $H_0$ ) holds, gender is not a significant factor in influencing deep learning outcomes in data analysis. Had this not been the case it might have lead one to encourage more females to take up the study of data analysis.

## Age

Behind some learning theories is the Jungian idea that as a person matures their learning preferences tend to develop previously inferior functions. If this were supported by the data in Table 1, then a new hypothesis ( $H_1$ ) would be that mature age students deep learn better than younger students. Table 3 is split according to the mature age classification of the Melbourne Student Assistance Center. Thus any student who was 23 years of age on or before March 1st, 1994 is classified as a mature age student below. The total range of ages was from 18 to 53. The mean age of mature age deep learners was 30.29, whilst that of other mature age students was 31.54. The mean age of young deep learners was 20.71, whilst that of other young students was 21.18.

Classification	Deep learner	Other	Total
Mature age students	13	13	26
Young students	20	39	59
Total	33	52	85

Table 3 Age and learning

Again the value for  $X^2$  (1.976), given 1 degree of freedom, is not significant at 90%, let alone above 95%. This leads one to conclude the null hypothesis ( $H_0$ ) holds, age is not a significant factor in influencing deep learning outcomes in data analysis.

Yet this result may show the need for further research into the effect of age. The reason being that the Australian High School study referred to above, appeared to show a decline in deep learning with age in boys. No correlation to age was found, which may suggest that the declining trend reverses itself at some point. Another avenue for further investigation is the combination of age with gender.

## Origin

Biggs, the originator of the SOLO taxonomy and now Professor of Education at Hong Kong University, has written of differences in learning between Australian and Asian students (Biggs, 1994). In DAD the majority of overseas students come from Asian countries. However, the data only shows whether students come from overseas or not. Hence if Biggs' findings are supported by the data in Table 1, then a new hypothesis ( $H_1$ ) would be that overseas students exhibit different learning outcomes than Australian resident students. Table 4 is split according to the origin of the student and their learning outcome.

Classification	Deep learner	Other	Total
Australian resident	22	25	47
Overseas student	11	27	38
Total	33	52	85

Table 4 Origin and learning

In this case the value for  $X^2$  is 2.818, which given 1 degree of freedom, is significant at 90%, but no longer at 95%. This leads one to conclude the null hypothesis ( $H_0$ ) holds, origin is not a significant factor in influencing deep learning outcomes in data analysis. The author tends to discount the significance at 90% given that some overseas students may not come from Asia. Also given that being Australian residents might mean they came from overseas originally, but became Australian residents before beginning tertiary study. Thus being Australian residents does not necessarily equate with exposure to the whole of the Australian school system.

Nonetheless, the results of Table 4 would encourage further investigation into the relationship of origin and learning outcomes in data analysis. Any follow-up study would need to capture data about student origins that carefully differentiates Australian and Asian students.

Thus the author returns to the original proposition that the context of the instructional environment was a major influencing factor in producing deep learning outcomes in these data analysis students. The data does not support the view that individual differences between students account for significant differences in their learning outcomes.

## APPROACH TO LEARNING

The approach to learning seeks to measure the level of student motivation and its nature, as well as the types of strategies employed by the student to satisfy that motivation. The Study Process Questionnaire was used to measure these factors both before the semester began and again after the semester's study had been completed. It gives a decile score for the deep approach to learning, which compares each student to a population of 1000 students across 3 Australian tertiary institutions (Biggs, 1987). An average decile score is 4 to 7 (within the middle 31 to 70 per cent of the population), 8 to 9 being above average and 10 being well above average.

As stated in the methodology section above, there were a number of unforeseen difficulties with the Study Process Questionnaire. These difficulties made comparisons between the before and after measures complicated, let alone comparisons to the SOLO assessed students responses. For this reason only comparisons for the 28 students who took both the SPQ measures and were assessed according to SOLO were chosen.

In the before measure 64.29% of students exhibited either an average, above average or well above average deep approach to studying, compared with 67.86% in the after measure. Whilst this shows a marginal

improvement, given this only compares 28 students it cannot be interpreted as being representative of all DAD students. However, in the light of a recently reported study in Hong Kong, the result obtained in DAD is encouraging. The Hong Kong study involved tertiary Accounting students covering first year, second year, third year and honors students. It is reported that from first to second year achieving and deep learning declined, but that from second year to honors 'mean deep-approach scores rose slowly' (Gow, Kember & Cooper, 1994, p 118). The study pointed to heavy assessment demands and didactic teaching styles in first and second year tertiary study, as the major reasons for the initial decline. In the light of these findings and that DAD is taught as the first subject in second year, for business (Account, Marketing, Economics and Computing) students, the marginal increase in deep learning is counter to what would be expected. Again, the author ascribes this outcome to the deliberate creation of a teaching environment that facilitated deep learning in DAD. A similar conclusion is reached by the Hong Kong study. 'This study provides evidence that approaches to studying are directly related to the learning context.' (Gow, Kember & Cooper, 1994, p 128).

The main usefulness of the SPQ is that it can give an indication of the *motivation* for study; whereas the SOLO taxonomy only reveals the *structure* of the student response. Eliciting the type of strategy employed by the student, but not the motivation. Hence based on the SOLO analysis above one can say that 38.8% of DAD students employed a deep strategy for one or more parts of the exam. However it is altogether possible that they were motivated by achieving or surface learning. The SPQ analyses learning approach in two ways - strategy and motivation. It would then be useful to examine the SPQ results to see what the dominating motivation was for studying DAD. Thus one can quantify the extent to which the instructional environment developed intrinsic motivation in students to engage in the subject (Davis & Bostrom, 1994).

Below is a comparison of the before and after learning approaches, purely from the point of view of motivation for the 28 students identified above. All three learning approaches identified by the SPQ (surface, deep and achieving) are shown.

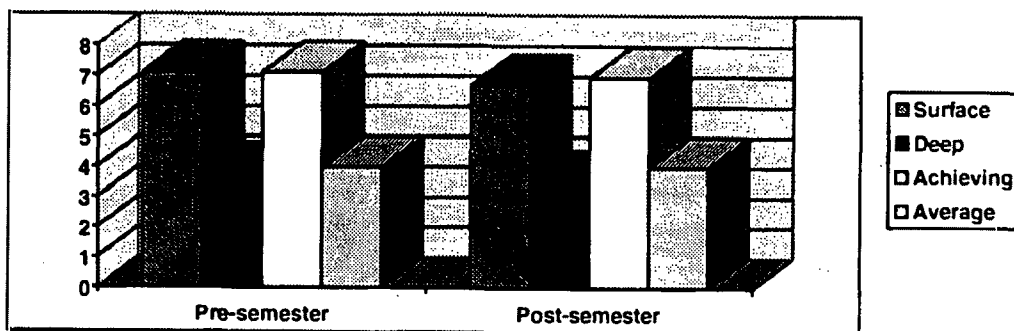


Fig. 2 Mean SPQ motivational decile scores of all respondents

Unfortunately nothing conclusive can be gleaned from this data. All three areas of motivation decreased from the before to the after study. The smallest decline was in the achieving motivation.

According to the SOLO analysis 15 of these students structured their responses in ways that suggest they used deep strategies and the other 13 did not. The graph below presents learning motivation only for the 15 students who gave level 4 or 5 SOLO responses.

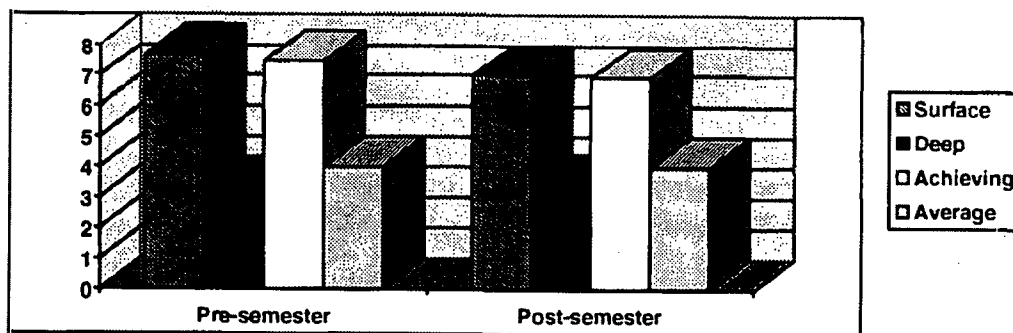


Fig. 3 Mean SPQ motivational decile scores of deep learners

For these 15 students there is a decline in both achieving and surface motivation, but a marginal increase in deep motivation. Interviewing all 28 of these students might shed further light as to the reasons behind the discrepancies between these two graphs. However, given the small numbers would be prudent to carry out another SPQ survey with many more participants and to draw meaningful conclusions from it instead.



## LEARNING PREFERENCES

The study revealed the following (N = 42):

ISTJ 21%	ISFJ 2%	INFJ 0%	INTJ 2%
ISTP 12%	ISFP 2%	INFP 5%	INTP 10%
ESTP 14%	ESFP 10%	ENFP 0%	ENTP 5%
ESTJ 17%	ESFJ 0%	ENFJ 0%	ENTJ 0%

**Table 5 DAD learning preferences**

Two things are clearly apparent. Firstly the strong predominance of Thinking ("T") types, a total of 81%. These are students who are good at putting things in logical order, who respond more to people's ideas than to their feelings, they anticipate the logical outcomes of choices. These are students who need to be treated fairly, are firm and tough minded, and have a talent for analyzing a problem or situation.

Secondly the fact that 79% were Sensing ("S") types (64% of the students were both "S" and "T" types). Sensing types are students who are aware of the uniqueness of each event, who focus on what works now; they like an established way of doing things and enjoy applying what they have already learned. They work steadily, with a realistic idea of how long it will take, and usually reach a conclusion step by step. Sensing types are careful about facts and are unlikely to trust their own inspirations; they may be good at precise work, yet can oversimplify tasks; they accept current reality to work with (Myers & Myers, 1987).

Unfortunately not being part of the original design, the impact of this measurement could not be capitalized upon as the MBTI test was not conducted until the final week of the semester - too late to impact upon teaching. Had this measurement been undertaken earlier in the semester, preferably at the very start, it could have significantly influenced the manner in which the subject was taught.

Analyzing the results summarized in the above table one can see that Sensing and Thinking types predominate, with the STJ combination accounting for the largest sub grouping within the student population (38%), Extroversion and Introversion types appearing in roughly equal numbers. However, given the significant degree to which team work skills are required in the systems discipline, this large proportion of Introversion types need to be encouraged to develop inter-personal relationship skills. Another conclusion that can be drawn from this data is that the use of learning maps should have proven beneficial to many of these students. Given the large "S" factor these students would be less likely to see the whole picture and would require encouragement to do so.

An attempt was made to find a connection between the MBTI and the SPQ results, however without success. A similar attempt to compare MBTI with SOLO results is given below. Of the 42 respondents to the MBTI 16 were among those who used deep learning strategies in the final exam, as assessed by the SOLO taxonomy, 10 were anonymous and 2 did not end up sitting for the final examination.

ISTJ 25.00%	ISFJ 0.00%	INFJ 0.00%	INTJ 6.25%
ISTP 6.25%	ISFP 6.25%	INFP 6.25%	INTP 12.50%
ESTP 12.50%	ESFP 6.25%	ENFP 0.00%	ENTP 0.00%
ESTJ 18.75%	ESFJ 0.00%	ENFJ 0.00%	ENTJ 0.00%

**Table 6 Learning preferences of deep learners**

As for the larger group, "S" and "T" types predominate; 75.00% are "S" (79% for the larger group) and 81.25% are "T" (81% for the larger group) types. However, whereas the larger group had roughly equal numbers of Introverts and Extroverts, this group is made up of 62.50% Introverts, people who prefer to study individually. In future as well as a larger sample size for students, it would be helpful to take up the MBTI of the teaching staff, to see where their strengths are as compared to the learning preferences expressed by their students.

### Comparing DAD MBTI results to those of informatic professionals

The only data that exists on professional MBTI scores comes from the US and was compiled over a period from 1971 to 1984. Also they were 'samples of convenience, not randomly selected samples' (Myers & McCauley, 1985). Hence given the nature of rapid changes in the computing industry since then, one cannot rely overly on this comparison. Yet no other such extensive data exists and therefore it will have to suffice. Informatics was represented under a variety of categories, of which the author chose 'Computer systems analysts, support

representatives' as the closest category to that of data analysis and design. Other categories included: 'Computer specialists', 'Computer programmers', and 'Computer and peripheral equipment operators'.

ISTJ 13.95%	ISFJ 5.81%	INFJ 2.33%	INTJ 12.79%
ISTP 3.49%	ISFP 3.49%	INFP 4.65%	INTP 9.30%
ESTP 1.16%	ESFP 1.16%	ENFP 1.16%	ENTP 10.47%
ESTJ 16.28%	ESFJ 1.16%	ENFJ 3.49%	ENTJ 9.30%

Table 7 Computer systems analysts, support representatives (1971 to 1984).

There are striking commonalities between tables 6 and 7. Most notable is the fact that the two most predominant types in both tables are ISTJ and ESTJ. Also the fact that "Thinking" types predominate in both tables (81.25% in table 6 and 76.74% in table 7). However there are also significant differences between these two tables. Whereas in table 6 75.00% were "Sensing" types, only 46.50% were in table 7. Also while introverts accounted for 62.50% of table 6, in table 7 they only marginally outnumber extroverts (55.81).

Again given the age of the data that is represented in table 7, one ought not be too definite about what the similarities and differences to table 6 signify. Perhaps one could argue that people most inclined to deep learn the principles of data analysis and to enjoying a career in analysis and design, are those who predominate in "Thinking" personality types. To repeat the definition given above, "Thinking" types are people who are good at putting things in logical order, who respond more to people's ideas than to their feelings, they anticipate the logical outcomes of choices. They are people who need to be treated fairly, are firm and tough minded, and have a talent for analyzing a problem or situation (Myers & Myers, 1987).

### CONCLUSIONS & RECOMMENDATIONS

The focus of this study was primarily on deep learning outcomes in data analysis students. Secondly the study sought to produce tertiary students who become autodidactic, life long learners, enriching themselves and others through their continuing studies. The SOLO taxonomy was used to measure accurately the strategies component, the SPQ to measure motivation. Use of the SOLO taxonomy revealed that 38.8% of students achieved deep learning outcomes. The SPQ showed a marginal increase in deep learning motivation, despite the fact that a similar domain specific study of Accounting students would lead one to expect a decline in deep learning motivation. This was in line with comments made by various students to the effect that their interest in the subject had increased. Nonetheless the study was not able to integrate the use of these two tools (SOLO and SPQ) as well as the author had hoped and it therefore remains possible that some students who evidenced deep strategies had achieving motivation or (less likely) surface motivation.

From an action learning perspective this study has been very rewarding personally. The author found new enjoyment and satisfaction in the area of specialty. It has been gratifying to have a number of students make comments to the effect that they were glad their teaching staff showed this level of interest in their learning. The author also found increased understanding of the learning task a rewarding experience and developed professionally as a result. In attempting to motivate students to develop a deep understanding and in formatting assessment that ensures SOLO level 5 is tested for, the author gained a far superior understanding of this area of systems work than before!

Further study is possible through revisiting this study at the time each of these students complete their degree study at Swinburne, to give a longitudinal vantage point to interpreting the success of this study. Also, follow-up interviews could be conducted with "deviant" cases, such as students whose SPQ results suggest they ought to perform very well/poorly, yet the reverse turns out to be the case. This could lead to further worthwhile enhancements in the teaching processes currently engaged in with DAD. Another area that can be investigated is the extent to which these same students apply their metalearning in other subjects; did they see it as something only applicable to DAD, or do they apply that metalearning in other contexts?

Another area for future study is to investigate the possibility of embedding deep learning data analysis behavioral teaching in a multi-media learning package. This might be accomplished by analyzing the neurological behavior of competent professionals and perhaps final year or postgraduate students. Then with appropriate instructional design, such behavior could be built into the teaching of a computer assisted learning package, thus becoming available to all students and professionals working in this area.

One recommendation to anyone interested in using the MBTI personality indicator questionnaire is to use it before the period of study begins. Thus the benefit of this analysis can be available to teaching staff and students alike early in the piece. It is a fountain of knowledge that was sought too late in this study.

The evidence presented here appears to indicate only a marginal increase in deep learning outcomes (see Figure 3). However, as indicated earlier DAD is taken by students either in the second semester of first year, or in the first semester of second year. In the light of the Hong Kong study (Gow, Kember & Cooper, 1994) the outcome should have been a decline in deep learning. Thus any increase, however marginal, is in fact significant! Deep learning outcomes were not based on individual characteristics such as gender, age or origin. Instead this study suggests that the major contributing factor to these deep learning outcomes was the instructional environment. In terms of personality preferences, deep learning predominated among students whose preferences were "Sensing", "Thinking" and "Introversion".

Not all DAD students engaged in the type of quality learning being investigated in this study. Yet it is the author's hope that amongst the 38.8% who did are likely to be many who will find great pleasure in pursuing their research interests well beyond their tertiary years. It is hoped too that some of them will make notable contributions to the area of data analysis in the years to come.

## REFERENCES

- Australian Council for Educational Research, (1985) *SPQ*, Hawthorn: ACER.
- Bachelor of Information Technology Quality Project**, (1994), Hawthorn: Swinburne University of Technology.
- Ballantyne, R., Bruce, C. & Packer, J. (1993) *Action Learning in Vocational Education and Training*, Chadstone: TAFE National Staff Development Committee, Dec.
- Biggs, J.B. (1978) "Individual and Group Differences in Study Process", *British Journal of Educational Psychology*, 48, pp 266-279.
- Biggs, J.B. (1979) "Individual differences in study processes and the quality of learning outcomes", in *Higher Education*, 8, pp 381-394.
- Biggs, J.B. (1991) "Student learning in the context of school", *Teaching for Learning*, Hawthorn: ACER, pp 7-29.
- Biggs, J.B. (1987) *Study Process Questionnaire Manual*, Hawthorn: ACER.
- Biggs, J.B. (1994) "What are Effective Schools? Lessons From East and West", *Australian Educational Researcher*, Vol 21, No. 1, Apr, pp 19-39.
- Bloom, B.S. (1956) *Taxonomy of Educational Objectives*, London: Longmans.
- Brown, A.R. (1993) "The Adventures of Jasper Woodbury and Anchored Instruction", *ASCILITE conference proceedings*, 6-8th Dec, Bruce W N Lo (ed), The University of New England-Northern Rivers Lismore, NSW, pp 55-64.
- Burmeister, O.K. (1994) "A critique of business computing hypermedia designs in terms of their efficacy in facilitating learning", *Applications of New Technologies in Business* conference paper, Monash University, Dec.
- Candy, P. (1991) *Self-Direction for Lifelong Learning*, San Francisco: Jossey - Bass Publishers.
- Chen, P.P. (1976) "The Entity-Relationship Model - Toward a Unified View of Data", *ACM Transactions on Database Systems*, Vol. 1, No. 1, March.
- Collis, K. & Biggs, J. (1986a) "Using The SOLO Taxonomy", *SET: Research Information for Teachers*, No. 1, Hawthorn: ACER.
- Collis, K. & Biggs, J. (1986b) *Evaluating the quality of learning: The SOLO Taxonomy*, New York: Academic Press.
- Courtney, J.F. & Paradice, D.B. (1992) *Database Systems for Management*, Homewood: IRWIN.
- Date, C.J. (1986) "A Practical Approach to Database Design", *Relational Database: Selected Writings*.
- Davis, & Bostrom, (1994) "The effects of an intrinsically motivating instructional environment on software learning and acceptance", *Information Systems Journal*, Vol 4 No 1, Jan.
- Davis G.B. and Naumann J.D., (1994) "A new second information systems course: personal productivity with information technology", *Proceedings of the IFIP TC8 Open Conference on Business Process Re-engineering: Information System Opportunities and Challenges*, May, pp 89-96.
- Eden, P. (1994) *Data Modelling*, Course notes, Hawthorn: Swinburne University of Technology.
- Evans, G. (1991). "Student control over learning", *Teaching for Learning*, Hawthorn: ACER, pp 126-145.
- Giampietro, N. (1994) *Developing Hybrid Managers to enable the effective use of Information Technology in Business*, Master of Business in Information Technology thesis, Swinburne University of Technology.
- Gow, L., Kember, D. & Cooper, B. (1994) "The Teaching Context and Approaches to Study of Accounting Students", *Journal of Accounting Education*, Vol. 9, No. 1, Spring, pp 118-130.
- IFIP ICC (1966) *Vocabulary of Information Processing*, North Holland.
- Jensen, S. (1987) "Learning Styles", in J. A. Provost & S. Anchors (eds) *Applications of the Myers-Briggs Type Indicator in Higher Education*.
- Kirby, J.R. (1991) "Reading to Learn", *Teaching for Learning*, Hawthorn: ACER, pp 103-125.
- Kroenke, D.M. & Dolan, K.A. (1994) *Database Processing*, Sydney: SRA, Third Edition.
- Laurillard, D. (1988) "Evaluating the Contribution of Information Technology to Students' Learning", *Evaluating Information Technology in the Arts and Humanities proceedings*, Birmingham, Computers in Teaching Initiative Support Service, pp 39-45.
- Lawson, M. (1991) "Managing Problem-Solving", *Teaching for Learning*, Hawthorn: ACER, pp 126-145.
- Martin, J. (1989) *Information Engineering*, New Jersey: Prentice Hall, Book I: Introduction.
- Martin, J. (1990) *Information Engineering*, New Jersey: Prentice Hall, Book II: Planning & Analysis.
- Marton, F. & Saljo, R. (1976) "On Qualitative Differences in Learning: II-Outcome as a Function of the Learner's Conception of the Task", *British Journal of Educational Psychology*, 46, pp 115-127.
- McCarthy, B. (1987) *The 4Mat System*, Barrington: Excel.
- McLeod, C. (1993) *Study Success Without Stress*, Hawthorn: Swinburne University of Technology.
- McLeod, C. & Burgess, Z. (1993) *Teaching and Learning Enhancement in a Multi-Sectoral Institution*, Hawthorn: Swinburne University of Technology.
- McLeod, Jr., R. (1994) *Systems Analysis and Design*, Sydney: The Dryden Press.
- Moody, D. & Shanks, G. (1994) *What Makes A Good Data Model? Evaluating the Quality of Entity Relationship Models*, Swinburne University Centre for Information Systems Research Seminar, Nov.
- Myers P.B. & Myers, K.D. (1987) *Myers-Briggs Type Indicator*, Palo Alto: Consulting Psychologists Press, Inc.

- Myers, I.B. & McCaulley, M.H. (1985) **Manual: A Guide to the Development and Use of the Myers-Briggs Type Indicator**, 2nd ed., Palo Alto: Consulting Psychologists Press, Inc.
- Olle, T.W. (1993) "Data modelling and conceptual modelling: A comparative analysis of functionality and roles", **Australian Journal of Information Systems**, Sep., Vol. 1, No. 1, pp 46-57.
- Ramsden, P. (1984) "The context of learning", in F. Marton (ed), **The Experience of Learning**, Edinburgh: Scottish Academic Press.
- Revans, R.W. (1982) "The Nature of Action Learning", in R. W. Revans **The Origins and Growth of Action Learning**, Bromley: Chartwell-bratt.
- Swinburne University of Technology, (1995) **Handbook**, Hawthorn: Swinburne University of Technology.
- Whitten, J.L., Bentley, L.D. & Barlow, V.M. (1994) **Systems Analysis and Design Methods**, Sydney: Irwin, Third Edition.
- Wills, S. & McNaught, C. (1993) "Evaluating computers in learning in tertiary education", **ASCILITE conference proceedings**, 6-8th Dec, Bruce W N Lo (ed), The University of New England-Northern Rivers Lismore, NSW, pp 691-703.
- Yates, G. & Chandler, M. (1994) "Prior Knowledge", **SET: Research Information for Teachers**, No 2, Item 6.