Y2K – Real problem or information technology's way to make money?

René Pellissier

Unisa Graduate School of Business Leadership

CC Pienaar

Abraxas Technologies

Addressing the year 2000 problem (also known as Y2K) and bringing systems into compliance is not an option, but a survival necessity for businesses and individuals. If there is anything to be learnt from this situation, it is that prior planning prevents panic – even catastrophe. It is feared that 15 per cent of businesses may not survive the problem. Furthermore, it is estimated that the problem will cost the international community billions of dollars not only in fixing the problem but also in legal cases that can arise from systems not being compliant. Problems range from the real-time clock in personal computers to escalators, high technology research equipment, traffic lights, video recorders (VCRs), critical life-support systems in hospitals, and banking services. Companies should not see the year 2000 bug as a problem to be fixed as quickly and cheaply as possible. They should see it as a challenge and opportunity to make the technological change that can provide them with a competitive edge. In doing this they should ensure that information technology remains an enabler and not a deterrent in the competitive business world.

Introduction

Y2K or the millennium bug. What is this and where did it originate? Is this a real problem or is this a fad of the information technology (IT) industry to generate more work and more money? Who will be affected by this bug and how will it impact on businesses and the world economy? How can risk be minimised and the problem fixed? The *timeous* solution to this question could save companies time and money.

Rock & Reynolds (1998) argue that the countdown to the year 2000 glitch began more than two decades ago when, in order to save computer memory space (then costing $600 000 but today a mere 10 cents), programmers used two digits rather than four to represent years. Kappelman (1996) states that this problem can be traced to the 1950s and early 1960s when the number of characters that fitted on an 80-column punched card was a major constraint. To save memory space the '19' of the century was abbreviated since users and programmers understood what it implied – but not the computers. With the new millennium approaching, programs should be changed to store and recognise the new century. Most of the dates are stored in six-character length. It is suggested that, to fix this, not only do programs have to change, but the underlying data should also be changed. This is standard in most applications. The fact that years have been uniformly stored in computers as two digits, and that the century was always assumed to be 1900, has created what is known as the *year 2000 problem* or *year 2000 bug*.

The problem does not only exist in application programs but starts at the lowest level in computers. Most computers have a real-time clock built into the hardware. This clock has a battery and memory and remembers the date and time set. Generally, the year in most of these clocks also uses two digits only. The century is assumed to be 1900. Furthermore, different manufacturers have used different baselines for the dates. In some clocks the baseline is 1900 and in others 1980. This means that the clock will tick over from 99 to 00 for a baseline of 1900 and from 99 to 80 in the second instance.

The next level where the problem materialises is in the BIOS of the computer. This is a special hardware chip that, among its many other functions, tells the computer the time, day, month and year. Since these chips are hardwired (a built-in program), they cannot be reprogrammed. Normally the BIOS will get the date from the internal real-time clock. In some cases, depending on the type of BIOS, the BIOS will use the date obtained from the real-time clock and add the century or change it to some other format to be used by the operating system.

From the operating system, the date passes to the application software. Many programs in the application software use dates in calculations. Consequently, after 31 December 1999, calculations will, for instance, deduct 97 from 00 giving negative 97, whereas what should have happened is that 1997 should have been deducted from 2000 giving the correct answer of 2000 - 1997 = 3 years. This is one example of incorrect date calculations found on older computers and programs. In most new computers the problem has been fixed. These computers store the year in four digits and date calculations will be correct.

Kappelman (1996) further states that, depending on the industry and application, three to six per cent of the data stored in organisational databases are in date form. Using four-digit rather than two-digit storage fields would have required 33 per cent more storage space for dates. This effectively means that a company with an average of 10 gigabytes of data over a 30-year span saved between $160 million and $240 million.

The problem does not end at computers and software. One is surrounded by systems that are linked to other systems. Many of these pass dates or use dates for synchronisation. Thus, outside links also need to be investigated.

According to Hoffman (1996), a survey by the Society for Information Management revealed that senior non-IT executives at most organisations have little knowledge of the potential impact of the year 2000 problem. This problem is compounded by the fact that many chief information officers are afraid to inform them. Leon A. Kappelman, head of the Society for Information Management studies, maintains that many information systems managers do not have a good explanation for board-level executives as to what caused the year 2000 problem. It is thus evident that there is a substantial education problem to be resolved before the year 2000 problem can be addressed. The problem is due, in part, to the lack of literature producing sufficient public discussion and solutions to this serious issue.

Impact

According to Bordwin (1998), some experts predict that 15 per cent of businesses will not survive the year 2000 problem. For the smart or lucky companies already engaged in resolving the problem, the next two years will only mean stepping up the repairs and testing. They will ensure that all their business partners are also addressing the same issue. Bordwin further states that one will have to ensure that staff is continuing with year 2000 projects. Staff with year 2000 project experience should become more and more in demand.

Scheier (1997) states that for the companies that have not started or have just started addressing the issue, 1998 and 1999 will be a time of triage – of ruthlessly deciding which systems should be fixed and which should be patched, scrapped or retained until later. They will be scrambling for computer resources and critical skills, such as programming and project management skills. According to a Cowen & Co. Research Report on year 2000 issues (Kador 1997), 28 per cent of companies are unaware of the extent of their problem.

Cost and time implications

Kappelman, Fent, Keeling & Prybutok (1998) argue that the cost of updating or replacing information technology for year 2000 compliance is highly situational and depends on an organisation's specific internal and external factors such as the applications they use and the links to other organisations. They identify four main categories of cost influences:

- The organisation and its structure
- The extent and complexity of its year 2000 problem
- How the organisation will solve its problem
- The availability of resources to achieve the solution, including people, skills, and funds.

Caldwell (1997) states that in a survey of 108 information systems managers, 82 per cent of the respondents believed that they underestimated total year 2000 costs. For the most part, the incorrect estimates resulted because project managers did not know how to measure desktop and network costs. Based on respondents who have revised their year 2000 estimates to include desktop and networking software, he believes that the cost of bringing distributed computing into compliance will be nearly half that of fixing mainframe systems.

Kappelman et al. (1998) further argue that cost estimates should include the cost of project management, direct labour and identification of affected software and data, and alternative solutions, testing, and implementation. The variance in estimated percentage costs found among top information systems managers and year 2000 project managers indicates that managers should work toward a shared vision of year 2000 projects within their organisations, especially since this difference appears to increase as projects progress into later planning stages.

A series of tracking polls taken in 1996 by Cap Gemini America showed that corporate awareness of the year 2000 date-change problem increased last year (*Computerworld* 1998). The following are some estimates of the cost to prepare for the year 2000:

- Gartner Group Inc. – $600 billion; closer to $1 trillion including litigation costs
- Capers Jones, Software Productivity Group – more than $3.6 trillion
- US government – its fixes alone will cost at least $3.6 billion
- Dataquest Inc. – US companies will spend a total of $14 billion on outside consultants and programmers. Internally, they could spend triple that amount (*Business Week* 1997
- According to the minister of Post, Telecommunications and Broadcasting, Jay Naidoo (*Telecommunications* 1997), the year 2000 problem could cost the South African economy between R7 billion and R25 billion.

It is clear that the window to fix the year 2000 computer programming problem cheaply or easily is already closing. From now on, the closer to 31 December 1999, the higher the cost to fix the problem and the more difficult it will become to fix it and to obtain and retain people to do so.

One of the biggest problems is obtaining funding for the testing and fixing phase. According to Hall (1998), one frequent scenario is to make use of existing funds. This is mainly because most companies have not budgeted for the year 2000 impact or do not realise the extent thereof. The closer one gets to the year 2000, the bigger the budget should be for the year 2000 project. Resources will be more in demand and steep price increases can be expected.

Management

Many firms are now engaged in urgent efforts to correct their software. Managers should not adhere to the belief that it is safe to wait until 1999 to focus on the year 2000 problems. Management should become aware of the problem, identify systems to be changed, predict the consequences of the errors, set time frames, and allocate resources. Fryer (1997) argues that information technology managers should show the CEO or board of directors what would happen to the company if a central system should fail. He believes that top management should–

- communicate the importance of the year 2000 problem to the rest of the company;
- appoint project managers to manage the year 2000 project;
- hire consultants, if necessary;
- review vendor contracts;
- protect the company against legal risks.

It was found that because of the fixed deadline that does not allow for schedule slippage, the use of powerful dependency modelling and simulation tools allows a company to make priority-based trade-offs in order to complete the most important tasks first. Doing so mitigates the possible damage to the enterprise and minimises the risk represented by the millennium problem (McArthur 1997).

Kelley (1998) argues that, because of the seriousness of the problem, managers face huge challenges, including the following:

- managing a highly complex project involving multiple interfaces with other systems;
- ensuring the readiness of vendor components;
- ensuring the readiness of applications;
- thorough testing;
- extensive communications;
- establishing contingency plans.

All of these should be managed above and beyond other day-to-day projects. This places enormous pressure not only on managers but also on everyone involved. As stated before, this is not only a software problem but impacts on the business world at large. A substantial part of the year 2000 problem involves business decisions and not only solving the software problems. Examples of business decisions that should be taken include replacement or fixing of a system that may cost millions or even termination of a service. In this, the manager should weigh up the risks and costs involved and the legal contracts held by the company.

According to De Jager (1998), actions that information technology managers could take to ensure positive outcomes include:

- formulating a detailed year 2000 project plan;
- identifying how many internally coded applications they have;
- identifying how many vendor applications they have;
- looking at the existing contingency plan and developing some if they do not have overall contingency plans;
- ensuring that the right person is in charge of their year 2000 project;
- looking at who that person reports to and how often;
- looking at who are on the year 2000 project and setting priorities if these persons are on other projects as well. (There is often a temptation to overstaff the year 2000 team but this should be resisted as smaller teams work better.)

Companies should look at how to measure and track their year 2000 project's performance. The META Group's Mike Egan tracks the company's year 2000 performance in the following categories (Gibbons 1998):

- Program management office (PMO)
- Code conversion
- Funding
- Testing
- Corporate effectiveness.

The program management office is the most critical element of the year 2000 project. Program management office personnel are often faced with the rigorous project management effort, including planning, internal and external communications, and burdensome paperwork.

Human resources

The year 2000 problem has a far-reaching impact on more than computers and software. Delancy (1998) argues that unless more tools are created, most year 2000 projects will become labour-intensive. The closer one gets to 31 December 1999, the more people are needed to fix the problem. Even with the appropriate tools, people are still required to test systems and fix problems. Delancy reiterates that companies will need to staff up, and subsequently staff down, within a relatively short period of time via new hires, independent contractors, or leased employees. Furthermore, many human resources departments are still totally uninvolved in the process or do not realise the extent of the problem.

Hall (1998) believes that, as year 2000 projects proceed into more formal analysis and remediation stages, companies should view outsourcing as an alternative. Because of the high volume of work (including testing and fixing) and the time constraint involved, more skilled employees are needed. It is almost impossible to derive long-term benefits from the find-and-fix process. In most cases, shops are better off leaving the find-and-fix to a qualified outsourcing factory. Hall further states that find-and-fix factories come in three basic forms, which are onsite, offsite, and offshore. For organisations with the number of lines of code in programs and systems below one million, it may be worthwhile to outsource.

According to Kimbark (1997), human resources professionals have three roles to play in year 2000 compliance:

- an internal role of ensuring that human resources systems are compliant;
- a bigger role of assisting information technology with the resources it needs to address the problem throughout the organisation;
- a role of communicating information.

Auditing and legal issues

According to Hock (1998), the directors and people working towards a solution to the problem should bear in mind the legal issues. To avoid personal liability, one is required to meet

the applicable legal standards of care to be taken. Hock suggests that efforts should be carefully documented to strengthen defences under the due diligence and business judgement rules in the event of litigation over the company's failure to solve the problem in whole or in part. According to Anthes (1998), companies should prepare for damage claims from customers, partners and shareholders. Some risk sharing should be negotiated in contracts, if possible. Companies should also review, and possibly strengthen, insurance policies for critical factors such as business interruptions, errors and omissions, product liability, and malpractice.

During the next few years, most companies will face year 2000 costs. Hock argues that the tax law treatment of these costs will in most cases have a significant impact on the company's after-tax income. He suggests that companies develop a tax strategy, including careful structuring of agreements with outside vendors, careful documentation of the plans for solving the year 2000 problem and documentation of costs incurred. These steps should also enable favourable tax treatment.

Types of problems and how to manage them

The types of year 2000 problems and places where it can be found range from personal computers and real-time clocks, BIOS and operating systems of the computers to embedded chips in non-IT equipment used in hospitals, oil drilling rigs, pipeline valves including water pipelines, escalators, and other non-IT equipment used in research and industrial facilities. More and more companies that, until recently, did not perceive any problem, now realise that they might have a serious one and should start assessing where and how big it is. One of the biggest problems is management's lack of understanding of the types of problems as well as where to look for them.

The effect of the year 2000 problem is already being felt worldwide. Seven per cent of information technology directors and managers maintain that their companies have already experienced a year-2000-related failure (Radosevich 1998). The problem initiated in the credit card industry when the expiry date on credit cards moved into the new millennium. Some of the credit card systems used were (and still are) not compliant. The systems cannot handle the expiry date of 2000 or more. Some credit cards even have an expiry date of '00. One way to manage this is to enter an expiry date of 1999. Alternatively, the retailers will have to call for verbal authorisation from the bank. This heralds bad news for credit card companies, not only because one of the controls built into their systems is not working but also because they lose their credibility as providers of top systems.

Kelley (1998) states that the year 2000 problem has the potential to disrupt seriously the infrastructure of computer systems and telecommunications that the world community depends upon for the free flow of funds and payments and, hence, virtually all of everyday commerce. The global nature of the financial services industry relies upon the interconnection of computer systems worldwide. Money is transferred internationally and nationally on a daily basis. International systems that are linked need to be synchronised because they might fall in different time zones. To accomplish this, programmers use date–time stamping where each record or transaction gets stamped with the date and time. By reading the date and time and knowing the difference in time between the different time zones, the local time of creating the transaction can be calculated. Without this technique, it will be difficult to synchronise systems in different time zones. This synchronisation technique is used not only in financial systems but also in international companies with branches in different countries using the same systems. These systems are then connected and kept in synchronisation by earmarking one system as the host and the others as slaves. Data capturing can take place in all the systems but are then 'downloaded' to the host system. Using this technique, duplication and consolidation of information can be managed.

There are still only a few manufacturers who are really aware of the impact of the year 2000 problem on their businesses or have assessed the magnitude of their problems. Some of these problems might have an enormous impact on their businesses and might even stop their operations completely. Like all other industries, they also face a scope of problems including financial, resource and skills scarcity problems. The complexity of some of the equipment used at the plant floor makes the year 2000 problems more difficult to assess and fix. To make matters worse, some manufacturing plants run 24 hours per day, seven days per week. To test and fix the year 2000 problems, the equipment must be stopped. This downtime costs time and money. According to a communication issued by AMR (Jesitus 1998), evidence indicates that the bill for addressing year 2000 problems at the plant floor level may be at least half of what a company spends to fix overall data centre issues. Unfortunately, if the year 2000 problems are not fixed, the company can loose more money and time from 1 January 2000. These date problems will most likely be manifested in subtle, rather than immediately catastrophic ways.

Other possibilities where problems might occur and testing that should be done include the following:

- Some older e-mail systems and mail gateways. It is suggested that companies upgrade to new versions that are year 2000 compliant. (Cole-Gomolski 1998). According to McNamara (1997), most of the problems are likely to be of the annoying variety, such as mislabelled headers, misorted in-boxes or premature license expiration notices.
- Mainframe-based mail systems, X.400 mail switches and directories (Cole-Gomolski 1998).
- Versions of Lotus Development Corporationís CC mail having a database called DB6. The problem with this system is that the software will delete messages and sort messages incorrectly after the year 2000 (Cole-Gomolski 1998).
- Banking transactions, including calculations of interest, transfer of funds between accounts (including debit and stop orders) and credit card transactions. Luckily, most banks in South Africa are well on their way to resolving the issue. This effort is driven by the South African Reserve Bank. The systems of the banks should be compliant by the end of November 1998. If not, they may be fined by the Reserve Bank. However, some smaller banks do not have the funds to fix their problems. One should ensure that one's banks comply with the regulations set by the Reserve Bank, which are:
  - Calculations of insurance premiums
  - Calculations of broker or mutual funds
  - In-house developed software
  - Spreadsheets. Not only should the spreadsheet package be compliant but also macros within the spreadsheet

that can use dates in calculations. These are found especially in accounting spreadsheets where dates are used to calculate escalations and budgets and forecasts.

- Telephone systems
- Delivery of goods
- Delivery of bills
- Delivery of government checks (*Computerworld* 1998)
- Stock Exchange computer systems
- Tax systems
- Hospital and life-support systems
- Airlines and train reservations
- Aviation systems
- VCRs and television sets.

According to minister Jay Naidoo (National Year 2000 Decision Support Centre 1998), 68 'mission critical' government systems have been identified. These include systems such as pension systems, examination systems, housing subsidies, transport, pay as you earn, UIF, population register, machine readable passport systems, and the drought disaster system.

There are three possible scenarios for unprepared systems (Carvell 1996):

- The system rolls over to 01-01-00 but assumes that the date is 1 January 1900 and not 1 January 2000. In systems that are not date sensitive, this is not critical. These are generally found in systems used in research environments where no transactions are created. In a high percentage of these cases, readings are taken and only elapsed time is recorded. Elapsed time does not normally use a date function but has a counter that counts milliseconds or seconds. The system is subsequently set to run for a period of time, sometimes even days or weeks. The counter counts until it gets to the specified time set.
- Instead of rolling over, the system crashes or halts. In this case, the system or computer does not know how to interpret the new date and halts. This can happen on any level discussed earlier. This means that the real-time clock, BIOS, operating system or application software discontinues to function. In a recent survey, a number of personal computers were found that halted when the clock ticked over to 1 January 2000. These personal computers were switched off, rebooted and started functioning again. A few of them even rolled over to 1 January 2000, but others rolled over to the year 00.
- The third possibility is that the system rolls over but when calculations are done, the system gives incorrect answers.

Corrective action

In addressing the problem, no fragmented approach will work. According to Hall (1998), the different stages in the conversion process include:

- Awareness of the problem
- Assessment of the extent of the problem
- Remediation to fix the problem
- Testing to ensure that the problem is resolved.

It is important to move past the assessment stage. Steps that can be followed to initiate the project are the following (Hall 1998): Take inventory of all the items that might have a problem. Estimate the cost to fix the problem. This can be done through estimating replacement values, a basic line-of-code count and estimating the cost to test and fix the code or to upgrade the item to a compliant version. Analyse the impact of the problem. Set priorities and plan the work. Select the most critical system and obtain the resources.

After an inventory has been conducted, Kador (1997) suggests that the list of all applications should be compiled and classified according to five categories, namely:

- Absolutely should have
- Need to have
- Nice to have
- Small-time
- Throwaway.

One example of a company that is following the steps and planning to address the year 2000 problem is Motorola Inc. At an Association for Manufacturing Excellence conference, Korus, an information technology employee at Motorola Inc., stated that the company is taking a responsible approach to the year 2000 problem. To minimise the impact of year 2000 problems, Motorola has developed the following six-phase plan (*Industry Week* 1998):

- Preliminary assessment
- Identifying all software applications that may be affected
- Determining the scope of the problem
- Conversion planning
- Conversion
- Development.

The process should be looked at more extensively than other normal problems. Moreover, it should be initiated at the vendors and their systems. If the vendors' systems are not compliant, there is a possibility that materials needed in the company to produce products will not be available, making production impossible. This process should also include provisions regarding the year 2000 issues in contracts, insurance to transfer risks and contingency plans.

It was found that less than 20 per cent of information systems groups have completed inventories of their personal computers and software. Most are testing each personal computer's BIOS for date turnover to check which personal computers could still be of use. This is done to save money and not to replace personal computers. Companies should make sure that they are not only checking personal computers but also spreadsheets and other software packages. However, testing may not be the hardest part. Replacing familiar software may take until the end of the millennium (Dryden 1998).

Testing the system and testing tools available

Malloy (1998) states that the test phase will represent 40 to 60 per cent of a company's year 2000 budget. In addition to the solid planning of the test phase, there are three stages of the testing phase of the year 2000 fix process, namely (Hammond & Carreon 1997):

- Functionality testing
- Year 2000 testing
- System testing.

Functionality testing ensures that changes made to fix the date problem do not affect existing production codes. Year 2000 testing involves running an application with data from the year 2000 to determine whether the fixes solved the problem. Testing can be done by taking existing data and ageing the data to specific predetermined dates. The minimum dates that should be tested for are:

- 9/9/1999
- 31/12/1999
- 1/1/2000
- 3/1/2000 (the first working day of the millennium)
- 28/2/2000
- 29/2/2000
- 1/3/2000

Year 2000 testing also requires finding ways to set the system clock forward as well as making sure that the operating system and vendor–supplier software on the system is year 2000 compliant.

According to El (1998), a growing number of companies are looking at automated testing and compliance tools to help accelerate the year 2000 conversion process. More and more testing tools are developed every day. Without these testing tools, the testing of larger systems in the limited time available is virtually impossible. Automation is essential to ensure consistent, repeatable results that make testing manageable. Testing tools range from tools that only test for dates in systems and reporting on these dates to tools that will change programs to increase the digits for dates to eight digits where the centuary is included. People should realise that these are only tools and the tools are there to help them with the testing of their systems and automate routine work and will not fix the year 2000 problem completely. The systems should still be tested manually and problems not fixed should be corrected by hand. Some of these tools will take data captured and age the dates with preset dates as defined by the user. One can then use these dates and data to test the programs and test the results to see if the systems can handle future dates. Criteria that date simulation tools should meet include the following (Malloy 1998):

- Intercept requests to all four date references
- Enable date selection from the mid-20th to the mid-21st century
- Require no changes to the system and job control language of the application
- Support all time/date formats
- Support all programming languages, including Cobol, PL/1, assembler, et cetera.

El (1998) states that some of the factors to consider when deciding what products or services to use or buy include the following:

- What types of systems you are running
- How many lines of code you have
- How much money the project will cost
- How long the project will take
- What resources are available.

Malloy (1998) argues that date simulation tools take care of an important part of year 2000 testing, but other testing tools, methodologies and environments are needed to conduct thorough testing. Time machines (stand-alone test machines with the systems clock set forward) are the most desirable but expensive option. Other options include lease of extra equipment, use of a service bureau or finding space and time on one's own machines.

Various tools and solutions are available on the market:

- Computer Associates International Inc. has a tool called CA-Fix/2000. It is a software fix tool that is available on a metered pricing scale based on the number of lines of code a company needs to change. The software fixes code on the mainframe and provides an audit trail so that information systems departments can see what was changed and what not (Weil 1998).
- Dallas Semiconductor and Resource 800 have launched a plug-in clock/CMOS RAM chip that changes the date field on desktop computers and file servers which, at this stage, can only handle a two-digit year.
- Cyberguard has a product called Millennium-Guard that adds year 2000 extension to the company's Certificate Authority and firewall products, protecting year 2000 compliant applications from contamination by noncompliant applications.
- Transcentury Analysis Tools of Platinum Technology Inc. uses parsing technology to help companies identify and analyse enterprise environments for faulty date occurrences. Frontier Corporation is using this tool for converting 21 million lines of code (El 1998).
- Turnkey 2000's Unravel 2000, an integrated code conversion tool, combines conversion technology with project management functionality and automates up to 90 per cent of the programmers' work (El 1998).

Certification level

According to the Software Engineering Institute in Pittsburgh (Yourdon 1998), approximately 75 per cent of US information technology organisations are at Level 1 on the 5-point Capability Maturity Model scale, which measures the maturity of an information technology organisation's software development process (refer to Figure 1). Fifteen per cent of US organisations are at Level 2 with a repeatable software process, and a little less than 10 per cent are at Level 3 where they could justifiably express confidence in the accuracy of their schedules. If an organisation has no metric about the performance of its previous software projects, there is no reason to believe that its performance on a year 2000 project should be any better than the industry average.

The Gartner Group created the 'COMPARE' (COMpliance Progress And REadiness) scale for year 2000 compliance and readiness and can be used as a standardised measure of an organisation's year 2000 progress and readiness.

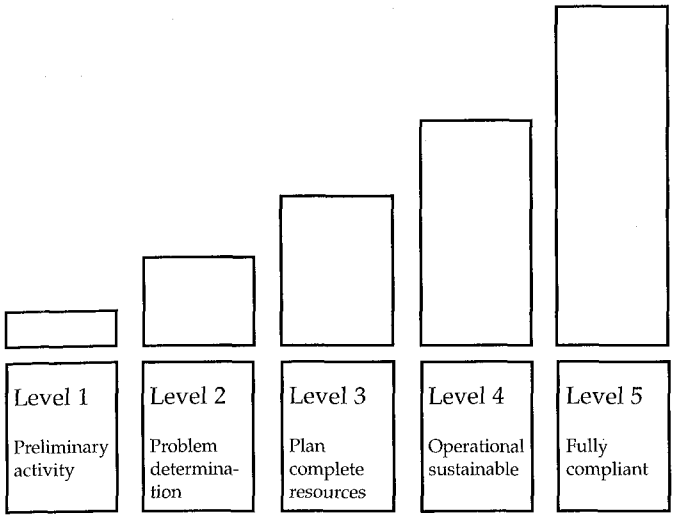

Figure 1. Compliance and readiness levels in the Capability Maturity Model (Gartner Group)

According to a survey of *Fortune* 500 companies (Hoffman 1997), only 20 per cent of the biggest companies in the United States have devised a full-fledged strategy to deal with the year 2000 date-change problem. Yet 87 per cent of the companies polled expect to have more than half their systems fixed by the end of 1998. Eighty-two per cent of the companies admitted they have underestimated the costs. The percentage of companies deferring new development work over the past eight months has increased from 11 to 18 per cent. This number is expected to rise closer to 31 December 1999. A survey of 1000 public and private sector companies in 15 countries found that half had yet to put forward a formal year 2000 programme (*Accountancy* 1998). A study conducted by the Gartner Group (1997) showed that 30 per cent of all companies worldwide have not yet begun to address year 2000 problems.

A nationwide survey done by the National Year 2000 Centre (Y2K Press Office 1998) indicated that 70 per cent of South African government departments and 90 per cent of large organisations anticipated completing their year 2000 conversions on time, while 80 per cent of local authorities will be ready. Minister Naidoo stated that the projected national compliance would be between 50 and 60 per cent, with 40 per cent of work expected to be completed in the last two quarters of 1999. The National Year 2000 Centre survey also showed that:

- 40% of government departments and 53% of local authorities have inadequately skilled government department project teams compared to 20% of large organisations;
- 36% of government departments and 40% of local authorities have inadequate action plans compared to 24% of large organisations;
- 42% of government departments and 39% of local authorities have disaster recovery plans in place compared to 78% of large organisations;
- 42% of government departments and 37% of local authorities have contingency plans compared to 60% of large organisations.

Possible solutions

According to Rock & Reynolds (1998), an estimated 15 per cent of software applications will not get fixed before the year 2000. Although there are software packages available to automate part of the year 2000 solution, there is no magic bullet to resolve the issue. Even with the use of these software packages, someone should still do the groundwork and unravel and rewrite a lot of tangled legacy computer code. Companies and individuals that can find or fix the millennium problem will be heavily in demand.

Castelluccio (1998) states that there are a number of ways to fix the year 2000 problem, some of them requiring years of very expensive code work. One of the more inventive quick fixes or methodologies is called 'windowing'. The name describes the process: creating a selected view that one can control. In this process a year is specified to use as the cutoff date. If the year in question is less than this year, a 19 is used to represent the century. If the year is the same as or follows this cutoff date, a 20 is used to represent the century. This cutoff date can be monitored and changed if necessary. However, the major drawback is that it only covers a span of 100 years. If dates outside this range are used, an alternative solution should be found. Another methodology is that of expansion. The expansion methodology requires that all date variables be increased in size to accommodate century information. It is a good idea to combine the two methodologies. If time permits, programs and data should be increased in size to accommodate century information. The windowing should be built in to handle data that were missed and not converted. If the windowing picks up data that were not converted, a warning or exception report should show or record this data.

Rock & Reynolds (1998) further argue that there is one across-the-board tactic to minimise the negative impact of the millennium bug. That is to keep paper records and save all receipts for the final few months of 1999 in order to reconstruct an account of errors in accounts detected after 1 January 2000. This should be done for all the accounts received including investments, mutual funds, bank and credit card transactions and insurance.

Unfortunately, bigger companies cannot benefit from this tactic. They process large volumes of transactions and it will be impossible to reconstruct all accounts. Many of these companies decided to replace their systems rather than to fix the problem. Some of them now find that time is running out and realise that they will not be able to install new systems in time. Hence, they are now trying to fix their legacy systems. This tactic of fixing systems is costing more and more because of the time and resource shortage and it will worsen as the year 2000 deadline approaches. Companies should thoroughly investigate and plan in order to minimise these problems. They should also learn from others.

According to Hoffman (1998), to help solve the problem some companies are making use of the Internet to tell customers which products are compliant or how to fix products that are not. One example is Microsoft Corporation's year 2000 resource centre that makes use of the World Wide Web (www.microsoft.com/year2000) to list products that are compliant, to show how products handle dates, to indicate how to work around problems and what tools are available to help fix individual products.

All companies should have contingency plans, even if their remediation programmes are going well. A Cap Gemini survey (*Computerworld* 1998) indicated that nine of the 25 companies had contingency plans in place as of the third quarter of 1996.

Many companies take too narrow a view of the year 2000 problem, concentrating only on application software. Contingency plans should be developed for application software packages, in-house developed software, outsource-developed software, internal and external infrastructures and external partners. One good approach is to have year 2000 contingency plans on top of existing disaster recovery plans, and, like all disaster recovery plans, year 2000 plans should be tested and evaluated under as realistic conditions as possible.

Conclusion

For businesses and individuals to survive, they will have to address the year 2000 problem and bring systems into compliance. This is not an option, but a necessity. There is more to making systems year 2000 compliant than widening the date fields by two digits or adding code that windows dates on the fly. Some organisations have not even finished assessing which software and hardware has a year 2000 problem. If there is anything to be learnt from this situation, it is that prior planning prevents panic.

Companies should not see the year 2000 problem as a problem that should be fixed as quickly and cheaply as possible. Management should commit themselves to addressing the issue and realise that this is not merely an information technology fad that will fade away in the year 2001. They should see this as a challenge and opportunity to make the technological mindshift that can give them the edge over their competition. Information technology should continue to be the business enabler for the new millennium.

References

Anthes, Gary H. 1998. 'Legal considerations in year 2000 contingency planning', *Computerworld*, 32(3): 83.

Bordwin, Milton. 1998. 'The millennium bug gets ready to bite', *Management Review*, 87(2): 27–29.

Caldwell, Bruce. 1997. 'Year 2000 cost estimates too low, reports say', *Informationweek*, 662(22 December): 32.

Carvell, Tim. 1996. 'Anatomy of a meltdown', *Fortune*, 134(4): 60.

Castelluccio, Michael. 1998. 'Y2K – the disaster viewed from a window', *Management Accounting*, 79(7): 58–59.

Cole-Gomolski, Barb. 1998. 'Year 2000 problem also threatens E-mail', *Computerworld*, 32(7): 4.

De Jager, Peter. 1997/1998. 'Entering the black zone', *Datamation*, 44(1): 39–40.

Delancy, Benjamin I. 1998. 'The Year 2000 problem: Are HR departments being left out in the cold?' *Employment Relations Today*, 24(4): 1–5.

Dryden, Patrick. 1998. 'Cleaning up LANs for year 2000', *Computerworld*, 32(3): 53–57.

El, Hakhi Alakhun. 1998. 'Conversion crunch', *Informationweek*, 663(5 January): 35–40.

Fryer, Bronwyn. 1997. 'Y2K stragglers', *Informationweek*, 660(8 December): 159–164.

Gibbons, Paul Lauren. 1997–1998. 'Y2K report card: How do you measure', *Datamation*, 44(1): 52–53.

Gross, Neil. 1997. 'Year 2000: The Meter's Running', *Business Week*, 3559(29 December): 41.

Hall, Bruce. 1998. 'Just do it!' *Software Magazine*, 18(2): 29–36.

Hammond, Eric & Carreon, Julia C. 1997. 'Cure for the uncommon test', *InfoWorld*, 19(51/52): 58–61.

Hock, Steven L. 1998. 'Legal considerations for solving the Year 2000 software crisis', *Information Strategy: The Executive's Journal*, 14(3): 15–20.

Hoffman, Thomas. 1996. '"Don't ask, don't tell" becomes year 2000 avoidance policy', *Computerworld*, 30(40): 12.

Hoffman, Thomas. 1997. 'If wishes were horses', *Computerworld*, 31(51): 2.

Hoffman, Thomas. 1998. 'Microsoft issues mea culpa on Y2K, readies resource center', *Computerworld*, 32(4): 6.

Internet. www.cinderella.co.za.

Jesitus, John. 1998. 'The real year 2000 nightmare: Manufacturing systems', *Industry Week*, 247(1): 22–26.

Kappelman, Leon. 1996. 'Year 2000 upgrades: A small price to pay', *Computerworld*, 30(48): 33.

Kappelman, Leon A, Fent, Darla, Keeling, Kellie B. & Prybutok, Victor. 1998. 'Calculating the cost of year-2000 compliance', *Communications of the ACM*, 41(2): 30–39.

Kelley, Edward W. Jr. 1998. 'The millennium bug', *Executive Speeches*, 12(4): 20–24.

Kimbark, William J. 1997. 'Are you ready for the year 2000?', *HR Magazine*, 42(6): 41–46.

'Making the year 2000 a national priority'. 1998. Seminar. CSIR.

Malloy, Amy. 1998. 'Testing tools: It's tomorrow – today', *Computerworld*, 32(3): 87–88.

Malloy, Amy. 1998. 'Tools, etc.', *Computerworld*, 32(3): 89.

McArthur, Larry. 1997. 'Countdown to 2000: Millennium bug risk assessment and management', *Risk Management*, 44(10), R6, R14.

McNamara, Paul. 1997. 'Most e-mail stamped OK for Y2K', *Network World*, 14(50): 40.

Radosevich, Lynda. 1998. 'Millennium bug already taking its toll', *InfoWorld*, 20(2): 1, 19.

Rock, Andrea & Reynolds, Tripp. 1998. 'The year 2000 bug', *Money*, 27(2): 48–52.

Scheier, Robert L. 1998. 'The pressure builds', *Computerworld*, 32(1): 30.

Sheridan, John H. 1998. 'Dealing with 'Y2K'', *Industry Week*, 247(1): 27–28.

'The year 2000 problem – what problem?' 1998. *Accountancy*, 121(1253): 52.

Weil, Nancy. 1998. 'CA fires off fix of 2000', *Computerworld*, 32(2): 20.

'Year 2000 scoreboard'. 1996. *Computerworld*, 30(40): 73–76.

Yourdon, Ed. 1998. 'Where's the basis for year 2000 optimism?' *Computerworld*, 32(7): 68.

In search of middle ground amidst the moral mess of worker participation

GJ Rossouw
Rand Afrikaans University

The focus of this contribution to the discourse on worker participation is on the moral dimension thereof — i.e. on the moral arguments raised to promote and oppose worker participation. Worker participation can be moulded into many forms. This paper will commence with a definition of what worker participation entails within the South African context. The most prominent moral arguments in support of and in opposition to worker participation will be identified. The objective of this paper is to determine against this backdrop whether there exists moral middle ground between the opponents and supporters of worker participation. It will be demonstrated that such middle ground indeed exists, but also that some irreconcilable differences remain. These differences provide the stimulus for an ongoing creative discourse on worker participation.

Introduction

Already at the beginning of the 1980s, researchers claimed that the staggering volume of literature on the topic of worker participation, which extends over centuries, is 'galloping out of control'. It was then also asserted that 'we may indeed be far beyond the point of any summarising overview, at least within the confines of any portable volume' (Alexander 1985: 337). Since then the debate on worker participation has escalated further. If these researchers were only vaguely correct then, they are acutely correct a decade and a half later. The volumes of scholarly and popular work on this topic render any comprehensive discourse on the topic a futile exercise. Contributions to this ongoing discourse need to be very well focused if they aspire to be meaningful.

In this article, a very specific niche in the discourse on worker participation will be examined. The focus will be on the moral dimension, i.e. the moral arguments raised to promote and oppose worker participation. However, worker participation can be moulded into many forms. The article will therefore commence with a definition of what worker participation entails within the South African context. Then, the most prominent moral arguments in support of and in opposition to worker participation will be identified. The objective is to determine whether there exists moral middle ground between the moral opponents and supporters of worker participation.

Worker participation in South Africa

In the international discourse on worker participation, a number of definitions and distinctions are made which are useful for a better understanding of the form and nature of worker participation in South Africa. These distinctions will be introduced briefly and their relevance for the South African situation will be indicated.

Names and definitions

The concept 'worker participation' is referred to by a number of names. These include the following: employee empowerment, participative or participatory management, codetermination, job enrichment, workers' control, and workers' self-management – to mention only the most prominent ones. Common to all these terms is the notion that worker participation consists of involvement of workers in decision-making processes that traditionally have been the responsibility and prerogative of managers (cf. Robson 1981: 4). In defining the concept of worker participation, it is useful to steer clear of definitions that are too wide, such as the one by Brannen,

SBL Research Review 1998 2(2): 9-14
In search of middle ground amidst the moral mess of worker participation

Batstone, Fatchett & White (1976: 1) who defines it as an increase in the level of worker influence upon management. The flaw of this definition is that it includes related phenomena such as collective bargaining and industrial democracy. For the sake of clarity, it is important to keep conceptual distance between these concepts.

Pateman was the first to introduce the distinction between worker participation and industrial democracy. According to him, industrial democracy would imply the electability and removability of persons in positions of authority. It thus democratises the entire organisation including authority structures. This inclusion of democracy in the workplace is what Iris Marion Young also advocates in her book, *Justice and the politics of difference*. A distinction drawn by Robert Archer (in Pagano & Rowthorn 1996: 29) is helpful for understanding the difference between industrial democracy and worker participation. He distinguishes between *control through ownership* and *control against ownership*. The former refers to a situation where workers gain control over the enterprise by becoming the owners (or majority shareholders) of the enterprise. The workers thus gain the majority vote on the board of directors and can, in principle, challenge any structure of authority or control within the enterprise. This control through ownership is sometimes labelled as worker participation, or is also referred to as ownership plans or ESOPs (Employee Stock Ownership Plans). In this article, Archer's control through ownership will not be included in the definition of worker participation. It will rather be seen as one route to industrial democracy. (According to Robert Dahl and Iris Marion Young there are also other routes.)

What Archer terms *control against ownership* is what will be referred to as worker participation in this article. In worker participation positions of authority are not at stake. Worker participation only deals with involvement in decision making within the existing framework of authority and control within an economic enterprise. Therefore, the basic structure of the capitalist system, where the majority of shareholders can, in principle, determine the structure and positions of authority, remains unchallenged by the concept of worker participation.

There is, however, a form of *control against ownership* that will be excluded from the definition of worker participation and that is collective bargaining. It cannot be equated with worker participation as it is not directed towards participation within the existing framework of decision making within enterprises, but rather exerts influence from the outside via union activity and pressure on management and corporate decisions. Table 1 summarises the above distinctions.

Forms of worker participation

Once worker participation is distinguished from related concepts, certain further distinctions can be drawn within the conceptual scope of the concept itself. These distinctions relate to the width of the concept. A popular distinction in this regard is the distinction between *task-centred* participation and *power-centred* participation – a distinction first introduced by Robert Salamon (cf. Van Niekerk & Le Roux 1995: 52; Maller 1992: 10). Power-centred participation refers to the levels at which workers have access to the decision-making process. A variety of levels can be distinguished in this regard, such as decision making on the level of supervisors, middle management, senior or top management or even on the level of the board of directors. Task-centred participation, on the other hand, refers to the issues on which workers can participate in decision making. There obviously is a close relation between these two distinctions as certain issues are dealt with on certain levels of management or governance.

Even within the distinction of power-centred participation, a finer distinction can be drawn with regard to the degree of power or control that can be exerted there. The degree of control can vary from minimal to substantial control on a certain level (cf. Bernstein 1976: 492). Minimal control will entail the right to be consulted on certain issues, while substantial control will entail the right to joint decision making or even a right to veto on certain decisions.

Worker participation and the South African Labour Relations Act of 1995

Against the backdrop of the above distinctions, the legal position in South Africa with regard to worker participation can be summarised as follows. The Labour Relations Act of 1995 makes a clear distinction between collective bargaining and worker participation and it allows for both. However, it does not make provision for industrial democracy. In chapter five of the Act, which deals with workplace forums, the scope of worker participation is clearly defined. It takes a task-centred approach to worker participation and only allows for minimal power-centred participation, because it determines that workers only need to be *consulted* on issues such as the following:

- restructuring of the workplace (introduction of new technologies and new work methods)
- changes in the organisation of work
- plant closures
- mergers and transfers of ownership

Table 1. Distinctions between concepts

	Industrial democracy	Worker participation	Collective bargaining
Full internal control			
Involvement in internal control			
External influence on control			

- dismissal of employees
- job grading
  - criteria of merit increases or payment of discretionary bonuses
  - education and training
  - product development plans
  - export promotion.

The Act allows for more substantial power-centred participation in the following tasks when it determines that *joint decision making* should take place on:

- disciplinary codes and procedures
- rules regulating the workplace that apply to conduct not related to work performance
- measures for advancing persons disadvantaged by discrimination
- changes to rules regulating social benefit schemes.

The employer is also obliged to *disclose information* to the workplace forum that will allow it to engage effectively in consultation and joint decision making.

Worker participation has been legislated into South African workplaces, and it takes the form of mostly task-centred participation coupled with fairly weak power-centred participation. This, however, is merely the legal status of worker participation. The moral justification of worker participation is a different issue, as there always remains a critical distance between the legality and morality of phenomena. It is to this moral justification that the focus will now shift.

The moral dispute on worker participation

The most prominent arguments for and against worker participation will now be raised. The purpose of this exposition of arguments is to unpack the dispute in order to determine whether opponents and proponents of worker participation might be able to establish some minimal form of moral consensus on the issue. It is thus done in the quest for moral middle ground on this issue.

Two objections might be raised against this approach, the first being an argument that is generally associated with the New Left. The objection will be, more or less, that it is senseless to focus on worker participation alone, because it still leaves the capitalist system unquestioned. What is required, according to this school of thought, is a much more profound challenge of the capitalist system as such, as well as the mechanisms of authority and control associated with it.

This objection is legitimate, but if taken seriously, also means the end of any discourse whatsoever on worker participation. The only kind of discourse that will be acceptable from this frame of mind is to reflect on how worker participation can contribute in preparing the way for an altogether different economic system with new mechanisms of authority and control (cf. Greenberg 1975: 200–204). The kind of discourse that is required is thus no longer a discourse on worker participation, but one on industrial democracy. This topic falls beyond the scope of this article and will therefore not be pursued further.

The second objection against this approach might come from those who support the idea of worker participation, but find themselves in disagreement with the form that it was given in the Labour Relations Act of 1995. This also is a tenable position – to support the idea of worker participation, but to disagree with the shape it has been given by the state. However, this discrepancy will have no bearing on the kind of moral arguments that will be formulated in the remainder of this section. Persons with this objection will not disagree with the kind of moral arguments that will be raised below. Their objection would rather be that the current state of affairs does not sufficiently accommodate their moral concerns.

Moral arguments in support of worker participation

Through literature research the following four categories of moral arguments were found in support of worker participation.

The anti-adversary argument

The traditional form into which economic enterprises organised themselves, according to this argument, is one of antagonistic relations between workers and owners of capital invested in business. In this frame of thinking, managers are seen as the representatives of the owners of capital. They thus align themselves with the owners of capital against the workers. This inevitably results in adversarial relations between workers and managers. This reality is well reflected in the adversarial collective bargaining approach which seems to be the rule, rather than the exception – also, and specifically so, in South Africa (cf. Van Niekerk & Le Roux 1995: 54; Pillay 1987: 60).

The moral argument in support of worker participation claims that it can contribute to transcending this adversarial relationship between workers and managers by involving them in management decisions. It will give workers and management the opportunity to build mutual understanding. This, in turn, will result in more harmonious interaction between workers and management.

The self-actualisation argument

This argument starts with the assumption that it is not only people who produce in the economy, but that the economy also produces people; i.e. how people work within economic enterprises determines how their natural capacities develop and also how they regard themselves. Various psychologists and sociologists regard workplace experiences as the cause for alienation and stunted personality development that rob employees of creativity, autonomy, and personal development (cf. Greenberg 1975: 194–197; Gini & Sullivan 1993: 132–138; Neilsen 1993: 151–155).

Maslow already did some experimental work in this regard that showed the correlation between a positive working environment and self-actualisation (cf. Greenberg 1975: 196). This positive correlation is also confirmed by other social scientists (cf. Giliomee 1990: 64; Young 1990: 27). The moral argument in support of worker participation claims that it will contribute towards reorganising the workplace such that it will become a more meaningful environment for workers. Worker participation will give workers the opportunity to have a voice in the creation of their working conditions. This inevitably will lead to more humane and interesting work that, in turn, will be conducive to the self-actualisation of workers within the workplace.

The political participation argument

Another strand of the self-actualisation argument focuses on the capacity of workers to participate in political life outside the workplace. A direct link is drawn between workplace organisation and the capacity to participate effectively in social processes and political life. According to this line of argument, hierarchical forms of production breed passivity, apathy, and ignorance which, in turn, undermine active citizenship in society at large. It is claimed that in order to participate effectively in political life, one's capacity for participation in decision making needs to be enhanced and developed in the workplace (cf. Pagano & Rowthorn 1996: 2; Young 1990: 38; Greenberg 1975; Dahl 1985: 54; Adamson 1990: 57). Worker participation and the opportunities it offers for participation in decision making once more is seen as a way of building this capacity for political participation.

The economic growth argument

The economic argument is based upon the assumption that the success of economies depends on their ability to promote consistent productive economic activity. This, in turn, requires cooperation and commitment within economic enterprises. It is believed that worker participation can foster such cooperation, responsibility, and trust and discourages 'free-riding' (cf. Kenworthy 1996: 51–56; Hickson & Oldham 1990: 95; Giliomee 1990: 64).

The basis for this conviction is closely related to the first two moral arguments already mentioned. Should worker participation succeed in breaking down adversarial relations within the workplace and should it further facilitate human development within the workplace, it is indeed not unrealistic to expect that the creative and productive potential of workers will be unleashed to a greater extent. It would also not surprise if it leads to more cooperation, trust, and commitment. This, in turn, can translate into improved productivity and, ultimately, economic growth. There is also empirical evidence to support this claim (cf. Alexander 1985: 343; Pagano & Rowthorn 1996: 209).

The moral dimension of the economic argument is provided by the expectation that the population at large will benefit from economic growth, because it can curb unemployment and some of the excesses associated with unemployment, such as poverty and soaring crime.

The equity argument

The equity argument takes its point of departure in the employment contract. That is the agreement in which workers surrender their productive capacity to an employer or manager in exchange for a wage. It is argued that this employment contract is authoritarian in nature and grants the employer/manager excessive power over the worker. The argument is often raised that the right of exit of the worker can curb this power of the employer/manager. In developing countries, like South Africa, with a high unemployment rate, such exit is extremely costly to the worker – if not unaffordable (cf. Pagano & Rowthorn 1996: 72; Dahl 1985: 115). It is therefore no real option for most workers.

Against the backdrop of this inequality that is built into the employment contract, it is believed that worker participation can be an important correction to restore some sense of equity in the workplace (cf. Deale 1995: 19). It can give workers a voice and the opportunity to bring their own perspectives and interests to bear on at least some of the decisions that affect them within the workplace. Worker participation can, in this way, restore some of the capacities that workers surrender when they enter into the employment contract.

Moral arguments in opposition to worker participation

The same literature research revealed the following three categories of moral arguments in opposition to worker participation.

The efficiency argument

The assumption underlying this argument is that the economy is driven by sound competition between economic enterprises. The competitive market, which is the result of this competition, is deemed to be the most efficient allocater of resources. Any factor that might impede the efficiency of economic enterprises or the competitive market system is thus regarded as unwanted and unacceptable.

Worker participation is seen as one such factor that can impede the competitive nature and competitive edge of economic enterprises. The reason for this expectation is that it can slow down decision-making processes within enterprises, because joint decision making between workers and management can be a slow and disruptive process. Furthermore, it is feared that workers might introduce concerns into the decision-making process that might distract economic enterprises from their main goal, which is the efficient delivery of goods and services to the market (cf. Block 1995: 409; Bernstein 1976: 512).

The moral dimension in this efficiency argument is provided by the conviction that it will be detrimental not only to the wellbeing of all employees in an economic enterprise should it lose its efficiency in the market, but also to the economy at large should there be a general decline in the efficiency of the economy because of worker participation. The argument thus is that worker participation can, despite its noble intention to assist workers, have the opposite effect of damaging the interest of workers and all other stakeholders should it result in less efficient performance within the market economy.

The union argument

Whereas the efficiency argument is mostly associated with managers and the owners of capital, the union argument is associated with workers, and more specifically workers who find themselves in organised labour unions. Unions, and their ability to challenge management, are seen as one of the great achievements of organised labour. It is one of the most effective mechanisms available to workers to force management to pay due consideration to the concerns of workers. Therefore, it is no surprise that there is a strong commitment to the protection and continued existence of unions amongst workers.

The argument against worker participation from workers themselves is that it can undermine the influence of unions. It is feared that worker participation can compete with unions and that it can eventually dilute the power base of unions. There is a particularly strong suspicion that management can deliberately exploit worker participation as a means of sabotaging union influence on the workplace (cf. Von Holdt 1995: 59–61; Evans 1992: 44; Levitan & Werneke 1984: 31). From this perspective, worker participation is seen either as unwanted

competition for unions or as a Trojan Horse used by management to the detriment of unions and workers in the long run.

This argument gains its moral dimension from the concern that worker participation might in the long run be detrimental to the interests of the workers who already find themselves in an unfavourable position via the employment contract. Worker participation is seen as a dangerous exercise that can rob workers of their most reliable instrument for safeguarding their interests, which is the labour union.

The sufficient protection argument

Another argument against worker participation regards it as superfluous because there already exists sufficient protection for workers in the workplace. According to this argument, the combination of a competitive market economy and liberal democratic state is more than sufficient to safeguard the interests of workers. It is believed that the state protects the rights of workers through the legal and constitutional framework, as well as through government controls and regulations on safety, minimum wages, et cetera. The market complements this protection of workers by granting them the freedom to associate or disassociate with whichever enterprise they deem fit or unfit. This protection of workers is further strengthened by the policy of social responsibility that most economic enterprises voluntarily subscribe to nowadays. This entails that companies commit themselves to interact in a morally responsible way with all their internal and external stakeholders, including their employees (cf. Pagano & Rowthorn 1996: 64; Simon 1983: 37–39; De Vos 1981: 44–49).

Against this background, worker participation is regarded as an unnecessary burden on economic enterprises, as the above measures make more than enough provision for the moral treatment of workers within economic enterprises. The proponents of this argument bemoan the fact that the primary control mechanism of politics, which is the vote, is forced upon the economic domain as a control mechanism for management decisions. They believe that the economic domain differs so substantially from the political that it warrants rather different control mechanisms in the economy (cf. Simon 1983: 39).

Moral middle ground?

Are the positions of the proponents and opponents of worker participation so mutually exclusive that the two can never meet, or can middle ground be found in this moral dispute? It is contended that there is sufficient moral middle ground available in this battle, although some rivalry will always remain, even when the middle ground has been discovered. The remaining rivalry need, however, not be destructive, but can be harnessed creatively.

Two concerns were raised in the dispute that need to be dismissed as unrealistic. The one is a concern mostly raised by workers whereas the other is mostly raised by managers or owners of capital. The latter will be discussed first.

The concern expressed in the *sufficient protection argument* needs to be dismissed as unrealistic. The legal provisions for workers in the constitution and government regulations are not sufficient to abolish the structural inequality of the employment relationship. This is even more acutely so in a labour market such as that of South Africa, where there is an oversupply of labour and a shortage of job opportunities. To rely blindly on the law and the market in such circumstances simply is unrealistic. Such a position suffers all the blind spots of the Hidden Hand dogma that Stone exposed in his epic debate with Friedman in the social responsibility debate (cf. Stone 1992; Friedman 1992). Managers and owners of capital need to realise that workers do not find themselves on equal footing with managers in the workplace and that the needs, interests, and concerns of workers have to be accommodated to a greater extent than is currently the case. This does not mean that the authority of managers should be discarded. It only means that more recognition needs to be given to the voice of workers.

A second concern that needs to be dismissed is the one expressed especially by union leaders in the *union argument* where worker participation was portrayed as a means to undermine unions and collective bargaining. Worker participation and collective bargaining do not need to stand in an either–or relationship. There can be a mutually complementary relationship between them. In order for this to occur, two kinds of change of heart are required. Managers and owners of capital should respect the role of unions and collective bargaining and should not use worker participation as a union-bashing instrument. Union leaders, on the other hand, should similarly not regard any attempt other than collective bargaining as a threat, because that would result in unions becoming a goal in themselves and no longer a vehicle for improving the situation of workers. They should also give due consideration to the fact that not all workers belong to unions and allow such workers the freedom to express their concerns via mechanisms other than unionised collective bargaining.

When these two concerns have been dismissed, quite a lot of moral middle ground becomes visible between the two sides in the moral dispute. If the proponents of worker participation:

- are correct in their claim in the *anti-adversary argument* that it can bring greater harmony into workplaces, and
- are correct in their claim in the *self-actualisation argument* that it can stimulate creativity, autonomy, and personal development in workers, and
- are furthermore correct in their assessment in the *economic growth argument* that it can result in improved productivity,

then it is hard to see why managers and owners of capital would resent worker participation. To the contrary, management theory encourages managers to facilitate harmony, creativity, autonomy, personal development, and improved productivity in workers (cf. Alexander 1985: 338; Broadwell 1995: 64; Levitan & Werneke 1984: 28; Hickson & Oldham 1990: 95). Instead of being a battlefield between workers and managers, worker participation can be an instrument for achieving mutually beneficial goals.

The *political participation argument* also does not need to divide proponents and opponents of worker participation. Any attempt to invigorate political participation that, in turn, can strengthen the legitimacy of the political process needs to be welcomed and appreciated by managers and owners of capital. The recent unsavoury history of political instability caused by an illegitimate political system and the negative impact thereof on the economy, should convince managers to embrace any attempt to strengthen the legitimacy of the current political dispensation in South Africa and participation of workers therein.

Despite this moral middle ground that has been uncovered, two seemingly irreconcilable concerns in the moral dispute

remain; i.e. workers' concern for equity in the workplace and management's concern for efficiency in the workplace. This tension cannot be resolved easily, as it is a realistic concern that the demand for equity can undermine efficiency, just as the demand for efficiency can undermine equity. Both equity and efficiency should be regarded as legitimate concerns for the reasons mentioned in the *equity argument* and *efficiency argument* respectively.

In order to find a way out of this impasse created by the tension between efficiency and equity, two ways of dealing with it will be suggested in concluding this quest for middle ground. The one route that could be taken is to give workers a much more direct stake in the efficiency of the economic enterprise in which they are employed. This can be done by means of a profit-sharing scheme, either in the form of direct payouts, or by stock-owning options. In this way, workers will benefit directly from greater efficiency. It will thus become less likely that their concerns for equity will run counter to the managers' concern for efficiency. One could rather expect a form of worker participation in which a sound balance will be struck between equity and efficiency. A study by Bernstein (1976: 498) on the success of worker participation schemes indeed found that the most successful ones are those in which participating workers were given a return of the surplus above their regular salaries.

The second route that could be taken is to cultivate certain values in the workplace that will facilitate worker participation in which both equity and efficiency can flourish. For equity to be respected, certain ethical values need to be fostered amongst workers and managers alike. Ethical values such as tolerance, justice, and respect for the dignity of each employee are needed in this regard. For the promotion of efficiency, work values such as commitment, reliability, and punctuality need to be cultivated. Laws cannot promulgate these ethical and work values. They can only be promoted within the confines of the economic enterprise (cf. Rossouw 1997). But once they are there, they might be the best safeguards for securing a fruitful interaction between the values of equity and efficiency.

Conclusion

In this article, the concept of worker participation was defined and then distinguished from two related concepts, namely collective bargaining and industrial democracy. The most prominent arguments in support of and in opposition to worker participation were then explored. Against this background, an attempt was made to find common moral ground between proponents and opponents of worker participation.

Worker participation can be seen as an attempt to bring democracy into the workplace without democratising the structures of authority within economic enterprises. There is, however, an international trend in the making that seeks to do exactly this (cf. Block 1995; Young 1990; Dahl 1985). This dispute on whether democracy should be brought to bear on the structures of authority within the economy, opens up even more complex problems than the ones that were analysed in this article. This phenomenon will invite further moral investigation.

References

Adamson, W. L. 1990. 'Economic democracy and the expediency of worker participation', *Political Studies*, 38 (Mar): 56–71.

Alexander, K. O. 1985. 'Worker ownership and participation in the context of social change', *American Journal of Economics and Sociology*, 44(3): 337–347.

Bernstein, P. 1976. 'Necessary elements for effective worker participation in decision making', *Journal of Economic Issues*, 10(2): 490–522.

Block, R. N. 1995. 'Labor Law, Economics, and Industrial democracy: a reconciliation', *Industrial relations*, 34 (Jul): 402–416.

Brannen, P., Batstone, E., Fatchett, D. & White, P. 1976. *The worker directors: a sociology of participation*. London: Hutchinson.

Broadwell, M. M. 1995. 'Why command and control won't go away', *Training*, 32 (Sept): 62–68.

Dahl, R. A. 1985. *A preface to economic democracy*. Berkeley, CA: University of California Press.

Deale, P. 1995. 'Meeting eye to eye: the workings of workplace forums', *People Dynamics*, 13(9): 18–23.

De Vos, T. 1981. *U.S. Multinationals and worker participation in management: the American experience in the European community*. London: Aldwych Press.

Evans, G. 1992. 'Worker participation at PG Bison', *South African Labour Bulletin*, 16(3): 40–51.

Friedman, M. 1992. 'The social responsibility of business is to increase its profits'. In J. Olen. & V. Barry, *Applying ethics*. Belmont, CA.: Wadsworth Publishing Company: 433–438.

Giliomee, H. 1990. 'The challenge of industrial relations in the 1990s', *Industrial Relations Journal of South Africa*, 10(1): 58–68.

Gini, A. R. & Sullivan, T. 1993. 'Work: the process and the person'. In T. I. White (ed.), *Business ethics: a philosophical reader*. New York: MacMillan: 132–140.

Greenberg, E. S. 1975. 'The consequences of worker participation: a clarification of the theoretical literature', *Social Science Quarterly*, 56(2): 191–209.

Hickson, M. & Oldham, G. 1990. 'Ten years of workers, participation in a South African producer co-operative', *Africa Insight*, 20(2): 93–99.

Kenworthy, L. 1996. 'Balancing competition and cooperation', *Challenge*, 39(4): 51–58.

Levitan, S. A. & Werneke, D. 1984. 'Worker participation and productivity change', *Monthly Labour Review*, 107 (Sept): 28–33.

Maller, J. 1992. *Conflict and co-operation*. Johannesburg: Ravan Press.

Nielsen, K. 1993. 'Alienation and work'. In T. I. White (ed.), *Business ethics: a philosophical reader*. New York: MacMillan: 151–155.

Pagano, U. & Rowthorn, R. 1996. *Democracy and efficiency in the economic enterprise*. London: Routledge.

Pillay, P. 1987. 'Worker control of enterprises: some theoretical considerations', *Industrial Relations Journal of South Africa*, 7(1): 60–70.

Robson, M. P. 1981. *Worker participation in the United Kingdom*. Bradford: MCB Publication.

Rossouw, G. J. 1997. 'Etiese kodes in ondernemings: ontwikkeling en implementering', *Koers*, 62(3): 319–331.

Simon, H. A. 1983. 'What is industrial democracy?' *Challenge*, 25 (Jan–Feb): 30–39.

Stone, D. S. 1992. 'The corporate social responsibility debate'. In J. Olen & V. Barry, *Applying ethics*. Belmont, CA.: Wadsworth Publishing Company: 438–446.

Van Niekerk, A. & Le Roux, P. A. K. 1995. 'Worker participation – sharing the right to decide", *Contemporary Labour Law*, 4(6): 51–56.

Von Holdt, K. 1995. 'Workplace forums: undermining unions?', *South African Labour Bulletin*, 19(6): 59–64.

Young, I. M. 1990. *Justice and the politics of difference*. Princeton: Princeton University Press.

Reaction to video technology for distant management training in South Africa

V Makin

Unisa Graduate School of Business Leadership

In a first-time South African management research study a total of 100 MBA/MBL students from the University of Pretoria (a residential university) and Unisa (a distance educator) took part in a television lesson taught either by prepackaged video or live interactive teleteaching. The students' reaction to television tuition was surveyed and five hypotheses were tested and rejected. Of the total sample group, when presented with a choice of whether they held a positive or negative view towards the new technology, 98 per cent of the respondents described their reaction as positive, while two per cent held a negative view. Differences were found within and between all groups. All the groups seemed to have a preference for the interactive television method of tuition over the prepackaged video. The study indicated that further research is needed on how best to use multimedia technology to enhance training, as new and better management skills are needed in South Africa.

Introduction

There is a vast shortage of suitably qualified managers in South Africa. At the same time the cost of education is increasing and traditional teaching methods are no longer adequate. However, although there are a number of technologies that can be of use, human resources professionals are not always clear which are the best and what the trainees' attitude to them will be. While there is much overseas research in this area, there has been limited research carried out within South Africa. South Africa has some unique infrastructure factors to consider. Demands such as 'affirmative action' from the African National Congress and the Black Management Forum are also putting a new emphasis on management training.

The nature of management in South Africa

The future of management training is in knowledge and high-level skills. In a survey reported by Hofmeyr, Rall & Templer (1995), 205 human resources professionals state that affirmative action is a priority in South Africa, but almost four out of five human resources managers feel that affirmative action progress is too slow. In a further study on strategic business issues, Perry & Associates (1996) found that 44 chief executive officers indicated affirmative action as the second most important issue after global competitiveness. The South African economy has been criticised as there is too much dependence on white males to fill highly skilled occupations.

The Breakwater Monitor

The Breakwater Monitor national database and information service was established in 1991 as a partnership between progressive South African employers and the University of Cape Town Graduate School of Business. The purpose of this longitudinal research service is to provide reliable labour market information, with the emphasis on tracking the implementation of employment equity practices in South Africa.

Rate of change: April 1994–March 1996

Management

A total of 64 organisations were selected to participate in a two-year longitudinal study (April 1994–March 1996). The sample was selected from those organisations that consistently reported stock and flow data for each of the six monthly reporting periods during the two years. A breakdown of management (defined as Paterson grades D, E and F including

SBL Research Review 1998 2(2): 15-24
Reaction to video technology for distant management training in South Africa

executive directors, senior executive management, senior management, middle-senior management and junior-middle management) by race over the two-year period is provided in Table 1.

Table 1. Management by race

Race	1994	%	1996	%
Asian	782	2.2	1049	2.8
Black	1003	2.9	1830	4.9
Coloured	603	1.7	955	2.6
White	32608	93.2	33274	89.7
Total	34996	100.0	37108	100.0

Blacks, Asians and Coloureds constitute the vast majority of the population in South Africa but this is not reflected in management. As indicated in Table 1, in 1996 only 10% of management came from the Asian, Black and Coloured population while 90% came from the White population.

When one looks at the situation with regard to executive directors (Paterson grade F) from September 1994 to March 1996 the picture is bleaker. If one compares the figures over this period, there is a dependence on white males to fill this category. In September 1994, 98.73% of executive directors were white males. The equivalent percentage for March 1996 is 96.96%. This clearly indicates that most of the increase in Black, Asian and Coloured representation at management level has taken place at lower levels than that of executive director.

Government policy

According to the White Paper on Reconstruction and Development (1994: 21):

> The Government's economic policies require human resources development on a massive scale. Improved training and education are fundamental to higher employment, the introduction of more advanced technologies, and reduced inequalities. Higher labour productivity will be the result of new attitudes towards work and especially new skills in the context of overall economic reconstruction and development. New and better management skills are urgently required.

The above background shapes this study in that how best to use advanced educational technologies in South Africa is the question facing human resources professionals.

The sample

The Unisa Graduate School of Business Leadership (SBL), a distance educator, has about 400 trainees for their MBL 1 course, of which about 80 come from the Pretoria area. The University of Pretoria Business School of Management, a residential university, has 250 MBAs with 103 students in their first year. All sample groups were drawn from the Pretoria area. The University of Pretoria (UP) trainees are broken down into 43 part-time (PT) and 60 full-time (FT) trainees. Of those, 29 part-time and 31 full-time took part in the experiment. There were 10 Unisa Pretoria groups of which 7 groups were selected, which consisted of 40 MBL students taking part in the experiment. Therefore, the total sample size was 183 of which 100 MBA/MBL trainees took part in the experiment.

The case study method is a common learning tool of tuition in both institutions. At the SBL and University of Pretoria the trainees are usually put into groups on the basis of geographical location but with some awareness that there is a mix of business expertise. This was a convenient sample although both groups had a vested interest in being part of the experiment as it was part of the MBA/MBL course. Of the sample 85% were male and 75% were Afrikaans–speaking, with most being between the ages of 25 and 29 years (45%). This sample group consisted of many busy executives studying mainly part-time and thus time for tuition was valuable and could not be wasted.

Background to the video programme

Prepackaged video – The case study method

The Unisa Graduate School of Business Leadership developed a video programme with the Department of Educational Technology (DET) on an American pharmaceutical company and the company's decision to increase or decrease investment in South Africa. It is a support video to the case study 'Medcare' which is available on request. The video was made in 1994 and updated and re-edited for this research in 1995. The video is available for anyone who wishes to view it. The video programme is 30 minutes in duration and consists of 3 sections:

- Lesson on the case study method by the instructor.
- Debate by the groups.
  - The group receives instructions to debate whether to invest in South Africa or not after viewing a 1994 group reaching a decision.
  - They turn off the television and hold their own debate.
  - They come to a conclusion.
  - The alternatives are then summarised on video by the instructor.
- An interview with a financial expert.

Teleteaching

Teleteaching is a combination of one-way video and two-way audio. At the University of Pretoria, teleteaching started broadcasting in the Pretoria area and to their satellite campus at Witbank. Teleteaching is a Limited Interactive Television (LITV) that comprises the use of microwave/fibre optical and/or cable television to transmit video and audio signals to remote sites and allows students to respond via telephone lines or on audio systems. This system is substantially cheaper than Unlimited Interactive Television (UITV) that involves television studio classrooms at both origination and remote sites, allowing audio and visual interaction between tutor and learner at all sites. However, South Africa is in the beginning stages of teleteaching and needs to start by assessing the reaction and satisfaction of present management trainees to teleteaching and prepackaged video.

In the teleteaching classroom, a similar structure to teaching the case study was followed as that of the prepackaged video. Footage, such as 'lesson on the case study method', was cut in with live instructions from the instructor. The live broadcast allowed live interacting by using audio systems by trainees before and after the debate.

Time scale and research design

Time scale

The experimental research was conducted from May to August 1995. The television tuition was scheduled as part of the MBA/MBL training and was done in a period lasting approximately one and a half to two hours. Other commitments to MBA/MBL classes and availability of the teleteaching facilities made it impossible to schedule it any other way without a total disruption of other instructors' classes. The prepackaged video was done independently of the instructor while the teleteaching class involved a 'live' presentation by the instructor.

Research design and the null hypotheses

Hypothesis testing

Five hypotheses were looked at and they are outlined below:

Within group differences

- University of Pretoria full-time MBAs

Ho_1 There is no difference in the attitudes of University of Pretoria full-time MBA trainees to teleteaching and prepackaged video.

- University of Pretoria part-time MBAs

Ho_2 There is no difference in the attitudes of University of Pretoria part-time MBA trainees to teleteaching and prepackaged video.

- Unisa distance education MBL

Ho_3 There is no difference in the attitudes of Unisa distance education MBL trainees to teleteaching and prepackaged video.

Between group differences

- Teleteaching

Three groups will be taught by teleteaching. The reactions will be assessed.

Ho_4 There is no difference in the attitudes to teleteaching between University of Pretoria full-time, University of Pretoria part-time and Unisa distance education.

- Prepackaged video

Three groups will be taught by prepackaged video. The reactions will be assessed.

Ho_5 There is no difference in the attitudes to prepackage video between University of Pretoria full-time, University of Pretoria part-time and Unisa distance education.

Measuring instruments

The research was conducted by an opinion survey. A questionnaire is available on request.

Design of the questionnaire rationale

Television tutoring reaction questionnaire

The questionnaire was developed with consideration to the work of Smeltzer (1986: 42). It was designed with input from STATOMET, Bureau for Statistical and Survey Methodology of the University of Pretoria, who was also responsible for the statistical analysis.

The questionnaire was divided into seven sections including areas such as tuition, the instructor, and group dynamics. There were two surveys – one for prepackaged video and another for teleteaching – but the questions were the same apart from V20, 22, 34–37, which only applied to teleteaching as they were about the interactivity of the tuition technology.

Table 2. Research design

	F	P	U	
LITV	University of Pretoria	University of Pretoria	Unisa	
Teleteaching (TT)	Full-Time UP(FT)	Part-time UP(PT)	Distance education U(DE)	
1	Pretoria 2 (2)*	Pretoria 1 (3)	Unisa 4 (3)	
	Pretoria 7 (2)	Pretoria 3 (3)	Unisa 5 (5)	
	Pretoria 8 (6)	Pretoria 4 (5)	Unisa 6 (6)	
	Pretoria 9 (5)	Pretoria 7 (5)	Unisa 7 (7)	
	F1 n = 15	P1 n = 16	U1 n = 21	52
Date of tuition	18/8/95	5/6/95	17/5/95	
Prepackaged video (PP)	Pretoria 1 (6)	Pretoria 2 (5)	Unisa 1 (7)	
2	Pretoria 6 (3)	Pretoria 5 (3)	Unisa 2 (4)	
	Pretoria 5 (4)	Pretoria 6 (4)	Unisa 3 (8)	
	Pretoria 10 (3)			
	F2 n = 16	P2 n = 13	U2 n = 19	48
Date of tuition	18/8/95	5/6/95	15/5/95	
			Total	100

* Number of students in a group

The sections comprising the questionnaire were as follows:

A. General V1–V7: General background on television literacy

B. Method of tuition V8–V17: Probing the reaction to television and traditional tutoring

C. Your feelings about teleteaching/prepackaged video V18–V30: Encouraged or discouraged about the experience

D. Physical impression V31–V37: The location and its suitability

E. The instructor – dislike or liking of the instructor

F. Group dynamics – more cohesive, less cohesive

G. Quality of the video – technical quality: graphs, sound.

Although the major objective was to obtain the trainee's reaction to the method of tuition, the other areas could influence the outcome, e.g. technical problems could overshadow the experience. If the group was anticipating a pleasurable experience, this might have acted as a reinforcer but technical problems could change this attitude to a negative one.

The null hypothesis

The null hypothesis proposed that there were no differences within and between the groups' reaction to the methods of tuition. The Wilcoxon two sample test was used for within group differences and the Kruskal-Wallis Test (chi-square approximation) was used for between group differences. A comparison within groups is first discussed and then between groups. Table 3 indicates that there were two variables significant at the 1% level, nine at the 5% level and nine at the 10% level when within-groups were compared, whereas there was one variable significant at the 1% level, eight at the 5% level and three at the 10% level when teleteaching was compared between groups and four at the 1% level, three at the 5% level and one at the 10% level when prepackaged video tuition was compared between groups. This seems to indicate that there is a difference in the reaction to the methods of tuition as well as differences between groups. Table 3 is a summary of significant variables with those that seem the most significant indicated first. The descriptive statistics indicated that when the total sample group was presented with a choice of whether they held a positive or negative view towards the new technology, 98 per cent of the respondents described their view as positive, while two per cent held a negative view.

Comparison of methods of tuition within groups

University of Pretoria full-time MBAs

H_{01} There is no difference in the attitudes of University of Pretoria full-time MBA trainees to teleteaching and prepackaged video.

Table 3. Summary of significant variables

	UP (Full-time)	UP (Part-time)	Unisa (Distance education)	Teleteaching	Prepackaged video
V39		**	***	**	
V18	***	*			***
V41	**			**	**
V23	*	**		**	
V47	*		**	**	
V45		**	**	*	
V13			**		**
V32			*		***
V46		**		*	
V33		*		**	
V28	**	*			
V40				**	*
V15					***
V20				***	
V10					***
V42				**	
V43				**	
V24					**
V31	*				
V14		*			
V9			*		
V22				*	
	6	8	6	12	8

Note: *** Significant at 1% (0,01) level (therefore, also at 0,05; 0,10)

** Significant at 5% (0,05) level (therefore, also at 0,10)

* Significant at 10% (0,10) level

As indicated in Table 2, the full-time (F1) teleteaching group consisted of 15 participants while the full-time (F2) prepackaged video group consisted of 16 participants. Most variables were not significant. As indicated in Table 3, only six of the 34 variables compared are significant. One variable was significant at the 1% level, two at the 5% level and three at the 10% level. These variables are discussed below.

*V18*** I think we were confused about what we were supposed to do.*

In the prepackaged video group, 73% agreed strongly with this statement compared to 20% in the teleteaching group. Although there were two on-site personnel, better training could possibly improve the situation. One would have expected that the teleteaching would have been the more difficult to operate as many people are used to using video recorders.

*V28** The video helped me gain an understanding of difficult ideas.*

The teleteaching group agreed with this statement more than the prepackaged video group. In the teleteaching group, 53% agreed with this statement as opposed to 13% in the prepackaged video group. Teleteaching seems to create more satisfaction in the group.

*V41** Did the instructor stimulate a creative environment?*

The prepackaged video group did not believe that the environment was as stimulating as the teleteaching group did. In the teleteaching group, 13% replied 'yes' while there was 0% in the prepackaged video group; 67% to 50% responded with ìsometimesî while 20% of the teleteaching group replied ìnoî as compared to 50% in the prepackaged video group. This would seem to indicate a preference for teleteaching.

V23 We covered the material at a comfortable rate.*

There was a slight significance on this variable with the teleteaching group being more content with the rate than the prepackaged video group. Most of the teleteaching group, 87%, agreed or agreed strongly with the variable as compared to 53% of the prepackaged video group.

V31 The equipment was easy to operate.*

The teleteaching group agreed or agreed strongly at 87% compared to the 43% of the prepackaged video group. This seems strange as more people are familiar with video tapes than teleteaching. Either the on-site personnel at the teleteaching venue gave a better explanation than those at the prepackaged video venue or the instructor being able to interact with the group to check that they understood the instructions, could have made the difference.

V47 I found my group interaction (warm ... unhelpful).*

The teleteaching group had a much more positive feeling towards their group than the prepackaged video had. In the teleteaching group, 60% found the group warm, friendly and supportive compared to 25% of the prepackaged video group, while 40% of the teleteaching group found their group helpful and sufficient compared to 75% of the prepackaged video group. Neither of the groups found their groups unhelpful or full of mistrust.

Synthesis of University of Pretoria full-time MBAs

At the University of Pretoria, full-time MBAs favour the teleteaching method of tuition compared to the prepackaged video. Therefore, H_{01} is rejected. Three variables, V18, V28, and V23 were found to be significant in the University of Pretoria part-time group and V47 was also in common with the Unisa distance education group. However, these were at different levels of significance. In conclusion, better training for the support staff of the prepackaged video session could possibly remove any significance difference, but the interactive teleteaching method of tuition seems to be favoured by the group.

University of Pretoria part-time MBAs

H_{02} There is no difference in the attitudes of the University of Pretoria part-time MBA trainees to teleteaching and prepackaged video.

As indicated in Table 2, the part-time (P1) teleteaching group consisted of 16 participants and the part-time (P2) prepackaged video group consisted of 13 participants. Most variables were not significant with only eight out of 34 being significant. As indicated in Table 3, none were significant at the 1% level, four at the 5% level and four at the 10% level.

*V23** We covered the material at a comfortable rate.*

This was also significant at the 10% level in the University of Pretoria full-time group. However, it was in opposite directions. In the part-time group, the prepackaged video group agreed more strongly with this statement than the teleteaching group did – 84% compared to 40%. Both sessions were in a normal scheduled teaching period and with more exposure it is believed that a comfortable rate for the majority of participants could be reached.

*V39** Did you enjoy the style of the presentation?*

The teleteaching group tended to like the style of presentation more than the prepackaged video group did – 86% to 46%. However, 7% of the teleteaching group compared to 38% of the prepackaged video group replied 'sometimes' while only 7% of the teleteaching group compared to 15% of the prepackaged video group enjoyed the style of presentation. This variable was also significant in the Unisa distance education group at the 1% level and will be discussed in the next section.

*V45** The instructor creates a sense of togetherness with a teleteaching/prepackaged video broadcast.*

The teleteaching group was more in agreement with this statement than the prepackaged video group was – 43% to 16%. This variable was also significant in the Unisa distance education group at the 10% level. It would seem that the interactive broadcast creates a better sense of togetherness.

*V46** I find it useful to share ideas with others.*

The teleteaching group found this less important than the prepackaged video group did – 14% to 58%. This could possibly be due to the prepackaged video group having to depend upon each other to a greater extent than the teleteaching group who could also ask for advice from the instructor.

V14 Did the teleteaching/prepackaged video lesson affect your opportunity to ask questions?*

The prepackaged video group felt that their opportunity to ask questions was affected a great deal – 54% compared to 20%. In the teleteaching group, 60% felt that it had affected their opportunity to ask questions very little. This could be expected if one considers that in the prepackaged video group they could not ask any questions. The tuition method felt to be the same as face-to-face class by 20% of the teleteaching group and 15% of the prepackaged video group.

V18 I think we were confused about what we were supposed to be doing.*

The prepackaged video group agreed or agreed strongly with this statement – 61% compared to 33% of the teleteaching group. The interactivity of teleteaching probably allowed for more questions and clarification compared to prepackaged video. However, better trained on-site personnel and more time to assimilate instructions could possibly improve the prepackaged video group's response and help them to feel more comfortable.

V28 The video helped me gain an understanding of difficult ideas.*

The prepackaged video group was more in agreement with this statement than the teleteaching group – 38% to 20%. The group dynamics possibly played a part. This area needs to be probed in further research. This variable was also significant in the University of Pretoria full-time groups at the 5% level but, in that case, the teleteaching group was more in agreement with the statement than the prepackaged video group – 53% to 13%. Possibly the teleteaching group, who had access to the instructor, felt that the instructor and not just the video had helped in the understanding of difficult ideas. This difference between groups also needs to be probed.

V33 The broadcast was boring.*

The teleteaching group disagreed more with this statement than the prepackaged video group did – 85% to 65%. The interactivity seems to have created a more stimulating environment.

Synthesis of University of Pretoria part-time MBAs

The teleteaching style of presentation seems to be favoured more than the prepackaged video style. The teleteaching group also believed that the presentation was less boring. The instructor was seen to create more of a sense of togetherness in the teleteaching group than the prepackaged video group. In the part-time MBA group of the University of Pretoria, teleteaching seems to be the preferred media. Therefore, the H_{02} is rejected.

Unisa distance education MBLs

H_{03} There is no difference in the attitudes of the Unisa distance education MBL trainees to teleteaching and prepackaged video.

As indicated in Table 2, the Unisa distance education (U1) group who received teleteaching consisted of 21 participants while the Unisa distance education (U2) group who received prepackaged video consisted of 19 participants. As indicated in Table 3, most variables were not significant with only six out of 34 being significant, but they mainly related to method of tuition. One variable was significant at the 1% level, three at the 5% level and two at the 10% level.

*V39*** Did you enjoy the style of presentation?*

The prepackaged video group did not enjoy the presentation to the same extent as the teleteaching group – 61% to 95%. In the teleteaching group, nobody disliked the presentation while in the prepackaged video group, 17% disliked it. This variable showed the same tendency in the University of Pretoria part-time group. This would seem to indicate that the more interactive method is preferred.

*V13** Did you feel more or less stimulated in class as a result of teleteaching/prepackaged video?*

The teleteaching group felt more stimulated than the prepackaged video group – 43% to 21%. The prepackaged video group felt less stimulated than the teleteaching group – 36% to 10%, while 48% of the teleteaching group and 42% of the prepackaged video group felt as stimulated as in a face-to-face class.

*V47** I found my group interaction was warm, sufficient ... unhelpful.*

The prepackaged video group found that their group was more helpful than did the teleteaching group – 56% to 19%. The teleteaching group felt their group was sufficient – 76% compared to 45% of the prepackaged video group. This variable was also significant in the University of Pretoria full-time group. There the teleteaching group was seen as more helpful than the prepackaged video group – 60% to 25%. Group dynamics could account for group differences and should be further probed.

V9 How was the class time spent in the teleteaching/prepackaged video sessions compared to the traditional sessions?*

The prepackaged video group seems to have felt that it was more effective than the face-to-face session when compared to the teleteaching group – 58% to 21%. However, 68% of the teleteaching group felt it was the same as a traditional class compared to 26% of the prepackaged video group.

V32 The venue was well suited for the course.*

The prepackaged video group was more in agreement with this statement than the teleteaching group – 83% to 71%. However, 14% of the teleteaching group did not know if it was suitable. This variable is not significant in the other groups.

V45 The instructor creates a sense of togetherness with the teleteaching/prepackaged video broadcast.*

The teleteaching group was more in agreement with this statement than the prepackaged video group – 57% to 28%. The same trend was prevalent in the University of Pretoria part-time group.

Synthesis of Unisa distance education MBLs

Variable 39 was significant at the 1% level for Unisa distance education and at the 5% level for the University of Pretoria part-time group. This would seem to indicate that there is a preference for the more interactive method of television tuition. Therefore, H_{03} is rejected. The instructor also seems to create more togetherness with this method. There are no variables that are common to all three groups, but in the next section we will be discussing the between group variations.

Comparison of methods of tuition between groups

The methods of teleteaching and prepackaged video tuition were compared between the University of Pretoria full-time groups, the University of Pretoria part-time groups, and the Unisa distance education groups. The null hypotheses were that there would be no differences between the three groups. The Kruskal-Wallis test indicated that 20 variables were significant: twelve in the teleteaching group and eight when respondents received a prepackaged video session. There were five variables significant at the 1% level, eleven at the 5% level and four at the 10% level. This amount of variation would seem to

indicate that there were differences in the groups' response to the tuition methods.

Teleteaching tuition between groups

H_{04} There is no difference in the attitude to teleteaching between University of Pretoria full-time, University of Pretoria part-time and Unisa distance education.

Twelve variables were found to be significant: one at the 1% level, eight at the 5% level and three at the 10% level. There were about 50 participants in the teleteaching session. The significant variables are summarised in Table 4.

Interpretation of results

*V20*** It was hard to know who was talking at the other venues.*

In the Unisa distance education group this variable is significant at the 1% level group. Most of the participants, 57%, agreed or agreed strongly with this statement compared to 20% of the University of Pretoria full-time group and 13% of the University of Pretoria part-time group. The Unisa distance education broadcast was the first broadcast and the technical side possibly did not run as smoothly as the two other broadcasts. With further training this should be sorted out.

*V23** We covered the material at a comfortable rate.*

The University of Pretoria full-time group seemed more in agreement with this statement than the Unisa distance education and University of Pretoria part-time group – 87%; 57%; 40%. This is at the 5% level of significance. Possibly, with more familiarity with the technology, a comfortable rate can be reached for all groups. However, it might be an indication of other factors operating. This variable is also significant when one compares methods of tuition teleteaching/prepackaged video in the University of Pretoria full-time and part-time groups at the 10% and 5% level respectively. The University of Pretoria part-time group seemed more comfortable with the prepackaged video material.

*V33** The broadcast was boring.*

The University of Pretoria full-time group was more in agreement with this statement than the other two groups – 33% compared to 7% of the University of Pretoria part-time group and 10% of the Unisa distance education group. This would seem to indicate that there could be a difference between full-time, part-time and distance education trainees. The University of Pretoria part-time group was also significant on this variable at the 10% level when methods of tuition are compared. The University of Pretoria part-time group found the teleteaching less boring than the prepackaged video.

*V39** Did you enjoy the style of presentation?*

The University of Pretoria part-time group and the Unisa distance education group enjoyed the style of presentation more than the University of Pretoria full-time group – 86%; 95%; 60% – while 13% of the University of Pretoria full-time group did not like the presentation. Only 7% of the part-time group and none of the Unisa distance education group disliked the presentation. It would seem that the University of Pretoria full-time group do not have much preference for teleteaching or prepackaged video as this is not significant, but in the University of Pretoria part-time and Unisa distance education groups teleteaching is preferred to prepackaged video and is significant at the 5% and 1% level.

*V40** How was the explanation of the case study from the instructor?*

Only 27% of the University of Pretoria full-time group found it acceptable, while 67% felt that it could be improved. In the University of Pretoria part-time group, this ratio was 43% to 43%, while in the Unisa distance education group it was 71% to 50%. In the University of Pretoria full-time group, only one person felt that it was poor, while there were none in the Unisa distance education group. This will need to be probed in further research.

Table 4. Teleteaching – between groups

Variable	Mean score			Level of significance
	UP(FT)	UP(PT)	U(DE)	
V20	19	19	34	***
V23	33	19	25	**
V33	33	21	22	**
V39	31	24	22	**
V40	31	28	19	**
V41	32	26	19	**
V42	31	27	19	**
V43	30	23	23	**
V47	20	21	31	**
V22	21	21	36	*
V45	19	25	29	*
V46	31	19	25	*

*** Significant at 1% (0,01) level (therefore, also at 0,05; 0,10)

** Significant at 5% (0,05) level (therefore, also at 0,10)

* Significant at 10% (0,10) level

*V41** Did the instructor stimulate a creative environment?*

In the University of Pretoria full-time group, 13% agreed with the statement compared to 36% of the University of Pretoria part-time group and 62% of the Unisa distance education group. However, 67% of the University of Pretoria full-time group felt that to some extent this was the case, while 50% of the University of Pretoria part-time group and 33% of the Unisa distance education group indicated the same. The most satisfied group is the Unisa distance education group. Variable V41 is also significant at the 5% level in the University of Pretoria full-time group in that the instructor is seen as stimulating a more creative environment in the teleteaching session as compared to the prepackaged video session. The reasons why the University of Pretoria full-time group is different will have to be investigated further.

*V42** Did the instructor confidently project the subject content?*

With regard to V42, 47% of the University of Pretoria full-time group agreed to 57% of the University of Pretoria part-time group and 90% of the Unisa distance education group. None of the participants of the Unisa distance education group disagreed with the statement whereas 3 (20%) of the University of Pretoria full-time participants did.

*V43** Was the length of the video presentation timed comfortably?*

Of the University of Pretoria full-time group, 67% agreed with the statement compared to 93% of the part-time group and 95% of the Unisa distance education group. The University of Pretoria full-time group was the least comfortable with the length of the presentation. The University of Pretoria full-time participants seem to be differing in some way from the University of Pretoria part-time and Unisa distance education participants.

*V47** I found my group interaction warm, sufficient ... unhelpful.*

The University of Pretoria full-time group felt their group interaction was 60% warm and 40% sufficient, the University of Pretoria part-time group experienced their group interaction as 57% warm and 43% sufficient, while the Unisa distance education group was 19% warm and 76% helpful and sufficient. Generally, all the groups seemed satisfied with their group interaction.

V22 I was sometimes uncertain whether the people at the other venues were listening.*

With regard to V22, 63% of the Unisa distance education group were in agreement compared to 20% of the University of Pretoria full-time group and 36% of the University of Pretoria part-time group. The Unisa distance education broadcast was the first one and it was possibly not as smooth as the other two. However, with further exposure any technical area could be improved upon.

V45 The instructor creates a sense of togetherness with a teleteaching broadcast.*

V46 I find it useful to share ideas with others.*

Only 20% of the University of Pretoria full-time group agreed with variable 45 compared to 43% of the University of Pretoria part-time group and 57% of the University of Pretoria full-time group. Of the University of Pretoria full-time group, 33% disagreed or strongly disagreed with this statement as compared to 29% of the University of Pretoria part-time group and 14% of the Unisa distance education group. This variable was also significant in the University of Pretoria part-time group and the Unisa distance education group at the 5% level. The teleteaching group is more in agreement that the instructor creates a sense of togetherness than the prepackaged video group.

With regard to V46, 87% of the University of Pretoria full-time group, 49% of the University of Pretoria part-time group and 81% of the Unisa distance education group were in agreement with this statement. In the University of Pretoria part-time group, the largest percentage, at 43%, did not know. The University of Pretoria full-time group seems a bit different from the other groups.

Synthesis of between group teleteaching

The University of Pretoria full-time group seems to be less favourably disposed to the teleteaching medium than the University of Pretoria part-time group and the Unisa distance education group. Therefore, H_{04} is rejected. However, although the University of Pretoria full-time group is less favourably disposed, generally, there is still a high apprecia-

Table 5. Prepackaged video – between groups

Variable	Mean score			Level of significance
	F	P	U	
V10	26	31	18	***
V15	30	26	18	***
V18	33	24	16	***
V32	15	25	30	***
V13	27	20	18	**
V24	30	17	23	**
V41	29	26	17	**
	29	24	19	*

*** Significant at 1% (0,01) level (therefore, also at 0,05; 0,10)

** Significant at 5% (0,05) level (therefore, also at 0,10)

* Significant at 10% (0,10) level

tion of teleteaching as a method of tuition. Individual characteristics of the University of Pretoria full-time group will need to be explored.

Prepackaged video tuition between groups

H_{05} There is no difference in the attitude to prepackaged video between the University of Pretoria full-time group, the University of Pretoria part-time group and the Unisa distance education group.

Eight variables were found to be significant: four at the 1% level, three at the 5% level and one at the 10% level. There were about 50 participants in the prepackaged video session. According to the previous results, the groups seemed to not respond as favourably to prepackaged video as teleteaching. The significant variables are summarised in Table 5.

Interpretation of results

*V10*** How dominant was the instructor in the prepackaged video session compared to the instructor in a traditional session?*

In the University of Pretoria full-time group, 19% felt the prepackaged video session was about the same as a face-to-face class, 15% of the University of Pretoria part-time group felt the same way along with 53% of the Unisa distance education group. In the University of Pretoria full-time group, 69% felt that the instructor was much less dominant with 85% of the University of Pretoria part-time group and 32% of the Unisa distance education group responding in the same way. The University of Pretoria full-time and part-time groups seem more similar than the Unisa distance education group. However, many viewed the prepackaged video session as very similar to a class situation.

*V15*** Did the prepackaged video session affect your willingness to ask questions?*

Sixty-nine per cent of the University of Pretoria full-time group felt it was the same as a face-to-face class, while 62% of the University of Pretoria part-time group and 59% of the Unisa distance education group thought so too. None of the University of Pretoria full-time group seemed very discouraged. Nevertheless, 15% of the University of Pretoria part-time group and 36% of the Unisa distance education group were discouraged, i.e. 7 participants out of 19. This needs to be probed in further research.

*V18*** I think we were confused about what we were supposed to be doing.*

The University of Pretoria full-time group agreed or agreed strongly with this statement – 80% compared to the University of Pretoria part-time group's 61%, and 37% of the Unisa distance education group. Possibly, with more exposure this variation could be reduced. In a prepackaged video lesson one has the opportunity to replay the instructions if one is not clear about the situation. This did not take place in this experimental situation and could possibly be improved upon in a more normal environment. This variable is also significant in the University of Pretoria full-time and University of Pretoria part-time groups when methods of tuition are compared. The teleteaching method is preferred to the prepackaged video at the 1% and 10% level respectively.

*V32*** The venue was well suited for the course.*

The University of Pretoria full-time disagreed or disagreed strongly with this statement: 50% compared to 30% of the University of Pretoria part-time group and 11% of the Unisa distance education group. A large lecture hall with bench row seating (+ 200) was used. Certainly, a better designed venue with a more comfortable seating arrangement, allowing round-table discussion, could alleviate the situation. This variable was also significant in the Unisa distance education group when methods of tuition were compared. The prepackaged video group agreed slightly more with the above statement than the teleteaching group.

*V13** Did you feel more or less stimulated in a class as a result of prepackaged video teaching?*

In the University of Pretoria full-time group, 50% felt more stimulated along with 61% of the University of Pretoria part-time group, while in the Unisa distance education group, only 21% were more stimulated. The University of Pretoria full-time group had 37% less stimulated participants compared to 15% and 13% of the University of Pretoria (full-time and part-time groups) respectively. However, the majority of the Unisa distance education group, 42%, felt it was the same as a face-to-face class.

*V24** I think that if the whole group had been together in the same room, the discussion would have been more productive.*

The University of Pretoria full-time group agreed or agreed strongly with this statement: 40% compared to 7% of the University of Pretoria part-time group and 21% of the Unisa distance education group. The University of Pretoria full-time group seemed somewhat different from the rest.

*V41** Did the instructor stimulate a creative environment?*

In the University of Pretoria full-time group, 50% felt this was not the case, while 46% and 17% of the University of Pretoria part-time and Unisa distance education groups respectively agreed with the statement. However, 50% of the University of Pretoria full-time group felt it was the same as a face-to-face class with 38% of the University of Pretoria part-time group and 44% of the Unisa distance education group feeling the same way. None of the University of Pretoria full-time group agreed with the statement, but 15% and 39% of the University of Pretoria part-time and Unisa distance education groups respectively did. It seems that there is a difference between the full-time group and the others and this needs to be researched. This variable was also significant when methods of tuition were compared in the University of Pretoria full-time group and teleteaching came out stronger. This variable was also significant when teleteaching was compared among groups. teleteaching was less favoured in the University of Pretoria full-time group than the other two groups.

V40 How was the explanation of the case study from the instructor?*

The explanation was felt to be poor by 31% of the University of Pretoria full-time group compared to 15% of the University of Pretoria part-time group and none of the Unisa distance education group. This variable is also significant in the teleteaching group comparison at the 5% level. The University of Pretoria full- time group found it less acceptable than the University of Pretoria part-time and Unisa distance education groups.

Synthesis of between group prepackaged video

The University of Pretoria full-time group seems to hold a more negative view to a prepackaged video session than the University of Pretoria part-time and Unisa distance education

groups. The confusion about what was expected of them could probably be relieved by further exposure to this kind of tuition or by allowing the trainees to do the lesson in their own time. There they will have an opportunity to replay the video for any instructions that might have been missed. The Unisa distance education group also seems to not be as stimulated by the prepackaged video session as the other groups but the majority saw it as good as a face-to-face class. Therefore, H_{05} is rejected.

There were differences between groups in both the teleteaching and prepackaged video method of tuition. Therefore, H_{04} and H_{05} are rejected. The University of Pretoria full-time group seemed the most negative towards teaching technology. This could indicate a difference between the more residential trainee compared to the more part-time or distant education trainee or could be attributable to individual differences.

Limitations of the study

This research was only conducted in the Pretoria area and thus could be a distorted sample which might not reflect national feelings. Limitations of field research, opinion survey and the small sample size (n = 100) could also have had an effect on the results. An attitude survey is a point-in-time opinion and would have to be monitored over the long term to ascertain whether the same trends exist over time. This study also only looked at the teaching of a case study in a management course, but the results could be different if other courses such as, for example, accountancy, are taken. Another limitation of the study is that the vast majority of participants (92%) had never taken part in a live interactive session and thus the Hawthorne or novelty effect could have been present. This could have contributed to the overall positive reaction to the new technology. Nevertheless, the study does seem to indicate some trends that could be followed up in further research.

Summarised hypothesis findings

There were differences within and between groups in their attitudes towards teleteaching and prepackaged video. All the groups seemed to have a preference for the interactive teleteaching method of tuition. Therefore, H_{01}, H_{02}, H_{03} are rejected.

There were differences between groups in both teleteaching and prepackaged video. Therefore, H_{04} and H_{05} are rejected. The University of Pretoria full-time group seemed the most negative towards teaching technology. This could indicate a difference between the more residential trainee compared to the more part-time or distant education trainee or could be attributable to individual differences.

Further research

The differences within and between groups seem attributable to the methods of tuition and the respondents' feelings about the methods of tuition. Most trainees seemed satisfied with the physical surroundings, the instructor, and the quality of the video. Although some comments were made about the sound and how some video shots could be improved, it would seem that these aspects did not affect the respondents in a negative fashion. Some further research areas are outlined below.

Television tuition

- Further research is needed on a national sample of MBA/MBL trainees to ascertain their reaction to video technology and whether it is similar to the Pretoria sample. Focus groups could be explored nationwide.
- Advanced television technology such as UITV could also be surveyed as could satellite broadcasts that could bring international business training to the homes of the MBL/MBA trainees.
- The whole area identified by Lemak & Miskin (1992: 17) needs to be explored thoroughly – what role can correspondence courses and noninteractive video tapes play in the future of distance education?
- This is a one-off study and in order to really benefit from the research, a longitudinal study over a number of years needs to be planned. This could entail comparisons of media, instructors, subjects and whether attitudes change over a time period. Given the advances in multimedia, it would be advantageous to include these developments in the study. This information would form a valuable input into a management information system.

Conclusion

This research indicates the unquestionable positive value of using teaching technology in South African management training. The management trainees hold positive views on the new teaching technologies and, therefore, management training in South Africa needs a new direction. The study indicates some trends for further research particularly in the area of individual differences. Distance education and teaching technology could be used to train new and better management skills in the context of overall economic reconstruction and development.

References

Breakwater Monitor. 1996. March.

Hofmeyr, K. B., Rall, P. J. & Templer, A. J. 1995.'The future challenges facing South African human resource management', South African Journal of Business Management, 26(13):108–113.

Lemak, D. J. & Miskin, V. D. 1992. 'Diversity, distance, learning and management education: A glimpse into the future', *Report submitted to the Academy of Management Annual Meeting, Management Education/Development Division,* Seattle: Washington State University: 1–17.

Perry & Associates. 1996. *Strategic business issues.* Johannesburg.

Smeltzer, L. R. 1986. 'An analysis of receivers' reaction to electronically mediated communication', *Journal of Business Communication,* 23(4): 37–54.

White paper on reconstruction and development: Government's strategy for fundamental transformation, 1994. September.

Towards a framework of firm responsiveness to turbulent environments

Mzamo P Mangaliso
Isenberg School of Management
University of Massachusetts

Raza A Mir
School of Business Administration
Monmouth University

Bradford J Knipes
Department of Economics and Business
Westfield State College

While posing a real problem for management practitioners, environmental turbulence finds only fleeting and fragmented mention in the mainstream academic literature of management. This article discusses turbulence, traces its origins, and subjects it to critique. Based on observations of responses to it, and several theoretical streams, including new theories such as hypercompetition and chaos theory, a framework of responsiveness to turbulence is presented and discussed, and a set of propositions presented. The rationale behind the framework is that successful responses to turbulence are contingent on sociocultural as well as technical-economic conditions in the environment.

Introduction

The relationship between planning and firm performance has been the subject of intense debate for many years (Andrews 1971; Miller & Cardinal 1994; Mintzberg 1994a). The debate reflects one of the major concerns of strategic management: aligning the skills and resources possessed by the organisation with opportunities and constraints in the environment (Jemison 1981; Miles & Snow 1994; Schendel & Hofer 1979). This alignment with an uncertain and fast-changing environment engages the organisation's core competencies in the pursuit of corporate goals (Andrews 1971; Mahoney & Pandian 1992; Prahalad & Hamel 1990; Thompson 1967). A variety of contingencies determine the fit sought by organisations, and one of these contingencies is environmental turbulence.

Environmental turbulence is a phenomenon that is increasingly being recognised as having an important moderating influence in the relationship between planning and performance (Ansoff 1991; Goodstein & Boeker 1991; Miller & Friesen 1983; Mintzberg 1994b; Theobald 1994). It has been mentioned with increasing frequency as researchers have grappled with how organisations should cope with rapid changes in the environment. For instance, according to some observations, environmental turbulence was involved in the changing managerial roles within Japanese corporations (Johnston 1995), the reconstruction of management in post-communist Eastern Europe (Johnson & Loveman 1995), the 'greening' of organisational studies (Shrivastava 1994), the systemic organisational upheavals in the United States and their effect on big corporations (Quinn & Mason 1994), organisational reengineering (Hammer & Champy 1993), and the TQM movement (Klimoski 1994). This may be the reason why some of the most prominent scholars in the field have warned that environmental turbulence will continue to challenge organisations (Ansoff 1984, 1991; Drucker 1980; Huber 1984).

Despite these admonitions, there appears to be little research on managing in turbulent environments. The academic literature aimed at helping managers learn to cope in turbulent environments is sparse (Camillus & Datta 1991; Ellis 1982; Theobald 1994). Consequently, most managers either suffer from the effects of paralysis by analysis, or simply resort to tried and tested approaches to strategy such as centralisation, standardisation, and efficiency maximisation, which are largely inappropriate in turbulent conditions (Bourgeois, McAllister & Mitchell 1978; Cameron, Kim & Whetten 1987; D'Aveni 1989; Khandwalla 1978). Largely following the threat–rigidity response pattern, these strategies often act as blinkers which provide focus on direction, but block out the peripheral vision so crucial in turbulent environments

(D'Aveni 1989; Mintzberg 1990; Staw, Sandelas & Dutton 1981).

Turbulence thus poses a fundamental challenge to strategic management by mediating the planning–performance relationship (Miller & Cardinal 1994). If turbulence can be accepted as real, the entire premise of strategic planning may be challenged as merely an exercise in organisational self-delusion (Mintzberg 1991). This article explores the theoretical streams of research that have informed the debate, and offers a map for the researcher studying the impact of strategic planning on firm performance. The intention is to engage the academy in discourse concerning the origins of environmental turbulence, its manifestations, and implications for organisations.

Towards understanding environmental turbulence

Characterisations of environmental turbulence are as numerous as they are disparate. Some sources define turbulence as a wild, unruly, and disorderly commotion, a great perturbation (Webster's English Dictionary), others as changeability, speed, and novelty in the environment (Ansoff 1984). Still others define it as environments in which the speed of change and interconnectedness of elements are high (Emery & Trist 1965). In empirical studies, some researchers have operationalised turbulence through the perceived relationship between centralisation and effectiveness (Huber, Miller & Glick 1990), while others see it as a key moderator in the planning–performance relationship (Miller & Cardinal 1994; Pearce, Robbins & Robinson 1987). According to Pfeffer & Salancick (1978), turbulence is change coming from anywhere, without notice, producing unanticipated consequences to both those initiating it and those experiencing it. Some characterise turbulence as difficult-to-predict discontinuities in the environment (Aldrich 1979; Keats & Hitt 1988), especially in situations where there is room for strategic choice within a highly deterministic environment (Hrebiniak & Joyce 1985). Environmental turbulence causes factors in the environment to become more uncertain and the value of important variables to become more erratic (Smart & Vertinsky 1984). As Metcalfe (1974) noted, the sequential escalation of the interrelatedness of the complexity and dynamism in the causal texture of the environment results in greater uncertainty which, in turn, creates problems for complex organisations. Bourgeois & Eisenhardt (1988) refer to the analogous high-velocity environment characterised by rapid discontinuous change in demand, competitors, technology, or regulation, so that information is often inaccurate, unavailable, or obsolete. In this article, environmental turbulence is defined as *the complex interconnectedness of environmental elements that exhibit rapid, unpredictable, and discontinuous change that makes the future hard to predict.*

Related literatures

As figure 1 demonstrates, several streams of literature explain or delineate environmental turbulence. These include resource dependency (Aldrich 1979; Aldrich & Pfeffer 1976; Lawrence & Lorsch 1967; Pfeffer & Salancick 1978), decline (Cameron, Kim & Whetten 1987; D'Aveni 1989; Harrigan 1985; Starbuck & Hedberg 1977), crisis management (Smart & Vertinsky 1977; Starbuck, Greve & Hedberg 1978; Zammuto & Cameron 1985), volatility (Bourgeois 1985; Snyder & Glueck 1982; Tosi, Aldag & Storey 1973), uncertainty (Duncan 1972; Thompson 1967), chaos theory (Baumol & Benhabib 1989; Levy 1994; Stacey 1991), and postmodernism (Harland 1987;Rosenau 1992).

The underlying premise of the resource dependency literature is that decisions are made within organisations under the assumption that the environment is a systemlike context of resources, social structures, and the natural environment (Aldrich & Pfeffer 1976; Pfeffer & Salancick 1978). In order to survive and prosper, the organisation must acquire the resources that are critical to its functioning from an environment that is uncertain. This makes the availability of resources uncertain, which, in turn, calls upon the organisation to find a way of reducing this uncertainty. But there are also other

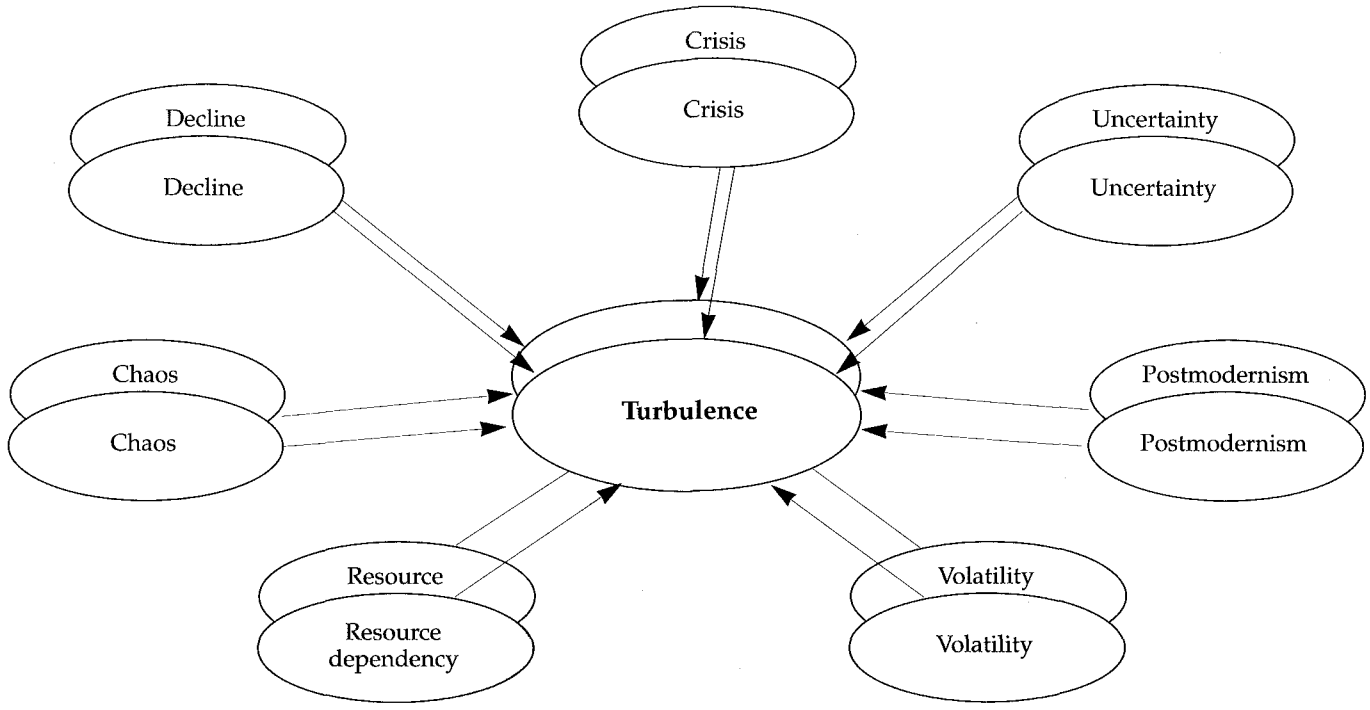

Figure 1. Theoretical influences of turbulence

organisations that are competing for the same resources. The organisation must reduce its dependence on these organisations and maximise dependence on itself (Pfeffer & Salancick 1978). Change is assumed to occur via autonomous variations in any element of the environment and is continuous and potentially predictable (Lenz & Engledow 1986). But when environments become turbulent in the manner defined earlier, the assumption of continuity and predictability of change breaks down. Resource dependency does, however, encourage organisational learning through opportunistic surveillance, and protection of the technical core of the organisation through buffering (Pfeffer & Salancick 1978; Thompson 1967).

The decline literature is one that many authors associate with diminishing resources in the organisation's task environment which tends to result in decreasing internal resource munificence over time (Cameron, Sutton & Whetten 1988; D'Aveni 1989). The typical response to declining resources is workforce reduction, which invariably leads to a variety of adverse effects (Greenhalgh, Lawrence & Sutton 1988). If the environment continues to be turbulent, downsizing and efficiency maximisation strategies are only 'fuel for temporarily perpetuating the lingering paralysis' (D'Aveni 1989: 599). While Harrigan's (1985) work recommends timely use of exit strategies as appropriate in declining industries, Zammuto & Cameron (1985) recommend domain consolidation and reduction of scale of operation in order to protect the firm's core product-market domain. In general, it can be said that both decline and turbulence have negative effects on the organisation. The difference is that decline tends to be felt by employees who are mostly at lower levels of the organisation, while turbulence is felt mostly by the top management team (Cameron, Kim & Whetten 1987).

The crisis management literature is grounded on the understanding that crises are events inside and outside the organisation that represent threats to organisational survival (Shrivastava & Mitroff 1987). They occur as a result of simultaneous and interacting failures in internal and external environmental elements (Perrow 1984). Turbulence is described more as a consequence of the crisis than its causal antecedent. The challenge for management is to devise ways of anticipating, preventing, and managing these corporate crises. Several models for doing this are suggested in the literature (Mitroff & Pauchant 1990; Meyers & Holusha 1986; Shrivastava 1987), most of which deal with the need for organisations to undergo a 'paradigm shift', or a fundamental rethinking of organisational goals and practices.

Volatility is typically defined as the change in the rate of change: as long as the rate of change is constant, it is predictable and therefore not problematic (Bourgeois 1978). Although research on environmental volatility can be traced back to the contingency theorists of the 1960s, the state of the art is still unclear (Snyder 1987). Debate seems to centre on the question of whether environmental volatility can be measured objectively. But, as with the crisis literature, the concept of volatility seems to involve change that is unpredictable and comes when it is least expected, with varying frequency and amplitude (Bourgeois 1985; Snyder & Glueck 1982; Tosi, Aldag & Storey 1973). The distinction between turbulence and volatility or crisis is in terms of duration. A crisis or a volatile situation is usually of shorter duration; with turbulence, no end is in sight (Smart & Vertinsky 1977).

The next important consideration of environmental turbulence comes from the uncertainty literature. Uncertainty has been defined as the absence of information; the difference between the amount of information required and the amount of information available; or a state that exists when an activity is embarked upon with less than complete knowledge about the probability of its outcome (Daft 1986; Downey & Slocum 1975; Galbraith 1977). Further reading of the literature on uncertainty (Bourgèois & Eisenhardt 1988; Downey & Slocum 1975; Duncan 1972; Thompson 1967) leads one to conclude that an implicit causal relationship exists between turbulence and uncertainty: environmental turbulence is regarded as a precursor to environmental uncertainty, varying in direct proportion to it. Yet, a distinction is made that turbulence is the state of the world whereas uncertainty is a state of mind. As Downey & Slocum (1975) note, uncertainty is a set of stimuli that lack meaning for informational value until perceived by an individual. A turbulent environment requires a high degree of abstraction by decision makers in order to produce manageable strategic maps. As the level of uncertainty increases, so does the individual's lack of knowledge and thus the need for higher quality strategic maps in line with the law of requisite variety (Ashby 1956). Responses to uncertainty may vary, from defensively buffering the organisational core (Thompson 1987), to proactively seeking potential synergistic benefits from the source of the turbulence itself, or capitalising on competitors' weaknesses (D'Aveni 1994).

Another stream of research is based on chaos theory (Baumol & Benhabib 1989; Levy 1994; Stacey 1991; Thietart & Forgues 1995). While not a new theory itself, the application of chaos theory to management research is a recent phenomenon that illuminates the concept of environmental turbulence (Levy 1994). Conceptualised mainly as a study of complex, nonlinear dynamic systems, chaos theory may be seen as an opportunity for extension of systems theory into the nonlinear realm of organisations (Katz & Kahn 1966). Organisations are enmeshed in chaos with tension between forces that push the organisation toward order and stability, e.g. planning, structure, coordination, and control; and forces that push the organisation toward instability and disorder, e.g. innovation, entrepreneurship, and experimentation. A balance of these forces is necessary. Organisations need a sufficient level of order and structure, but too much would destroy the organisation's creativity. Organisations also need a certain amount of chaos since it is the only way to find new forms, but too much would undermine their very existences (Thietart et al. 1995). Chaotic systems exhibit a sensitive dependence on initial conditions: small errors in the estimation of the present state result in exponentially large errors in the estimation of future states. As a result, long-term forecasting becomes impossible under chaos. Turbulence tends to drive the system towards chaos. In such situations, successful firms are those that adapt and react faster than the competition, and the best way of coping is through learning. But, according to some observers, systems do not remain chaotic for long (Phelan 1995). Business environments appear to go through long periods of incremental change punctuated by the occasional discontinuous disruption, the essence of the 'punctuated equilibrium' paradigm (Gersick 1991; Tushman & Anderson 1986). Under normal, stable conditions, however, it is recommended that firms shift the allocation of resources more towards strategic planning.

The last stream of research, postmodern theory, drawing from Baudrillard's assertion that '[t]he secret of theory is, indeed, that truth does not exist' (1986: 141), suggests that all attempts by managers to 'know' are rendered futile by the

existence of an ultimate unknowability (Harland 1987). In effect, postmodernism equates strategic planning with a 'truth claim' made by the organisation, and suggests that in an enacted environment, such claims are at best a product of organisational navel gazing, and at worst, an enforcement of discipline and order in the organisation, that ends up stifling all organisational creativity (Rosenau 1992).

The above discussion was an attempt to demonstrate that the discourse on turbulence is informed by numerous streams of research. Table 1 is a summary of the discussion. While this list of research streams is not complete, it represents the diversity of sources. Despite arguments to the contrary, it is our belief that environmental turbulence is on the increase (Drucker 1980; Mangaliso & Lane 1992; Miles & Snow 1986). Since its impact on contemporary organisations will be telling, we believe it is imperative that management researchers become more deliberate in studying the concept. The radical sociopolitical changes sweeping the globe, rapid technological breakthroughs, and wild economic up- and downswings will continue to put immense strain on organisations' capacity to cope.

Evidence of turbulence

From the foregoing discussion it can be understood why an analysis of the firm's environment should be a *sine qua non* in the overall schema of strategy making. The history of organisations in the US and elsewhere bears testimony to the fact that ignoring this can be fatal to the firm. Many organisations have been driven out of business because of ignorance of environmental developments and inability to adapt to them. Large firms such as Paramount Studios, Warner Bros., and MGM in the movie industry were shaken when the television networks CBS, NBC, and ABC entered the same scene. Facit was unprepared for the onslaught of the word processors. GTC failed to capitalise on the miniaturisation revolution of the 1980s, losing its leadership role in the microprocessor market (Prahalad & Hamel 1990). Some firms have been able to harness turbulence to their advantage by strategic product entries, divestments, mergers and acquisitions. For instance, the recent takeover of Lotus by IBM, or AT&T's use of value-based management systems to turn the divestiture of the 'Baby Bells' to advantage may be seen as *turbulence-mediated turnarounds* rather than mere reactions to turbulence.

Turbulence has not been limited to the economic, political and technological dimensions. In the legal or regulatory sphere, many airlines did not quite fathom the impact the 1978 Airline Deregulation Act. The opening up of Eastern European markets may be seen as turbulence-mediated opportunities in the international arena. These opportunities have led multinational corporations to redeploy their resources swiftly and strategically in these markets (Johnson & Loveman 1995).

Table 1. Related literatures

Stream	Authors	Premises
Resource dependency	Aldrich & Pfeffer (1979) Pfeffer & Salancick (1978) Thompson (1967)	Environment is a system-like context of resources, social structures, and a natural environment. Organisations use their resources to actively shape their environments or protect their core by buffering.
Decline	Cameron et al. (1987, 1988); D'Aveni (1989); Greenhalgh et al. (1988); Harrigan (1985); Zammuto & Cameron (1985)	Diminishing resources in organisations' task environments lead to adverse effects such as layoffs and mutual blame. Typically felt at lower levels of the organisation. Turbulence is felt at higher levels.
Crisis management	Meyer & Mitroff (1987) Mitroff & Pauchant (1990) Perrow (1984); Shrivasta & Mitroff (1987)	Events inside and outside organisations represent threats to organisations' survival. Organisations must anticipate, manage, and prevent corporate crises.
Volatility	Bourgeois (1978, 1985) Smart & Vertinsky (1977) Snyder (1987); Tosi et al. (1973)	Problem with volatility is that change in the rate of change in the environment leads to unpredictability. Unlike turbulence, volatility is of finite duration.
Uncertainty	Bourgeois & Eisenhardt (1988) Daft (1986) Downey & Slocum (1975) Duncan (1972) Galbraith (1977); Thompson (1967)	Exists when knowledge about activity is incomplete, probability of outcome not known, a set of stimuli that lack meaning until perceived by individuals. Uncertainty is a state of mind, turbulence is the state of the world.
Chaos theory	Baumol & Benhabib (1984) Levy (1994) Stacey (1991) Thietart & Forgues (1995)	Organisations viewed as complex, nonlinear, dynamic systems that exhibit characteristics of both chaos and stability. High levels of turbulence lead to chaos that makes long-term forecasting practically impossible. Short-term planning enhanced by simulation.
Postmodernism	Baudrillard (1986) Harland (1987) Rosenau (1992)	Truth is unknowable. Attempts at long-range planning like 'navel gazing'. Only stifle organisation creativity.

In the turbulent social environment, who could have predicted the myriad stakeholder coalitions that pressured Nestlé, SA to stop marketing infant formula to third world countries as a substitute for breast feeding? Or the focus of the international anti-apartheid movement on the shareholder resolutions which pressured many multinational corporations to withdraw from South Africa (Mangaliso 1992)? Evidently, social responsibility, or the imperatives derived therefrom, has the potential to generate turbulence, as Exxon and Union Carbide have learned.

The overwhelming evidence of turbulence in real-life situations, as well as the fundamental impact it has on planning systems, has led many scholars to concur that the environment is shaped by specific conditions within it, and that better understanding of organisational dynamics requires a fuller understanding of the determinants of environmental contexts, i.e. the economic, technological, sociopolitical and ecological environmental contexts (Glueck 1976; Whittaker 1978).

Criticisms of turbulence

While turbulence has been studied intensely, its premises have not gone unquestioned. As has been pointed out, the metaphor of turbulence was borrowed from the natural sciences by Emery & Trist (1965). While it offers explanatory value, it may not be invested with the analytical rigour that goes into the making of a theory. Scholars have criticised the usage of turbulence in the management literature on two premises: relativity and metaphorical generalisability.

Relativity

Some scholars (McCann & Selsky 1984; Mintzberg 1991; Woodward 1982) point out that the Emery & Trist typology does not give a complete account of the real-world manifestation of the environment. They do not see turbulence as a universal phenomenon. Indicating that turbulence is not a discrete state passed through by all members of the environment, these authors assert that turbulence is a relative phenomenon. Other organisations may see dynamic, complex conditions simply as opportunities for innovation and growth. The major shake-out in the computer industry and its move away from mainframes may be a serious threat to IBM, but a golden opportunity for corporations like Dell and Compaq. In short, turbulence is in the mind of the strategist. Mintzberg (1991, 1994b) has gone so far as to state that turbulence is a figment of the imagination since no one has made an attempt to measure it. Still others (Ansoff 1991; Mangaliso & Lane 1992) have argued that because of the evidence provided by upheavals in several industries in the late 1980s and 1990s, turbulence must be studied.

Metaphorical generalisability

In the physical sciences, turbulence is the movement of fluid masses in all directions, with much mixing and many collisions. When this happens, eddies, vortices, waves, and 'white water' are some of the patterns that form. Turbulent flows occur when the combination of density, depth, and velocity overcomes the viscosity (Hurlbut 1976). A closer reading of the physics literature reveals that two kinds of turbulence are distinguishable: streaming flows and shooting flows. The Emery & Trist (1965) metaphor of turbulent environments seems to be based on streaming flows. Shooting flows, as generated in rapids and falls at very high velocities, may have opposite implications for management, whereby collaborative and planning-based systems are indicated. Although Emery & Trist do refer to the system developing collaboration and other adaptive mechanisms for coping with increased turbulence, they do not explain what happens when the coping capacity of the members is completely overwhelmed. This issue was taken up by McCann & Selsky (1984) who proposed the existence of another type of environment that they called the *Type 5* or hyperturbulent environment. In this environment, the degree of turbulence is seen to have exceeded the coping capacity of the system as a whole. New and radically different phenomena such as social triage begin to manifest themselves. Coping mechanisms in hyperturbulent environments include the formation of social enclaves by survivors and social vortices that suck the failed organisations down. On the whole, corporate strategies for dealing with hyperturbulent environments can be seen in such real-world phenomena as mergers (social enclaves) and divestitures (social vortices). The notions of environmental turbulence and hyperturbulence have important implications for the strategy making process.

Responses to turbulence

What, then, are the strategic implications of environmental turbulence? It would seem that in an atmosphere characterised by turbulence, the challenge lies in the ability to create meanings that would reduce equivocality and facilitate the emergence of real and actionable choices. Under such conditions, for a firm to sustain its entrepreneurial identity, it is appropriate to revisit a number of implications in the organisational framework.

The field of strategic management offers many recommendations for coping with changing environments. These may be arranged into two categories: *normative* responses and *pre-emptive* responses. Normative responses are those which follow strategies prescribed by conceptual theorists and proved to work through empirical research and observation. In pre-emp-

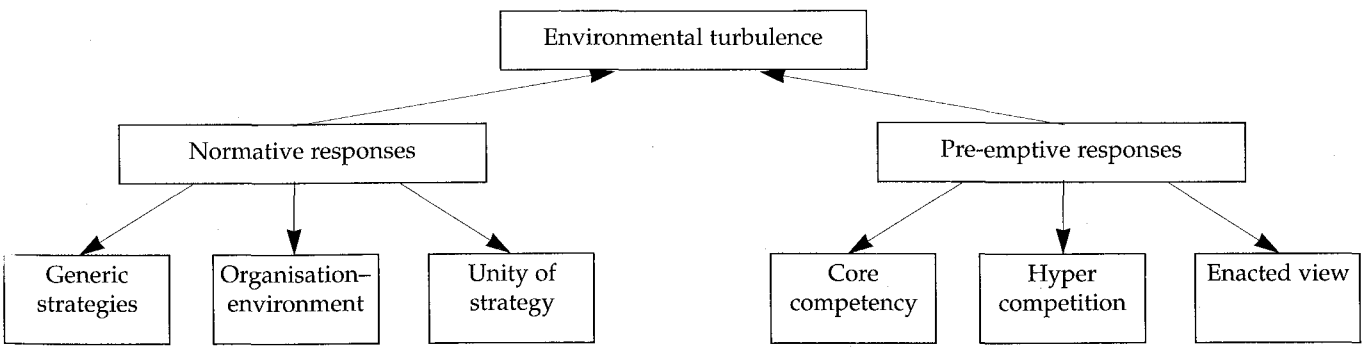

Figure 2. Categories of responses to turbulence

tive responses, organisations focus on specific segments of the environment where they will be able to direct unique strengths with which to mediate, dictate, and take advantage of environmental turbulence. These categories are presented in Figure 2.

Normative response perspective

The normative/prescriptive perspective primarily derived from the works of Porter (1980, 1985), Miles & Snow (1978, 1994), and other scholars who have subjected the prescribed typologies to empirical analysis (Anderson & Zeithaml 1984; Hill 1988; Jones & Butler 1988; Miller 1988; Nayyar 1994). While these theorists do not agree as to what strategic actions are appropriate under the different environmental conditions, what they propose can be summed up into three categories, namely generic strategies, organisation–environment fit, and unity of strategy.

Generic strategies: Given the likelihood of intense competitive rivalry among firms, the basic focus of the firm should be to develop a position of competitive advantage in the industry, either through a variation of economies of scale as is seen in the case of defenders and cost-leaders, or through some variation of economies of scope that is evident in prospectors and differentiators (Miles & Snow 1978; Porter 1985).

Organisation–environment fit: The successful firm is one that is able to develop a two-way fit between itself and the environment. This it does through 'forward-fit' strategies such as product–market strategies, R&D investment, and environmental scanning, and 'retro-fit' strategies such as streamlined organisational structure, adaptive control processes, and optimal production systems.

Unity of strategy: While mixed strategies exist in a diversified firm, it is important for the firm to know and implement a *master* strategy that will guide the organisational structure. Organisational structure often reflects a single strategy (Hill, Hitt & Hoskisson 1992), which is imperative for unambiguity.

While the normative perspective towards turbulence provides a valuable framework for analysis and strategising, it is perceived to be weak on several counts. The notion of a formal strategy has come under severe attack by the proponents of 'logical incrementalism' (Quinn 1978), who suggest that total posture planning for firms is a waste of corporate energy, and should be eschewed in favour of a more inductive and incremental approach to precipitating events. Theorists of 'emergent strategy' (Mintzberg 1987) advocate a 'crafting strategy' where the previous action of managers alters their future action as surely as a potter's tacit knowledge shapes every pot in a uniquely different manner.

Pre-emptive response perspective

This is a contingency-based perspective that incorporates both the resource-based view (Barney 1986; Hansen & Wernerfelt 1989) and the core competency view (Prahalad & Hamel 1990) of the firm. These views see the firm more as a bundle of imperfectly substitutable resources and knowledge-based strengths that need to be harnessed and sustained in order for the firm to grow and prosper. Theorists who hold this perspective believe that organisational cultures (Barney 1986), routines (Nelson & Winter 1992), core competencies (Prahalad & Hamel 1990), competition-based advantages (D'Aveni 1994), human resource systems (Lado & Wilson 1994), and entrepreneurial strengths (Rumelt 1987) reside uniquely within the individual firm. Firms that develop them as a strategic weapon may be in a stronger position to reap advantages from turbulence than those that adopt a reactive posture. Despite the disparity in the interest areas of these scholars, certain salient concepts emerge, including core competencies, hypercompetition, and enactment focus.

Core competencies: Conceptualised mainly as essential strengths, core competence translates into a map of action (Prahalad & Hamel 1990). The core competence enables an organisation to shape its environmental reality, focus its inputs, build on innovative and entrepreneurial strengths, foster learning systems, promote organisational culture, develop a corporate image, and create human resource systems to enhance such competence (Lado & Wilson 1994).

Hypercompetition: Though the term hypercompetition is primarily associated with the work of Richard D'Aveni (1994), various elements of hypercompetition logically connect responses based on the pre-emptive view. Normative/prescriptive responses to turbulence characterised the fading competitiveness of US firms in the 1980s. Hypercompetitive firms not only develop their advantages they actively direct them against the strengths and weaknesses of the competition. Arguing that every competitive advantage is ultimately eroded, the theorists of hypercompetition argue for the need to aggressively position a product in consumers' minds, often by displacing the competition (D'Aveni 1994; Ries & Trout 1986).

Enacted view: While the normative perspective carries static connotations, and has often been seen as a framework of responsive strategies, the pre-emptive perspective is perceived as representing a more 'enacted' view of the environment (Weick 1979). Firms are seen to create their environments, projecting the cognitive maps of their top managers, who are seen as repositories of their dominant logic (Prahalad & Bettis 1986). The environment is thus stripped of its veneer of objectivity and autonomy, and provincialised within the actions of organisational members (Smircich & Stubbart 1985). Therefore, the pre-emptive response perspective suggests that solutions to environmental turbulence lie *within* the organisation rather than in the fictive environment.

While pre-emptive strategies are intuitively appealing, they have not been supported by as much empirical research as the normative theories. Pre-emptive prescriptions may be more difficult to operationalise for research; for the same reason they are more effective in operation.

Towards a framework of responsiveness

As the preceding sections suggest, responses to turbulence traverse a contingent and negotiated realm. Organisations need to make customised choices depending on internal resources and external contingencies. Given the contested nature of the debate, it is important to resolve the normative and pre-emptive perspectives so that responses to turbulence are intelligible to both practitioners and researchers.

In suggesting a framework of responsiveness to turbulence, it is important to anchor the concept in specific environmental contexts. In today's globalising world, organisations operate in environments that are characterised by a great amount of diversity, and this diversity cannot be fully captured in a single continuum. In this article, two environmental dimensions have been chosen that are believed to be most fundamental in their impact on the functioning of the business firm. Following Mitroff & Kilmann (1984), these dimensions have been named the technical-economic dimension and the sociocultural

dimension. While these dimensions are not exhaustive and may thus be said to be an oversimplification of the reality of the business environment, we believe that they achieve a useful parsimony. Unlike Mitroff & Kilmann (1984), it is believed that cognitively, technical-economic and sociocultural issues do not exist as opposite ends on the same continuum, but are independent dimensions.

The technical-economic dimension refers to the nonliving domain of the business environment, including technology, machinery, manufacturing, production, distribution, information, finance, accounting, and marketing. Indeed, TQM and reengineering may be seen both as responses to technical-economic turbulence by those who practice them, and as sources of turbulence to those who are victims of them. While it can be said that technical-economic turbulence is an all-pervading artifact of the business environment, its differential impact on different industries is apparent. For instance, the steel industry does not have to contend with the rapidity of technological innovation that is the norm in the computer industry. Similarly, marketing-based innovation would impact the beverage industry far more than it would the oil industry. Clearly, proactive strategic responses to the environment will be mediated by the nature of technical-economic turbulence present in the industry.

The sociocultural dimension deals primarily with the human domain of the business environment, which is driven mostly by noneconomic considerations. Humans act opportunistically – within bounded rationality – in the pursuit of their selfish goals and objectives (Williamson 1985). The sociocultural dimension encompasses political (in)stability, culture, morality, and ethics. Even with globalisation, organisations must alter their production, distribution and marketing strategies in response to local sociocultural quirks. Strategic decision making must respect the sociopolitical turbulence that pervades the global marketplace. Multinational corporations have far fewer issues of political uncertainty and policy upheaval in mature secondary markets like the EC than they do in the incipient market economies of Eastern Europe. Conversely, embryonic markets offer far greater potential for growth than mature markets.

Based on the above two dimensions, a framework of strategic choices in a turbulent environment is presented in Figure 3.

Quadrant 1: Placid–placid environment

The simplest environments are those characterised by both technical-economic and sociopolitical placidity. Such environments are usually seen in protected economies or in industries where entry barriers are very high, such as the diamond industry (Bain 1956). Here the strategic choice would be to focus on maximising the Ricardian rents that can accrue from the business. The top management team would be able to assume a long-range planning horizon focusing on such goals as resource generation, allocation, and growth. Firms operating in such environments may be best served by *bureaucratic* strategies, emphasising consolidation, control, and cost reduction to maximise efficiency and sustain competitive advantage. The term bureaucratic is considered to be appropriate, even though in contemporary popular literature it has been misconstrued to focus only on its negative aspects such as extreme pigeon-holing of responsibility, insensitivity, maliciousness, rigidity, and hostility (Goodsell 1985). There are also positive aspects of bureaucracy that ensure that millions of routine jobs are performed reliably, including mail delivery, approval of applications, quality control checks, and answering consumer complaints. The term bureaucracy have been used in the Weberian sense to connote division of labour, emphasis on efficiency, dedication to profession, and respect for authority and position in the chain of command (Weber 1958), all of which are positive attributes in environments that are placid.

Three key substrategies can be identified for firms operating in placid–placid environments. Given the certainty of resources in this situation, the issue of *resource allocation* becomes important. The success of a firm is predicated upon internal and reflexive activities aimed at improving efficiency, such as managing cash flows, inventory control, and labour relations. Similarly, according to transaction cost economics, the market or hierarchy question would favour the hierarchy because control would be internalised (Coase 1937; Williamson 1985). In the make-or-buy decision, where a firm has to weigh the benefits of integration against subcontracting

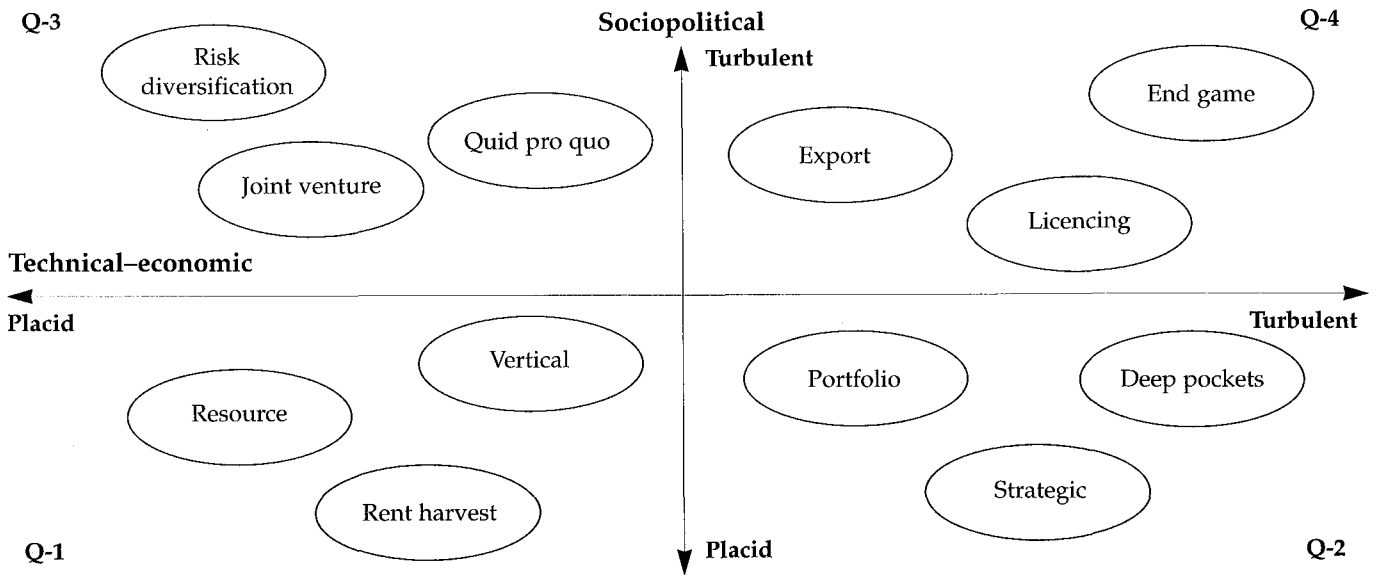

Figure 3. A framework of responsiveness to turbulence

or using other market agencies, the firm would choose to make something. Generally, the *vertical integration* strategy would be preferred, with the firm taking charge of upstream or downstream activities in the value chain (Porter 1985).

Finally, the placid–placid environment must be recognised as ephemeral. In a world characterised by creative destruction, such a benign environment may conceal the storm that follows the calm (Schumpeter 1934). Organisations operating in placid–placid environments therefore are advised to exploit their *rent harvesting* strategies before the available rents are eroded by competitive entry. This may be very clearly seen in the pharmaceutical industry where firms try to maximise the rents from various star drugs while they are patent-protected, and then invest these rents in product development in anticipation of severe falls in market share and revenues following the expiration of patent protection. The foregoing discussion may be summed up in the following three propositions:

P1: In placid–placid environments, successful firms focus on resource allocation, which includes managing cash flows, inventory, and labour deployment.

P2: In placid–placid environments, firms will seek to minimise costs and maximise efficiency through vertical integration.

P3: Because of the ephemeral nature of placid–placid environments, successful firms will adopt rent harvesting strategies in order to prepare for future uncertainties.

Quadrant 2: Placid–turbulent environments

While the placid–placid environments may have been artificially prevalent in the 1970s with the best examples in the command economies of Eastern Europe, such environments are few and even more ephemeral these days. Many organisations that operated in Eastern Europe's command economies have had to deal with fundamental paradigm shifts that they were scarcely equipped for. In market economies, technology has often been observed to step in and upset the homeostatic conditions that would normally be found in a mature marketplace.

A scenario in which sociopolitical placidity coexists with technical-economic turbulence – a scenario that describes a significant section of the US economy – gives rise to the concept of *option-based strategies*. The strategic choices of organisations operating in such environments can be viewed through what has come to be known as the 'option lens' (Bowman & Hurry 1993), which considers strategy as a process of resource deployment options. A superset of strategic options is created, which can be held in abeyance until the right opportunity arises. When this occurs, full-scale deployment of resources is unleashed. The option lens offers a better view of organisational resources in placid–turbulent environments.

Three key substrategies can be identified in the option-based realm. Firstly, *portfolio analysis* can be seen as one way of keeping various strategic choices alive and viable. A portfolio may be seen at once as the firm's investment in a basket of securities (Majd & Pindyck 1987; Myers 1984), or as the actual, minimal presence in various industries in anticipation of future increased involvement (Bowman & Hurry 1993; Hurry, Miller & Bowman 1992). For example, various pharmaceutical companies keep minimal investments in a variety of biotechnology firms in anticipation of major breakthroughs. Such a strategy is especially useful when the organisation is operating in a broad industry environment characterised by technological upheaval, like the computer industry.

This strategic choice is also influenced by the amount of capital at a firm's disposal. The amount of available resources is not entirely a function of a firm's size, but of the existence of *deep pockets* (D'Aveni 1994). The issue is not how much a firm has, but how the reserves are deployed within a universe of opportunities. Deep pockets provide the ability for swift global entries, defensive price strategies and political leverage in a hypercompetitive world.

Thirdly, firms that exist in environments characterised by technical-economic turbulence would benefit from synergistically exploiting the advantages of *strategic alliances* (Gulati 1995; Parkhe 1993). Strategic alliances have been proliferating at an exponential pace in the 1990s, and one reason for this has been the need to pool organisational resources, especially in high financial or intellectual capital arenas, where individual entry is daunting. For instance, in the pharmaceutical industry, where the cost of new product innovation has been estimated to be in excess of $500 million (Malknight 1995), it is not uncommon to see two large multinational corporations announce collaborative research proposals to work for new drug development. Such an arrangement is usually followed by either comarketing arrangements or the division of the global marketplace for marketing rights. In the semiconductor industry, which can be characterised as turbulent because of rapid, unpredictable technological changes, firms have increasingly turned to collaborative and cooperative interorganisational ties for research and development (Aldrich & Sasaki 1994; Tushman & Anderson 1986). Sometimes the costs become so great that the government must step in to facilitate the process, as seen in the case of the semiconductor industry in the US. A government-supported consortium, SEMATECH, was established in 1987 when it became obvious that none of the individual firms had sufficient resources to commit at the scale required to remain competitive with Japan (Spencer & Grindley 1993). SEMATECH may have been influential in sparking a resurgence of the US semiconductor industry by 1992. In sum, strategic alliances form a powerful tool for operating in technical-economically volatile markets, but usually need an environment of sociopolitical stability to be fully effective. The following three propositions follow from the above discussion:

P4: The most effective strategies in environments characterised by sociopolitical placidity and technical-economic turbulence are those based on portfolio analysis techniques.

P5: In environments with sociopolitical placidity and technical-economic turbulence, firms enhance the long-run sustainability of their competitive advantage by diversifying their investments.

P6: Firms that compete in environments with sociopolitical placidity and technical-economic turbulence enhance their competitive positions by engaging in collaborative strategies such as strategic alliances.

Quadrant 3: Turbulent–placid environments

The issue of sociocultural turbulence in an atmosphere of relative technical-economic stability is often seen in the case of multinationals which are planning to expand into areas that are relatively unexplored for their product range, with a severe shortage of local and specific information. The foray of the

Western multinationals into Eastern Europe in the mid-1980s was characterised by this tentativeness. It is argued that such situations offer *collaborative strategies* as the most logical strategic choice for organisational survival and success.

The unequal distribution of knowledge among organisations can best be tackled by entering into partnerships, which may offer both stability and a pool of knowledge. However, given the delicate and transient nature of some strategic alliances, they may not be the most suitable form of partnership. Organisations operating in sociopolitical turbulence but technical-economic stability may need to concede the 'political imperative' (Doz & Prahalad 1991) to the local subsidiary in the joint venture.

Three substrategies may be identified in this realm. The first, *quid pro quo*, holds the local subsidiary accountable for organisational performance, either through equity sharing or another form of compensation. For instance, Pepsi was able to enter Russian markets with a franchising arrangement despite the extreme shortage of hard currency in Russia, because it actively helped the Russian government to find international markets for naval defence equipment.

Such a strategy may not always be feasible, especially in regions that lack a dominant business with which to enter into a quid pro quo relationship. Another strategy is to enter into a joint venture with the organisation, whereby a third organisation is created. When liability for the *joint venture* is limited to the newly created entity, turbulence may be minimised. Most organisations operating in the Middle East surrender majority equity rights to their local subsidiaries in exchange for marketing rights and bureaucratic liaison.

In environments subjected to selective, industry-based sanctions, companies may reduce their risks by a conglomerate diversification strategy, spreading their investment into a variety of areas. For instance, Japanese organisations facing a threatening US tariff escalation on luxury cars would benefit from active presence in other segments of the automobile market, or from diversifications like those of Honda (generators, lawn mowers). General Electric Company, faced with protests to get out of nuclear warhead manufacturing, benefitted from their acquisition of the National Broadcastic Corporation (NBC). The following three propositions sum up the thoughts presented above:

P7: In environments characterised by sociocultural turbulence and technical-economic placidity (Q-3), the most appropriate strategies are the *quid pro quo* strategies in which a firm gives in order to receive.

P8: In order to stave off the effects of sociocultural turbulence in Q-3 environments, the most appropriate strategy to adopt is the joint venture.

P9: To reduce risk in Q-3 environments, firms seek unrelated or conglomerate diversification.

Quadrant 4: Turbulent–turbulent environments

When sociopolitical turbulence is accompanied by a great deal of technical-economic turbulence, the chaotic system that develops makes long-range planning impossible. This is the quadrant where learning becomes important, a situation which gives credence to Mintzberg's (1991) scorecard of 'Learning: 1, Planning: 0'. Here the firm must develop *distancing strategies*, whereby the gains of entry and presence are significantly given up in an attempt to reduce the risk associated with operating in highly uncertain markets. For instance, firms in industries such as computer software, music, and motion picture distribution find it difficult to operate marketing alliances in relatively unknown markets because of lax copyright laws and the consequent fear of infringement of intellectual property rights. Distancing strategies may be seen as traversing a spectrum, from an organisation seeking to limit its knowledge-based and capital-based commitment to a region, to outright divestiture of resources in an attempt to protect existing markets.

One strategy that may be adopted by new entrants into markets in the turbulence–turbulence quadrant is to develop *export relations* with local trading houses. Such arrangements are often characterised by low margins and tenuous market presence, but offer at least a limited opportunity in an area where a more organic presence is unfeasible. On the other hand, for firms already present in the market, technology may be *licensed* out to other areas, where potential is limited, but misuse is nonthreatening. Many Hollywood movies, which outlive their product life cycle as entertainment vehicles in the US and Europe, generate revenue in global markets. Their simultaneous release in the US would have risked piracy. Such strategies, when misused, may lead to the release of substandard, or even dangerous products to third world countries. This has been observed in the pharmaceutical industry, with the marketing of drugs banned by the World Health Organisation. Such strategies may backfire on their perpetrators as the global marketplace leads to globalised information exchange and its image consequences.

When neither export options nor licencing arrangements are feasible, a firm may consider the option of *endgame strategies*, whereby total divestiture from a particular market may be necessary either to prevent further escalation of losses or to protect proprietary information. The decision by the EC not to do business with non-ISO9000 certified firms might be a response to environments that are turbulent on both dimensions of Figure 3. Another example is offered by Polaroid Corporation and later Eastman Kodak who decided to exit the South African market when it became clear that their products were being used to further the political objectives of the apartheid regime. The following propositions capture the essence of the preceding ideas:

P10: Firms entering turbulent–turbulent, or Q-4, environments will seek export relations with local trading partners.

P11: For firms already operating in Q-4-type environments, the best way to protect their product and preserve their presence in the market is through licensing.

P12: When conditions of turbulence escalate beyond the organisation's coping capacity, the most appropriate strategies are endgame strategies such as divestiture that help to minimise losses.

Implications

Before discussing the implications of the framework offered in this article, it is important to acknowledge its weaknesses. The framework suffers from the same weakness as others (e.g. the BCG matrix) in that it is simplistic: a two-by-two representation of reality is bound to leave out many of the in-between

conditions, the grey areas. For instance, between the extremes of placidity and turbulence, Emery & Trist (1965) suggested two other states: placid randomised and disturbed reactive environments. If these two were to be added, we would end up with a much more complex four-by-four matrix. A two-by-two matrix was chosen in the interest of parsimony. As Mitroff & Kilmann (1984) note, it is the essence of the discussion rather than the subtleties that is important. Others have advocated simplification as a necessary starting point in the process of theory development (Lant & Mezias 1990). Obviously, the strategic responses are not strictly confined to their categories in the model. Rather, these are likely associations, and the suggested propositions need to be evaluated and empirically tested.

Since environmental turbulence is now a fact that organisations have to live with, several implications may be drawn. As a recent *Fortune* magazine cover story warned, organisations will have to 'hold on, it's going to be a bumpy ride' (Richman 1995). Firstly, managers and researchers must recognise that turbulence will place demands on organisations that are not necessarily consistent with their organisational strategies. This implies that organisation leaders must infuse a culture of adaptations, flexibility, and speed of response. Secondly, managers will continuously have to explore new ways of doing things through innovation, entrepreneurship, and experimentation. It will be necessary from time to time to establish temporary equilibria as a precondition for adapting to the vagaries of a constantly changing environment. Thirdly, the increased unpredictability of the environment implies that managers must have in their armour a battery of responses that they can utilise when needed. Furthermore, a new organisational paradigm would permit the development of catalogues of configurations from which organisations will have to choose when environmental forces are more powerful than organisational viscosity (Thietart & Forgues 1995). Fourthly, the complexity, diversity, and turbulence in the environment mandate a duplication of those conditions within organisations – a sort of 'law of requisite turbulence' to paraphrase W. R. Ashby (1956). As stated earlier, organisations must have the right mix of stability and chaos to be able to survive under conditions of turbulence.

Implications of a paradigmatic shift as a prerequisite for managing under turbulent conditions are profound. For one thing, the process of organisational sense making – in the same sense as the economic principle of comparative advantage – seems to suggest that when two firms are adept at making sense of different environments, they can offer critical support to each other in conditions of turbulence. This does not mean firms must coalesce, alter market structure, or arrange monopoly power as a prescription for survival. It means that in a world of 'interculturalising' business interests, a firm entering a new market often gains greater advantage in a collaborative relationship with another firm which may already be in possession of the necessary adaptive information. Examples of such strategic alliances include AT&T with Olivetti (Italy), and MCI with British Telcom in telecommunication; US AIR with British Airways, and Continental Airways with Alitalia in the airline industry; and Alfa-Romeo (Italy) with Nissan (Japan) in automobiles.

The impact of a combination of sociopolitical and technical-economic forces originating in the organisation's environment is best illustrated by South Africa in the 1980s when a number of multinational corporations reluctantly succumbed to the divestment movement and withdrew their business from that country (Mangaliso 1988). In Eastern Europe, the creation of new 'market economies' has led to a variety of options for diversifying multinational corporations. Within organisations, it has been demonstrated that cultural heterogeneity can play a part in economic performance (Franke 1991). Thus, a firm's strategic responses must respect not only the technical-economic imperatives of survival, but the sociopolitical imperative of legitimacy as well. The firm that is prepared for the impact of public policy is most suited to survival.

Conclusion

It has often been observed that organisations operating in turbulent environments neglect experimentation, risk taking, and the examination of noncore activities. The planning process becomes short-term crisis management, making long-term strategic planning a casualty. The most difficult job for a management team facing turbulence may be the traditional one: re-evaluating and redefining what business they are in, what other businesses they should be in, or even whether they should be in a particular business at all. They should maintain the methods and techniques that help them to succeed in relatively placid environments, while being vigilant of signals that indicate a paradigm shift for emerging environmental contingencies. The graveyards of organisation history are filled with 'successful' organisations which failed to undertake paradigm shifts in conditions of turbulence. It is perhaps this activity, above all, which determines whether a firm succeeds or even survives. Therefore the authors concur with those who believe that the only sustainable competitive advantage organisations have in turbulent environments is the ability to enhance learning (Mintzberg 1991; Stata 1989). Learning will be enhanced through discourse about how organisations will function in the ever-increasing turbulence of tomorrow's environments. It is hoped that this article will advance that process.

References

Aldrich, H. E. 1979. *Organizations and environment*. Englewood Cliffs, NJ.: Prentice-Hall.

Aldrich, H. E. & Pfeffer, J. 1976. 'Environments of organizations', *Annual Review of Sociology*, 2. Palo Alto, CA.: Annual Review.

Aldrich, H. E. & Sasaki 1994. 'R & D consortia in the United States and Japan', Paper presented in the Distinguished Visiting Scholars Series, School of Management, University of Massachusetts, Amherst, MA. Forthcoming in *Research Policy*.

Anderson, C. R. & Ziethaml, C. P. 1984. Stage of the Product Life Cycle, Business Strategy, and Business Performance, *Academy of Management Journal*, 27(1): 5–24.

Andrews, K. R. 1971. *The concept of corporate strategy*. Homewood, IL.: Irwin.

Ansoff, H. I. 1984. *Implanting strategic management*. Englewood Cliffs, NJ.: Prentice-Hall.

Ansoff, H. I. 1991. "Critique of Henry Mintzberg's 'The design school: reconsidering the basic premises of strategic management'", *Strategic Management Journal*. 12(6): 449–461.

Ashby, W. R. 1956. *An introduction to Cybernetics*. New York: Wiley.

Bain, J. S. 1956. *Barriers to new competition: their character and consequences in manufacturing industries*. Cambridge, MA.: Harvard University Press.

Barney, J. B. 1986. 'Strategic Factor Markets, Expectations, Luck, and Business Strategy', *Management Science*, 42: 1231–1241.

Baudrillard, J. 1986. 'Forgetting Baudrillard', *Social Text*. 15 (Fall): 140–144.

Baumol, W. & Benhabib, J. 1989. 'Chaos, significance, mechanism, and economic applications', *Journal of Economic Perspectives*, 3: 77–105.

Bourgeois, L. J. 1978. 'The environmental perceptions of strategy makers and their economic correlates', Working Paper 273. Graduate School of Business, University of Pittsburgh, Pittsburgh, PA. (March).

Bourgeois, L. J. 1985. 'Strategic goals, perceived uncertainty, and economic performance in volatile environments', *Academy of Management Journal*, 28(3): 548–573.

Bourgeois, L. J., McAllister, D. W., & Mitchell, T. R. 1978. 'The effects of different organizational environments upon decisions about organizational structure', *Academy of Management Journal*, 21: 508–514.

Bowman, E. & Hurry, D. 1993. 'Strategy through the option lens: integrated view of resource investments and the incremental choice process', *Academy of Management Review*, 18(4): 760–782.

Burns, T. & Stalker, G. M. 1961. *The management of innovation*. London: Tavistock.

Cameron, K. S., Kim, M. U., & Whetten, D. A. 1987. 'Organizational Effects of Decline and Turbulence', *Administrative Science Quarterly*, 32: 222–240.

Cameron, K. S., Sutton, R. I. & Whetten, D. A. 1988. 'Issues in organizational decline'. In K. S. Cameron, R. I. Sutton & D. A, Whetten (eds.), *Readings in organizational decline: Framework, research, and prescriptions*: 3–19. Boston, MA.: Ballinger Publishing Company.

Camillus, J. C. & Datta, D. K. 1991. 'Managing strategic issues in a turbulent environment', *Long Range Planning*, 24(2): 67–74.

Child, J. 1972. 'Organizational Structure, Environment, and Performance: The Role of Strategic Choice', *Sociology*, 6: 1–21.

Coase, R. 1937. 'The nature of the firm', *Econometrica*, 4: 386–405.

D'Aveni, R. A. 1989. 'The aftermath of organizational decline: A longitudinal study of the strategic and managerial characteristics of declining firms', *Academy of Management Journal*, 32(3): 577–605.

D'Aveni, R. A. 1994. *Hypercompetition: Managing the dynamics of strategic maneuvering*. New York: Free Press.

Doz, Y. & Prahalad, C. K. 1991. 'Managing DMNCs: A search for a new paradigm', *Strategic Management Journal*, 12 (Summer): 145–164.

Duncan, R. B. 1972. 'Characteristics of Organizational environments and Perceived Environmental Uncertainty', *Administrative Science Quarterly*, 17: 313–327.

Ellis, R. J. 1982. 'Improving Management Response in Turbulent Times', *Sloan Management Review*, 23(2): 3–12.

Emery, F. E. & Trist, E. L. 1965. 'Causal Texture of Organizational Environment', *Human Relations*, 18: 21–32.

Faur, T. 1993. 'Foreign Capital in Hungary's Privatization'. *In Privatization in the transition to a market economy*. New York: St. Martin's Press: 203–214.

Gersick, C. J. G. 1991. 'Revolutionary change theories: A multilevel exploration of the punctuated equilibrium paradigm', *Academy of Management Review*, 16(1): 10–36.

Glueck, W. F. 1976. *Business policy: strategy formulation and management action*. New York: McGraw-Hill.

Goodsell, C. 1985. *The case for bureaucracy: a public polemic*. Chatham, NJ.: Chatham House.

Goodstein, J. & Boeker, W. 1991. 'Turbulence at the Top: A New Perspective on Governance Structure Changes and Strategic Change', *Academy of Management Journal*, 34(2): 306–330.

Greenhalgh, L., Lawrence, A. T. & Sutton, R. I. 1988. 'Determinants of work force reduction strategies in declining organizations', *Academy of Management Review*, 13(2): 241–254.

Gulati, R. 1995. 'Does familiarity breed trust? The implication of repeat ties for contractual choice in alliances', *Academy of Management Journal*, 38(1): 85–112.

Haleblian, J. & Finkelstein, S. 1993. 'Top Management Team Size, CEO Dominance, and Firm Performance: The Moderating Roles of Environmental Turbulence and Discretion', *Academy of Management Journal*, 36(4): 844–863.

Hammer, M. & Champy, J. 1993. *Reengineering the corporation: a manifesto for business revolution*. New York: Harper Business.

Hansen, G. S. & Wernerfelt, B. 1989. 'Determinants of Firm Performance: The Relative Importance of Economic and Organizational Factors', *Strategic Management Journal*, 10: 399–411.

Harland, R. 1987. *Super structuralism: the philosophy of structuralism and post structuralism*. London: Metheun.

Harrigan, K. R. 1985. *Strategic flexibility: a management guide for changing times*. Lexington, MA.: D.C. Heath.

Heenan, D. A. & Perlmutter, H. V. 1979. Multinational development: a social architecture perspective. Reading, MA.: Addison-Wesley.

Hill, C. W. L. 1988. 'Differentiation versus low-cost or differentiation and low-cost: A contingency framework', *Academy of Management* Review, 13(2): 159–178.

Hill, C. W. L., Hitt, M. A. & Hoskisson, R. E. 1992. 'Cooperative versus competitive studies in related and unrelated diversified firms', *Organization Science*, 3(4): 501–521.

Hrebiniak, L .G. & Joyce, W. F. 1985. 'Organizational adaptation: Strategic choice and environmental determinism', Administrative *Science Quarterly*, 30: 336–349.

Huber, G. 1984. 'The nature and design of post-industrial environments', *Management Science*, 30: 928–951.

Huber, G. P., Miller, C. C. & Glick, W. H. 1990. 'Developing more encompassing theories about organizations: The centralization-effectiveness relationship as an example', *Organization Science*, 1: 11–40.

Hurlbut, C. S. 1976. *The Planet we live on*. New York: Harry Abrahams.

Hurry, D., Miller, A. J. & Bowman, E. H. 1992. 'Calls on high-technology: Japanese exploration of venture capital investments in the United States', *Strategic Management Journal*. 13(2): 85–101.

Jemison, D. B. 1981. 'Organizational versus environmental sources of influence in strategic decision making', *Strategic Management Journal*, 2(1): 77–90.

Johnson, S. & Loveman, G. 1995. 'Starting Over: Poland After Communism', *Harvard Business Review*, (March–April): 41–58.

Johnston, S. 1995. 'Managerial Dominance of Japan's Major Corporations', *Journal of Management*, 22(2): 191–204.

Jones, G. R. & Butler, J. E. 1988. 'Costs, Revenue, and Business-Level Strategy', *Academy of Management Review*, 13(2): 202–213.

Katz, D. & Kahn, R. L. 1966. *The social psychology of organizations*. Chisester: John Wiley and Sons.

Keats, B. & Hitt, M. 1988. 'A causal model of linkages among environmental dimensions, macro-organizational characteristics and performance', *Academy of Management Journal*, 31: 570–598.

Khandwalla, P. N. 1978. 'Crisis response of competing versus non-competing organizations'. In C. F. Smart & W. T. Stanbury (eds.): 158–178. *Studies on crisis management*. Toronto: Butterworth.

Klimoski, R. 1994. "A 'Total Quality' Special Issue", *Academy of Management Review*, 19(3): 390–392.

Lado, A. A. & Wilson, M. C. 1994. 'Human Resource Systems and Sustained Competitive Advantage: A Competency-Based Perspective', *Academy of Management Review*, 19(4): 699–727.

Lant, T. K. & Mezias, S. J. 1995. 'Managing continuous change: A simulation study of organizational learning and entrepreneurship', *Strategic Management Journal*, 11 (Special Issue): 147–179.

Lawrence, P. R. & Lorsch, J. W. 1967. *Organization and environment*. Boston, MA.: Harvard Business School.

Lenz, R. T., & Engledow, J. L. 1986. 'Environmental analysis: the applicability of current theory', *Strategic Management Journal*, 7(4): 329–346.

Levy, D. 1994. 'Chaos Theory and Strategy: Theory, Application, and Managerial Implications', *Strategic Management Journal*, 15 (Summer): 167–178.

MacCann, J. E. & Selsky, J. 1984. 'Hyperturbulence and the emergence of type 5 environments', *Academy of Management Review*, 9(3): 460–470.

Mahoney, J. T. & Pandian, J. R. 1992. 'The resource-based view of within the conversation of strategic management', *Strategic Management Journal*, 13: 363–380.

Majd, S. & Pindyck, R. S. 1987. 'Time to build, option value and investment decisions', *Journal of Financial Economics*, 18: 7–27.

Malknight, T. W. 1995. 'Globalization of an ethnocentric firm: An evolutionary perspective', *Strategic Management Journal*, 16(2): 119–142.

Mangaliso, M. P. 1988. 'The relationship of environmental turbulence, strategy preference and performance: A study of the perceptions of North American and South African corporate executives', Unpublished Ph.D. Dissertation. Amherst, MA.: University of Massachusetts.

Mangaliso, M. P. & Lane, P. J. 1992. 'The role of turbulence in the strategic management literature'. *Proceedings*. Eastern Academy of Management Meeting. Baltimore, MD.

Metcalfe, J. L. 1974. 'Systems models, economic models and the causal texture of organizational environments: An approach to macro-organization theory', *Human Relations*, (September): 639–663.

Myers, S. C. 1984. 'Finance Theory and Financial Strategy', *Interfaces*, 14(1): 126–137.

Meyers, G. C. & Holusha, J. 1986. *When it hits the fan: managing the nine crises for business*, Boston: Houghton Mifflin.

Miles, R. E. & Snow, C. C. 1978. *Organizational strategy, structure, and process*. New York: McGraw-Hill.

Miles, R. E. & Snow, C. C. 1994. *Fit, failure, and the hall of fame*. New York: The Free Press.

Miller, C. C. & Cardinal, L. B. 1994. 'Strategic Planning and Firm Performance: A Synthesis of More Than Two Decades of Research', *Academy of Management Journal*, 37(6): 1649–1665.

Miller, D. 1992. 'Environmental fit versus internal fit', *Organization Science*, 3(2): 159–178.

Mintzberg, H. 1990. 'The Design School: Reconsidering the Basic Premises of Strategic Management', *Strategic Management Journal*, 11: 171–195.

Mintzberg, H. 1991. 'Crafting Strategy'. In H. Mintzberg & J. B. Quinn (eds.), *The strategy concept*: 105–114. Englewood Cliffs, NJ.: Prentice-Hall.

Mintzberg, H. 1991. 'Learning 1, Planning 0: Reply to Igor Ansoff', *Strategic Management Journal*, 12(6): 463–466.

Mintzberg, H. 1994a. "That's not 'Turbulence', Chicken Little, It's Really Opportunity", *Planning Review*, 22(6): 7–9.

Mintzberg, H. 1994b. *The rise and fall of strategic planning: reconceiving roles for planning, plans, planners*. New York: The Free Press.

Mintzberg, H. & McHugh, A. 1985. 'Strategy formation in an adhocracy', *Administrative Science Quarterly*: 160–197.

Mitroff, I. I. & Pauchant, T. 1990. *We are so big and powerful nothing can happen to us*. New York: Caroll Publishing.

Mitroff, I. I. & Kilmann, R. H. 1984. *Corporate tragedies: product tampering, sabotage, and other catastrophes*. New York: Praeger.

Moor, J. F. 1993. 'Predators and prey: A new ecology of competition', *Harvard Business Review*, (May–June): 75.

Nayyar, P. 1993. 'On the measurement of competitive strategy: Evidence from a large, multi-product, U.S. firm', *Academy of Management Journal*, 36(6): 1652–1669.

Nelson, R.R. & Winter, S. 1982. *An evolutionary theory of economic change*. Cambridge, MA.: Harvard University Press.

Parkhe, A. 1993. 'Strategic alliance structuring: A game theoretic and transaction cost examination of inter-firm cooperation', *Academy of Management Journal*, 36(4): 749–829.

Pearce, J. A., Robbins, D. K. & Robinson, R. B. 1987. 'The impact of grand strategy and planning formality on performance', *Strategic Management Journal*, 8: 125–134.

Perrow, C. 1984. *Normal accidents*. New York: Basic Books.

Pfeffer, J. & Salancick, G. R. 1978. *The external control of organizations: a resource dependence perspective*. New York: Harper & Row.

Phelan, S. E. 1995. 'From chaos to complexity in strategic planning', Paper presented at the 55th Annual Meeting of the Academy of Management. Vancouver, BC., Canada. August 8.

Porter, M. E. 1980. *Competitive strategy*. New York: Free Press.

Porter, M. E. 1985. *Competitive advantage*. New York: Free Press.

Prahalad, C. K. & Bettis, R. A. 1986. 'The Dominant Logic: The New Linkage Between Diversity and Performance', *Strategic Management Journal*, 7: 485–501.

Prahalad, C. K. & Hamel, G. 1990. 'The core competence of the corporation', *Harvard Business Review*, 3: 79–91.

Quinn, J. B. 1991. "Strategic Change: 'Logical Incrementalism'". In H. Mintzberg & J. B. Quinn (eds.), *The strategy concept*: 96–104. Englewood Cliffs, NJ.: Prentice-Hall.

Richman, L. S. 1995. 'The Economy: Managing through a downturn', *Fortune*, August 7: 59–64.

Rosenau, P. M. 1992. *Postmodernism and the social sciences*. Princeton, NJ.: Princeton University Press.

Rubin, I. S. 1979. 'Retrenchment, Loose Structure, and Adaptability in Universities', *Sociology of Education*, 52: 211-222.

Rumelt, R. P. 1987. 'Theory, Strategy, and Change'. In D.J. Teece (ed.), *The competitive challenge*: 137–158. Cambridge, MA.: Ballinger.

Schumpeter. J. A. 1934. *The theory of economic development*. Cambridge, MA.: Harvard University Press.

Shrivastava, P. 1987. *Bhopal: Anatomy of Crisis*. Cambridge, MA: Ballinger Publishing.

Shrivastava, P. 1994. 'Castrated environment: greening organizational studies', *Organization studies*, 15(5): 705–721.

Shrivastava, P. & Mitroff, I. I. 1987. 'Strategic management of corporate crises', *Columbia Journal of World Business*, 22(1): 5–11.

Smart, C. F. & Vertinsky, I. 1977. 'Designs for Crisis Decision Units', *Administrative Science Quarterly*, 22: 640–657.

Smircich, L. & Stubbart, C. 1985. 'Strategic Management in an Enacted World', *Academy of Management Review*, 10: 724–736.

Snyder, N. H. 1987. 'Validating measures of environmental change', *Journal of Business Research*, 15(1): 31–43.

Snyder, N. H. & Glueck, W. F. 1982. 'Can Environmental Volatility be Measured Objectively?', *Academy of Management Journal*, 25: 185–192.

Spencer, W. J. & Grindley, P. 1993. 'SEMATECH after Five Years: High-technology consortia and U.S. competitiveness', *California Management Review*, 35(4): 9–32.

Stacey, R. D. 1991. *The chaos frontier*. Oxford: Butterworth-Heinemann.

Starbuck, W. G., Greve, A. & Hedberg, B. L. T. 1978. 'Responding to Crisis', *Journal of Business Administration*, 9: 111–137.

Starbuck, W. G., & Hedberg, B. L. T. 1977. 'Saving an organization from stagnating environments'. In H. Thorelli (ed.): 249–258. *Strategy + structure = performance*. Bloomington, IN.: Indiana University Press.

Stata, R. 1989. 'Organizational learning– the key to management innovation', *Sloan Management Review*: 63–74.

Staw, B. M., Sandelas, L. E. & Dutton, J. E. 1981. 'Threat–rigidity effects in organizational behavior: A multilevel analysis', *Administrative Science Quarterly*, 26: 501–524.

Terryberry, S. 1968. 'The evolution of organizational environments', *Administrative Science Quarterly*, 12: 590–613.

Theobald, R. 1994. 'New Success Criteria for a Turbulent World', *Planning Review*, 22(6): 10–13.

Tosi, H., Aldag, R., & Storey, R. 1973. 'On the measurement of the environment: An assessment of the Lawrence and Lorsch Environmental Uncertainty Subscale', *Administrative Science Quarterly*, 18: 27–36.

Tushman, M. L. & Anderson, P. 1986. 'Technological discontinuities and organizational environments', *Administrative Science Quarterly*, 31: 439–465.

Weber, M. 1958. *The protestant ethic and the spirit of capitalism*. New York: Charles Scribner's & Sons.

Weick, K. E. 1979. *The social psychology of organizing*. Reading, MA.: Addison-Wesley.

Whittaker, J. B. 1978. *Strategic planning in a rapidly changing environment*. Lexington, MA.: D.C. Heath.

Williamson, O. E. 1985. *Economic institutions of capitalism*. New York: Free Press.

Woodward, S. N. 1982. 'The myth of turbulence', *Futures*, (August): 266–279.

Zammuto, R. F. & Cameron, K. S. 1985. 'Environmental decline and organizational response' In L. L. Cummings & B. Staw (eds.), 7:223–262. *Research in organizational behavior*. Greenwich, CT.: JAI Press.

The impact of the national payment system on the South African payments industry

Joy A Rees & M Anton Ferreira

Unisa Graduate School of Business Leadership

The implementation of the new National Payment System (NPS) in the banking sector has significant implications for the four major banking institutions in South Africa. The payment industry, which consists of all organisations offering any service along the payment value chain, will be significantly impacted by the NPS. The payment and banking industries are undergoing significant changes worldwide because of such issues as globalisation, technology, customer sophistication, and deregulation. It is against this background that 17 people involved in the implementation of the NPS at various banks or organisations involved in payments were interviewed to assess the significance of the NPS. The study found that the NPS, together with the other forces facing the industry, opened numerous opportunities for the 'big-four' banks as well as imposed certain threats. In order to keep their stronghold over the payments industry, the banks would need to adapt to the new processes imposed by the NPS. They would need to become more competitive in the earlier stages of the payment value chain by offering value-added and innovative payment products otherwise they would be separated from their customer base – a previously captive market. Although currently still protected by fairly large barriers to the entry in the clearing and settling part of the value chain, even here the 'big-four' banks face significant new competition from the niche and foreign banks who might choose to enter this arena.

Introduction

This article assesses the impact that the implementation of the new National Payment System (NPS) in South Africa will have on the four major banking groups (Standard Bank, First National Bank, ABSA, and Nedcor), the nature of competition between them and other players, and the structure of the industry in general. The objective is to suggest steps that these banks need to take in order to exploit the opportunities and to reduce the risks associated with the NPS. In order to do this, it will be necessary to examine the present payments industry structure and the potential impact of the NPS on it, the present nature of competition in the industry, and possible responses of the four major banks to protect their existing positions.

The payments industry is a subset of the financial services industry. It includes any organisation offering payment services, such as banks and credit card houses, as well as all organisations that provide value along the payment value chain (including bureaus, merchant banks, retailers, et cetera). The many stages that a payment must go through in order to be properly effected typically include the initiation of the payment (usually by the client), the validation of the payment details, the capturing of these details, the processing of the payment, and, finally, the clearing and settling of cross-bank transactions. For many years, the payments value chain has been under the control of the major banks. However, because of the introduction of many more payment mechanisms, such as electronic funds transfers and the use of personal computers and other devices to effect payment, it has become possible for new entrants to compete for various parts of the payments value chain. Nevertheless, clearing and settling have remained largely the domain of the major banks, so far, arguably, protected by their huge investments in the infrastructure required for clearing (such as Bankserv), coupled with the capital requirements for settling.

The financial services industry worldwide is currently in a state of flux (Deloitte Touche Tohamatsu 1995; Vessey 1996). In South Africa, a number of forces for change have conspired to place enormous pressure on the industry. These include the effects of globalisation and the reduction of exchange control requirements (national banks now find themselves up against large, well-established banks from other countries) (Volschenk 1997), the increasing sophistication of bank customers who are increasingly demanding more specialised banking and payment products (*Financial Mail* 1997: 35), and the massive impact of new information technologies (such as networks, smartcards and the Internet) on the industry (Deloitte Touche Tohamatsu 1995: 13–25; Vessey 1996; Melody 1996; Rozwadowski 1996).

SBL Research Review 1998 2(2): 38-48
The impact of the national payment system on the South African payments industry

Overview of the payments industry in South Africa

The major banks in South Africa are all participants in the payments industry. Other players include the newer foreign and niche banks that have entered the industry, as well as retailers, car dealers, and payments bureaus. Typically, payments can be subdivided into payments by individuals (these payments are usually small but occur in large volumes), payments by businesses to individuals, and payments from one business to another. Payments for shares and foreign transactions typically result in the smallest number of payments, but generally represent the greatest value of transactions. Payment mechanisms include cash, cheque, card, bank-owned device (ATMs) and client-owned device (personal computer using modem and Internet).

The value chain concept (Porter 1985) presents a particularly useful angle on the payments industry. For this industry it consists of the origination of the payment, the validation of the payment, the capture of the payment, the capturing of customer information, and the transmission of the payment (Boston Consulting Group 1996). These steps are then followed by the processing and clearing of the payment and, finally, the settlement of the payment. The value chain allows the identification of the most profitable parts of the chain to control, which are where most competition is to be found, as well as gaps or niches that could be filled by new entrants.

As shown in Figure 1, the domain of clearing and settling has thus far strictly been the province of the banks. However, in the other segments of the value chain, many different players have already entered the arena and many more are possibly poised to do so in the near future. This is a drastic change from a few years ago, when the same figure would have reflected 'bank only' across all types of transactions along the value chain. Thus, it could be argued that the erosion of the banks' control of the industry is still in its initial stages. While the current competitive arena appears to be the first four stages of the value chain (see Figure 1), the final two stages could also become contested in the near future.

From a rivalry point of view, the four major banks in the country (Standard Bank, Absa Bank, First National, and Nedcor, hereafter referred to as the 'big-four') presently dominate the competitive arena (*Financial Mail* 1997: 36). They own the requisite infrastructure, have a multiplicity of payment products and mechanisms and, in most cases, other service providers rely on their provision of specific payment mechanisms as well as on their clearing and settling services. Thus far, as a strategic group (Grant 1995: 90–106), they have cooperated and competition within the group has been restrained with respect to payments. Evidence of this can be found in parallel pricing and the similarity of product and mechanism offerings. At this stage, the 'big-four' have the capability to offer all payment products through all mechanisms to all clients. Their dominance of the payments industry is therefore overwhelming. Where a payment product is not offered to a particular market, it is usually because it is not required in that market. While the 'big-four' achieve benefits such as economies of scale from offering services across the board, these are probably off-set by the inevitable cross-subsidisation of services to essentially unprofitable clients.

This all-things-to-all-people strategy is noticeably not followed by the other players in the industry who tend to focus on specific markets with particular products and mechanisms. It is important to note that the late entry of these players into the payments industry has enabled them to select their target products, mechanisms, and clients. They have not had the burden of trying to meet the requirements of the whole population of market segments. This freedom could present significant advantages in terms of target marketing, cost reduction, and the avoidance of cross-subsidisation. Currently, new competitors in the competitive arena (refer to Figure 1) include niche banks, foreign banks, retailers, and car dealers seeking increased volumes and values of payments. However, possible future new entrants could include insurance companies, infrastructure providers such as Eskom and Telkom, software package providers such as Microsoft, Electronic Data Interchange (EDI) providers, computer companies such as

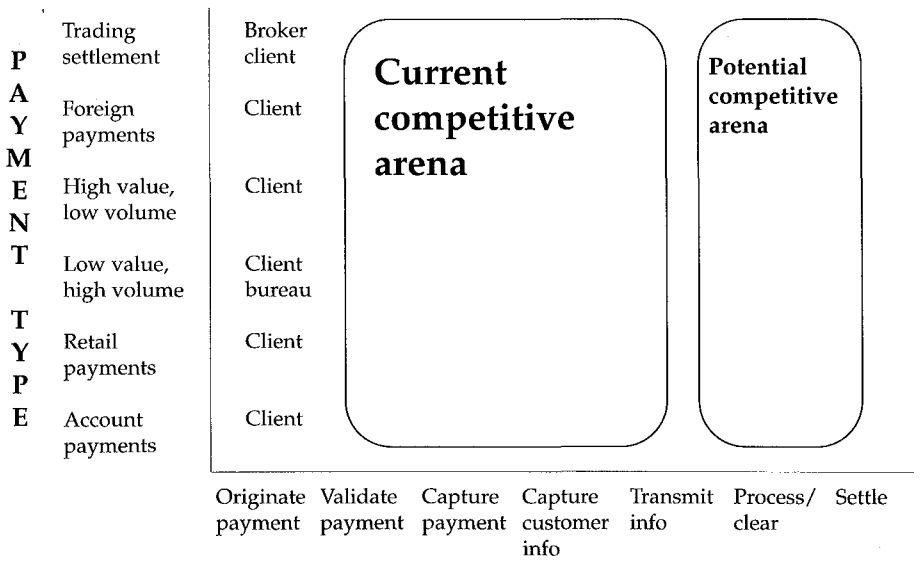

Figure 1. The current and potential competitive arenas

EDS, which provide network infrastructure for payments, and nonbank card issuers.

The niche banks (e.g. Investec, RMB, Citibank) are clearly focused on high net-worth individuals and the corporate/commercial markets. They are utilising a variety of product types but are avoiding the paper and cash products, except where paper is currently the only form of payment available such as in trading systems. Cards have been introduced (e.g. by Mercantile-Internet Banking) on a limited basis for the high net-worth individuals for access to automatic teller machines (ATMs) and retail payments. The retailers (such as Pick 'n Pay and Edgars) target the individual segment of the market. They offer a limited range of products and mechanisms to this segment, focusing particularly on card-based solutions (*Financial Mail* 1997: 35; *Saturday Star* 1997: 23). Having recognised that there is value in the volume business, many members of this group have already built up a reputation for extending credit in a convenient, friendly manner, whereas perceptions of the banks' lending policies often tend to be the opposite. (*Finance Week* 1997). Thus, control of the middle section of the value chain will give retailers easy entry into other bank strongholds (such as lending). A significant percentage of Edgars's income, for example, already derives from the interest margin charged on card lending. In the last financial year, the firm's extended credit customer base increased from 370 000 to over 440 000 (Edgars 1997: 30).

The foreign banks are very narrowly focused on the corporate market at present. Although they offer a limited range of payment products and mechanisms, these tend to be high value products through electronic mechanisms. One of their competitive advantages is that they can lend money at lower rates than local banks, because their superior credit ratings allow them to raise funds cheaply offshore (*Financial Mail* 1997: 35; Cloete 1990: 16; *Finance Week* 1997). This then often gives them a platform from which to offer payments services as well. Thus, niche banks and foreign banks have entered the payments industry at points where they are dealing with the large, well-funded corporate clients (and, in some cases, high net-worth individuals) to whom they provide excellent, flexible and personalised service. Furthermore, through the careful selection of payment mechanisms, they manage to keep their costs low (*Financial Mail* 1997). The 'big-four', with their broad-band approach and reputation for average service and inflexibility, no doubt face formidable competition in these market segments (*Financial Mail* 1997; *Finance Week* 1997; Cloete 1990: 15).

On the demand side, although clients will continue to be the originators of payments, a vital question becomes whose clients will they consider themselves to be. Previously, a client originating a payment would clearly have been a client of the bank, whereas today, the client could be a Microsoft, Telkom or Eskom client rather than a bank client. This highlights an extremely important aspect of the payments value chain. If banks are separated from their client's payment business by alternative service providers, they will no longer have a direct payment relationship with that client (see, for example, the success in the UK of Chartered Bank). This will not only be detrimental to their understanding of and relationship with that particular client, but will also impact their ability to cross-sell and to manage credit-granting using behavioural methods. Nevertheless, banks do have some notable advantages. They have been in the business of payments for a long time and they are perceived to be trustworthy by the market.

Another demand-side factor that seems to impact the industry is the increased sophistication of customers. There is little doubt that the sophisticated end of the retail market and corporate and commercial customers are looking for better payment solutions (*Financial Mail* 1997: 35–36). This means that new entrants into the competitive arena will have to add value to their payment services. Companies that can, for a competitive price, streamline the payment process through either added security or assistance in integrating data or applications with payments solutions (such as through EDI, which streamlines the communication between supplier and buyer from order to payment), will be chosen above those who do not have access to these capabilities.

From a supply point of view, there is a threat that suppliers of services to the payments industry might integrate vertically into the competitive arena. Suppliers can be segmented into two groups: suppliers of technology and suppliers of 'utilities', such as Eskom and Telkom. Whereas technology suppliers could concentrate on the provision of solutions in the corporate market or infrastructure in the retail market (Morgan 1996: 4; Deloitte Touche Tohamatsu 1995: 18; *Electronic Payments International* 1997: March), suppliers of utilities could attempt to attract retail clients by providing a mechanism to effect payments in a simple and efficient fashion. However, neither group is likely to seek to attack the clearing and settling stages of the value chain. Another potential substitute of the services provided by the 'big-four' is the purchasing of prepaid services directly from a supplier. This is already happening with Telkom phone cards, and it is possible that this type of solution could be used more extensively for other applications as it removes the risk of nonpayment and reduces administrative load significantly.

The potential competitive arena (refer to Figure 1) is the current stronghold of the 'big-four'. These stages of the value chain are presently strongly regulated and only registered banks are permitted to settle payments (Morgan 1996; South African Reserve Bank Act 90 of 1989). Barriers to entry into this area of the payment value chain are high. A prospective entrant would have to register as a bank and comply with the capital requirements and other regulations attached thereto. Although clearing would also require a significant capital outlay, this would not be as stringent as the requirements for settling. In order to be able to clear transactions, the entrant would need to have access to the existing clearing infrastructure of the 'big-four' (i.e. Bankserv switches, et cetera) or would have to participate in an alternative infrastructure (De Klerk 1997). However, the establishment of such an infrastructure and the cost of participation pose significant barriers to entry, as most clients would still require links to the clearing offered by the Bankserv infrastructure. This would call for agreements with the owners of the established infrastructure, the major banks, which would clearly have the objective of raising as many barriers as possible to protect this area of their business.

Thus, entry into the clearing and settlement stages would be an option only for companies of substantial size (e.g. insurance companies, major retailers or large service providers such as Telkom or Eskom). On the other hand, however, the possibility cannot be excluded of attempts by the SARB to 'nationalise' the clearing infrastructure given their view that the payment infrastructure is a national asset. If successful, this would provide relatively easy access to the clearing stage of the value chain.

Potential impact of the implementation of the National Payment System (NPS)

The majority of the proposals of the NPS document (South African Reserve Bank 1995) deal with issues concerning clearing and settlement (what has been termed ëthe potential competitive arenaí). However, there are some aspects that also cover the earlier stages of the value chain. These are addressed first.

The impact on the competitive arena

The NPS proposals include a strategy to formalise and provide a regulatory framework for payment service providers (South African Reserve Bank 1995: 34). Since it is currently not clear what is required to become a payment service provider, this proposal strives to ensure the integrity of the players in this market and to provide comfort to customers and make payment services more accessible (South African Reserve Bank 1995: 21; 1994: 34). The clarification of the requirements for participating in the competitive arena and the regulation of service providers are likely to result in more players entering these stages of the value chain, and to remove some of the concerns about aggressive retaliation by established players such as the major banks. In addition to the clarity provided by the proposal, the new formal position of these service providers will also allow access to the clearing and settlement stages of the payments value chain.

The impact of a significant increase in new entrants to the competitive arena on the nature of competition will be immense. As discussed earlier, such new entrants are likely to start separating banks from their clients. This will impact the banks in a number of ways that are not necessarily easily measurable. Banks will not only lose the understanding of their clients that can be achieved with complete information (which will, in turn, affect their ability to make educated credit decisions), but the loss of direct contact will also significantly reduce the banks' abilities to offer products proactively to their clients. Thus, this arena is critical to the banks to enable them to cross-sell so as to achieve a return on the significant cost of developing infrastructure. (It will not be sufficient to operate only in the clearing and settling stages.) Competition in this arena is likely to be intense with banks facing competitors who do not have to comply with the same onerous regulatory requirements.

Figure 2 identifies the factors likely to impact the nature of competition in the competitive arena, both from the point of view of incumbents and new entrants. It also indicates that the banks have good reason to stay in this part of the value chain and to protect their part of the franchise. However, they will have to understand the calibre of the competition they face and adapt appropriately. It is likely that the onslaught on this section of the value chain will either split up the strategic group of the 'big-four' or further cement their current relationship. Breaking out of the group will allow the banks to form alliances with new partners (retailers, software companies or infrastructure providers) and will assist in raising some of the barriers to entry (Grant 1995: 98; *Electronic Payments International* 1997: June). On the other hand, staying within the confines of their strategic group will tend to stifle creativity and they will run the risk of being outmanoeuvred, from either a pricing or strategy perspective, by more flexible competitors.

The impact on the potential competitive arena

The majority of the proposals of the NPS document, which comprise the core objectives for the proposed changes to payments processes, deal with the reduction of systemic risk (South African Reserve Bank 1995: 28–33). They introduce a number of extremely significant changes to the clearing and settlement aspects of the payments industry.

Firstly, there is currently a perception that the SARB will provide support to any bank in the settlement system that cannot meet its liabilities in order to ensure the stability of the monetary system in South Africa. Under the new NPS, the

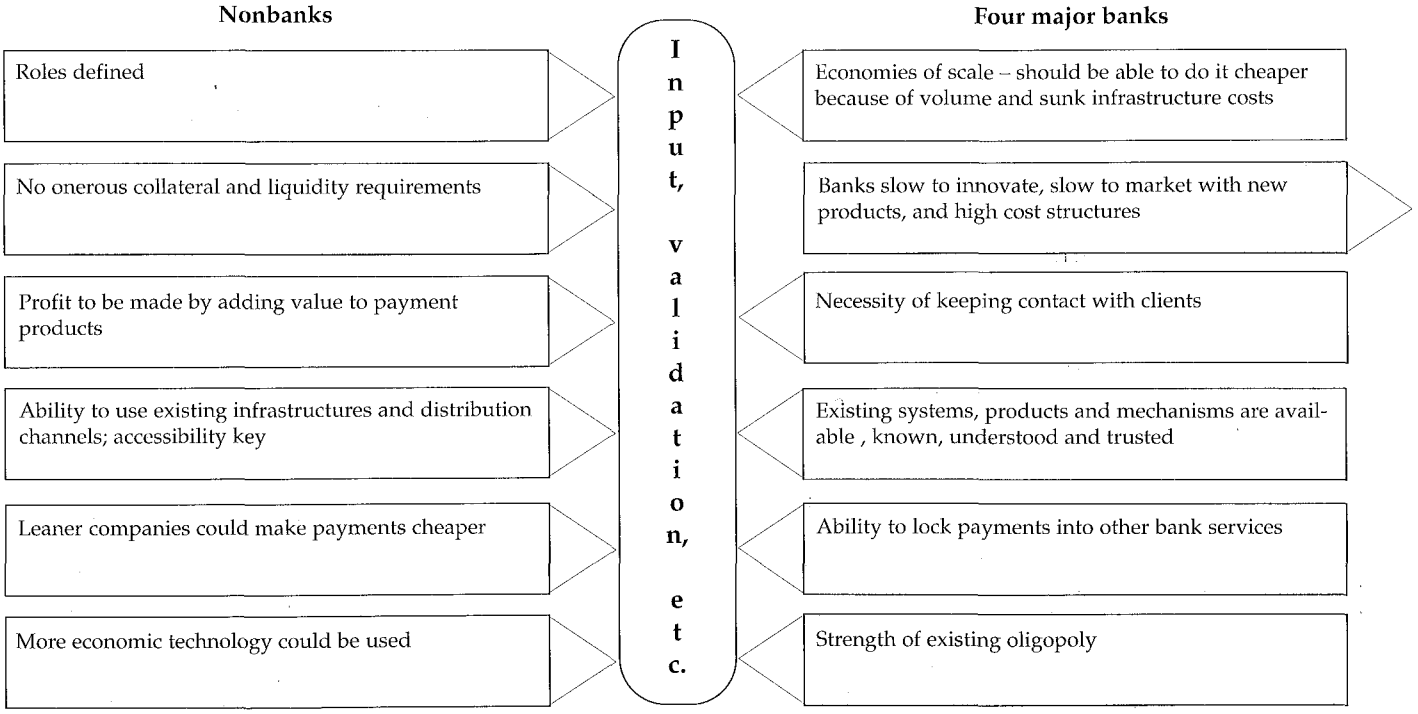

Figure 2. Push and pull factors: The 'big-four' and nonbanks in the competitive arena of the payments value chain

payment system will no longer be underwritten and supported by the SARB. Although each clearing and settling bank will become responsible for its failure to meet debt, the new NPS strategy is designed to reduce the risk of failure of a settlement bank and, should such failure occur, trap the situation at the point of failure rather than allowing a bank to continue accruing liabilities (Bank for International Settlements 1997: 1). This categorical statement on SARB nonintervention is likely to see a reduction in any potential risk of 'moral hazard' (Milgrom & Roberts 1992: 166).

Under the new system, the SARB will monitor the banks' positions, but the onus will be on the banks themselves to understand and manage their own exposures. This will require the development of systems to monitor and predict payment flows from all areas. Although a costly exercise, banks will benefit from being able to understand their payment flows at a far more granular level than is currently the case. The expense of this operation, however, will be particularly felt by the established banks which have a multitude of legacy payment mechanisms, most of which are based on an overnight processing methodology rather than the real-time concept prevalent in the NPS strategies. The likely shortage of liquidity and the necessity of managing cash flow will possibly increase rivalry. Short-term advantages will possibly exist for banks which can develop applications to provide central liquidity information and cash flow modelling capabilities, and which have an ongoing ability to stay liquid in order to make payments. A bank that comes to a standstill because of illiquidity will not be looked on favourably by clients whose own creditworthiness might be put at risk.

Secondly, the SARB will develop a computer application to support interbank settlement, which will be the only mechanism to effect settlement. Each bank participating in the settlement system and every payment clearing house will be required to develop an interface with the system. This will allow the banks to have instantaneous access to their positions with the SARB and immediate notification of transactions that affect those positions.

Thirdly, the NPS strategies point to a split of low-value, high-volume payments and high-value, low-volume transactions. This is justified by the higher risk profiles of high-value transactions and follows trends elsewhere in the world (Quinn 1993; *Electronic Payments International* 1996: May; *Payment Systems Worldwide* 1997: 34). A high-value payment system would require real-time settlement for risk to be reduced significantly. It would be generated by a client, presented by the client's bank, settled at the SARB, and then updated in the other bank's account immediately. Such a system does not as yet exist and requires significant application development by the banks and the SARB. The removal of the high-value payments from bulk (netting) clearing should significantly reduce risk incurred by receiving banks, if this were totally enforceable.

A fourth aspect covered in the NPS proposals concerns the rules governing netting (that is, when banks settle a net amount rather than on a transaction-by-transaction basis). The Bankserv infrastructure currently nets transactions wherever this is possible. Thus, the settlement of banks' accounts at the SARB the day after value date (payments made today will currently only be settled tomorrow at the SARB), is already a netted position. The policy of the SARB will be to allow netting to continue, but only where agreed risk reduction measures have been taken. These could take the form of loss-sharing agreements, whereby the banks agree to share the loss if one party to the netting agreement is unable to meet its debt. While certain foreign payment systems operate on this basis (Balino, Johnson & Sundararajan 1996: 3), it is unlikely that any of the 'big-four' would enter into such agreements with smaller players and new entrants without established credibility.

Fifthly, the introduction of an RTGS system, combined with the SARB's reluctance to underwrite any losses, will impose a significant intraday liquidity management requirement on all banks participating in the settlement system. The SARB will require a bank to have a prefunded settlement account at the SARB, based on the bank's balance as at the end of the previous day and any transactions that have occurred during the day. In the event that a bank requires further liquidity to fund payments that it wishes to make, the SARB will grant overdraft facilities but only against pledged liquid assets. If all of these facilities have been utilised, a bank will not be able to pass payments until transactions are received from other banks (these could be in the form of loans), or unless further arrangements have been made with the SARB (e.g. borrowing using other financial assets as collateral). This process signals a significant change for two reasons. Firstly, the vast majority of transactions are currently netted, thus avoiding the criticality of the timing of debit and credit transactions in an RTGS system. Secondly, final settlement is currently delayed until a number of hours after the banks become aware of their positions. This allows them to determine the cheapest source of funds, should they be short of liquidity.

The requirement to collateralise the borrowing from the SARB could potentially involve significant cost to the banks, especially if they are forced to acquire liquid assets over and above existing requirements (Singer 1996: 3; Trumble 1995: 18). Thus, the management of positions and the timing of passing transactions for settlement will become critical to reduce as much as possible the costs of participation. Also, the banks will need a real-time capability to track and manage the liquid assets that they hold. However, this requirement for collateralisation is likely to have an even greater impact on smaller banks, niche banks, and some foreign banks. If any of these banks have settlement accounts at the SARB, they will have to comply with the liquidity requirements for payments. This could be onerous on these banks whose asset bases are limited and who do not have the ability to raise liabilities easily (the payment resulting from taking on the liability will provide a positive liquidity flow). Foreign banks, permitted only to attract deposits over R1 million, may also find it difficult to attract liabilities quickly if required. These banks, with less reliance on legacy applications than the 'big-four', would therefore seem to be prime candidates to initiate an intraday money market. The implementation of an RTGS system could also be a factor that increases competition in this arena. This is because banks that are only able to meet their liquidity requirements at end of day will not be able to participate in this market. On the other hand, the onerous requirements to be in clearing and settling will make this a relatively elite group. Is the SARB protecting this arena for the major banks or is it attempting to level the competitive environment by forcing the large banks to participate in an unattractive section of the value chain?

Another impact of the proposed NPS on the potential competitive arena concerns the move of the final settlement of a day's transactions from the next morning to the same day. This

will impact every payment stream, even the netted (bulk) low-value transactions. Recent trends have been to allow clients to initiate transactions for same day value later and later in the day. This has been made possible by improvements in technology and improved communications capabilities between the banks and the various Bankserv services. The change to same day settlement will bring forward the deadlines for same day value for certain transactions. There will be two major impacts as a result of this strategy. Firstly, the banks will start to charge variable amounts for transactions depending on when they are initiated during the course of the day. This implies a significant change in the pricing regime of the industry, as branch deposits, other than cash, are currently free. It is likely that this will change. Secondly, there will be an enormous impact on the banks whose processing methodology is typically batch (overnight) rather than real-time. The impact of changing this will be massive as it will require significant enhancement of the current systems of the banks, with large cost and timing implications.

The SARB's target of March 1999 (South African Reserve Bank 1997) to achieve same day settlement would appear to be extremely optimistic. In order to achieve a same day clearing capability which would feed the settlement system, the major players will have to replace significant infrastructure or, at the very least, develop complex methods of working around the current application limitations. With the huge amount of work required to enhance applications for the approaching millennium change, it is unlikely that any of the major banks will have the necessary resources to create the capability for same day settlement by March 1999. A likely scenario would be a phased daily settlement process, with high-value transactions being closed in the RTGS system on a same day basis and bulk processing continuing to be settled overnight (effectively next day). However, this would have to be managed extremely carefully to avoid the situation where different payment products have different 'end-of-day' times. Such a situation would create difficulties in terms of value dating and could lead to participants attempting to arbitrage between payment types.

A further issue relates to the ending of financial sponsorship, as indicated by the SARB in the NPS proposals. Presently, there are numerous smaller banks which issue payments in their own name using their own infrastructure, but who are sponsored in the clearing and settlement systems by a larger bank. Thus, in the clearing and settlement systems, these transactions are processed as though they were generated by the sponsoring bank and the final settlement thereof will pass across the account of the sponsoring bank at the SARB. This has the effect that a sponsoring bank is not always fully aware of the risk generated by the transactions that it is required to settle.

The impact of this ruling will be twofold. Certain organisations that are currently sponsored will decide to maintain their current ability to offer an end-to-end payment service to their clients and submit to the regulations for clearing and settling in their own name (Stegman 1997). Others, however, will examine the potential cost of collateralisation and their inability – artificially regulated or otherwise – to raise liabilities swiftly, and decide to make use of the client or bureau facilities of a larger bank. The number of organisations likely to follow these different routes will depend on two issues, namely the increase in liquid asset requirements on becoming a participant in the settlement process, and the perception among these organisations of the availability of relatively cheap alternatives (compared to liquid assets) to collateralisation. A number of organisations that are currently sponsored will enter the arena, probably concentrating on electronic payments. There will be the segment of banks that do not wish to clear and settle in their own names. They could be lucrative targets as corporate clients for the major banks that understand the cost of the risk that they would be taking on. There will also be opportunities for banks which decide to settle and clear in their own name to act as bureaus. One such alternative would be the establishment of an intraday money market. The possibility of an intraday money market cannot be underestimated and this would certainly add a dimension to competition between banks.

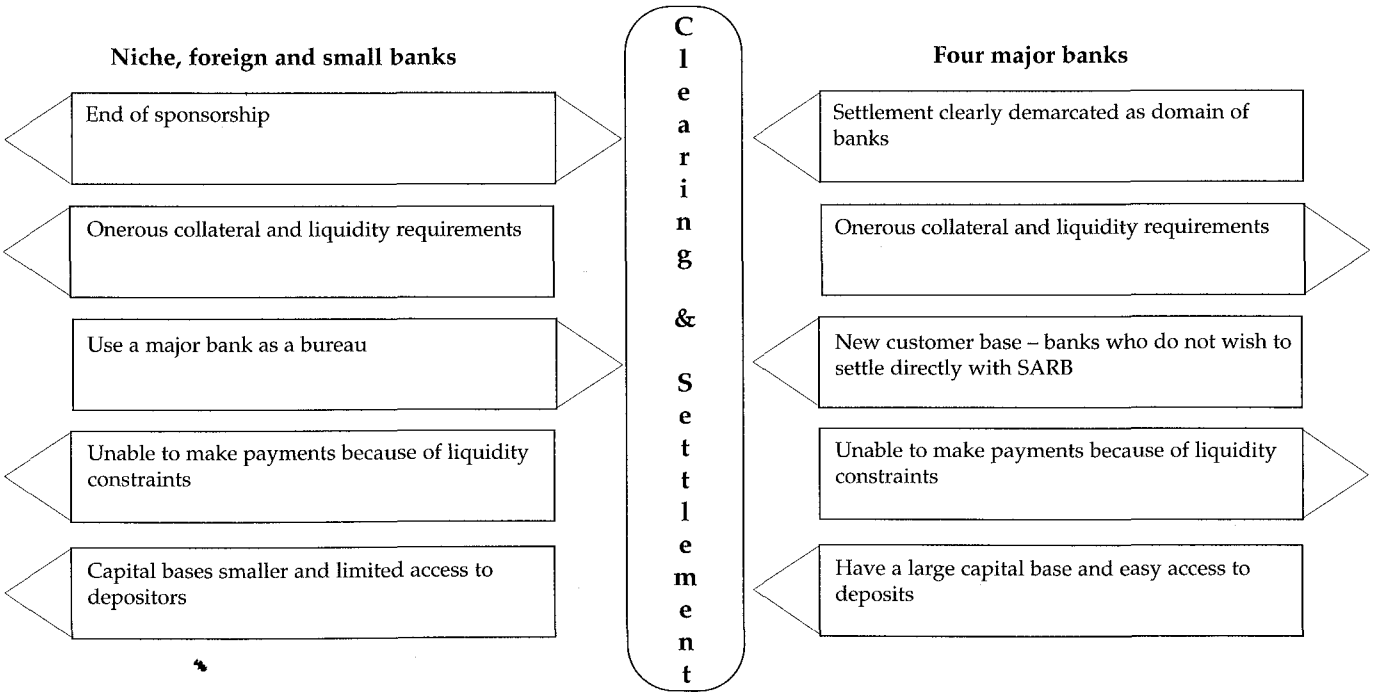

Figure 3. Push and pull factors: The 'big-four' and possible new entrants in the potential competitive arena of the payments value chain

Rivalry around intraday liquidity could also impact on players in the industry. This is because cash rich organisations such as insurance companies and pension funds might well be able to command higher interest rates in such a market. Alternatively, this might be restricted to the existing interbank loans market.

Finally, as previously mentioned, the major banks have spent significant time and money collaborating in the development of the various infrastructures that now make up Bankserv. Possibly one of the most contentious provisions of the SARBís NPS proposals is the statement that the existing infrastructure of Bankserv will be allowed to continue operations as long as it is deemed to be part of the NPS and its facilities are accessible to all players on an equal basis. It is not hard to see that the 'big-four', who have invested millions in the infrastructure and in the development and maintenance of their interfacing in-house applications and their own branded ATMs, could take offence to this.

The outcome of this debate will have a significant effect on the structure and nature of competition in the industry, whatever the outcome. Should the 'big-four' resist the direction implied by the SARB, they will be able to use access to the infrastructure as a barrier to entry or, at the very least, run the infrastructure at a profit (which is currently against the constitution of Bankserv). At the same time, the possibility of a new competing infrastructure could become an option (for example one that allows the clearing and settlement of electronic transactions, as paper clearing and credit card and ATM transactions are more complex). Such an infrastructure would compete with elements of the Bankserv offerings. Alternatively, competing banks could outsource their infrastructure requirements to a network company. Should the outcome be that all parties are to have equal access, then the competitive advantage of owning this network and ATMs disappears. The banks will then have to find alternative ways to differentiate themselves.

Figure 3 summarises the impacts that the changes envisaged could have on the various industry players in respect of the clearing and settling aspects of the value chain. The figure highlights the aspects that will inform organisations' decisions whether to enter the potential competitive arena. Arrows pointing towards the middle of the diagram indicate that the factor will induce an organisation to participate in the arena; arrows pointing away from the centre indicate that the factor will discourage such participation. It is likely that there will be new entrants, but they will only be those who believe that they can meet the collateral and liquidity requirements on an intraday basis. If the NPS did not impose these new requirements, it is likely that there would have been more entrants and thus it could be concluded that the NPS has actually increased the barriers to entry and movement within the industry. With respect to the 'big-four', it is extremely unlikely that they will exit this arena, despite the negative factors associated with liquidity and collateral management. They are likely to view participation as part of their core business and will thus find ways to circumvent the negative impact of the NPS proposals, rather than withdraw.

Research methodology

The aim of the empirical part of the study was to gauge the potential impact of the NPS proposals on the competitive and potential competitive arenas of the payments value chain. To this end, the proposed impacts, as discussed, and summarised in Figures 2 and 3, serve as the study's implicit propositions. The research was undertaken on the basis of 16 in-depth, structured interviews. Two test interviews were conducted using an interview question-schedule based on the work above. Feedback from these interviews was incorporated in the final question-schedule. The questions addressed five main areas: the forces impacting the banking industry in South Africa; the importance of payments to the banking industry; the possible impact of the NPS on the structure of the industry; the possible effect of the NPS on the nature of competition; and the possible responses of the 'big-four' to these developments.

Table 1. Sector representation in interview sample

Sector	Number
'Big-four' banks	7
Niche banks	1
Foreign banks	2
Bankserv	2
Retailer	1
SARB	3
Total	16

Table 1 shows the composition of the group of interviewees. The choice of respondents was driven by two considerations. Firstly, it was important to interview people who had a thorough understanding of the NPS. Thus, many of those interviewed had participated in the working groups set up by the SARB and the banking industry to address the various areas covered by the NPS. However, it was also considered necessary to canvas opinions from the niche banks, foreign banks and retailers. Care was taken to ensure that the interviewees selected occupied positions where they would at least have been exposed to the literature emanating from the NPS.

Secondly, the selection was driven by the requirement to obtain a reasonable cross-section of opinion. Given that this study is focusing particularly on the impact of the NPS from the perspective of the existing clearing banks, more people from these institutions were interviewed. The SARB has played an important role in setting up the NPS and, for this reason, a number of central bank representatives were approached. The clearing banks currently own the clearing infrastructure of the payment system. It was therefore deemed important to gain insight from the people who run this infrastructure (Bankserv) on behalf of the banks. They would have been party to many of the discussions and would possibly have an independent view of the importance of payments to the banking industry. Certain foreign banks, niche banks and retailers with a particular interest in payments were also approached for interviews.

Results

The perceived impact of the NPS proposals on the payments value chain is summarised in Figure 4. The competitive arena is likely to see many new entrants, formalised as Customer Payment Service Providers (CPSPs), which could potentially increase the number of people who have access to payment mechanisms other than cash. This would mean that persons

previously only using cash for payments might move to other forms of payment such as debit card, et cetera. Although these entrants would be regulated, they will be able to deal directly with traditional bank customers, thus creating a barrier between these banks and their clients. The potential competitive arena was still seen as only the domain of banks. This would mean that potential new entrants into this arena would be restricted to the foreign and niche banks. Furthermore, the already high barrier to entry into settling would be further raised because of the intraday collateral requirements for payments.

Based on the responses obtained from interviewees, some of the proposed push and pull factors for the 'big-four' and their possible competitors in the competitive and potential competitive arenas (refer to Figures 2 and 3) will now be revisited. The regulations to govern CPSPs were viewed by 62 per cent of the interviewees as a positive step towards formalising and regulating these service providers. Interviewees concurred that organisations acting in this capacity need to be responsible and trustworthy, as they handle clients' money. Organisations viewed as having a role to play in this arena included retailers at all levels, including nonformal organisations such as spaza shops.

A factor highlighted by 31 per cent of the interviewees was the necessity for constant innovation in order to bring new payment products to market. Some of those interviewed felt that this would be where the more traditional banks might not be able to perform adequately and that the niche banks, foreign banks and retailers might be better able to devise the necessary innovative payment products. This would require a particular type of management skill not normally associated with the usual conservative 'banking' manager. However, none of the 'big-four' respondents believed this to be of importance.

Apart from the ability to innovate, two other factors were thought to become drivers of competition in the competitive arena of the payments value chain. The isolation of the apartheid era protected the local banking industry from any outside competition. Costs were not contained and banks became 'fat'. The inability of South African banks to contain costs could mean that they would find it difficult to respond to competition from foreign banks which, after years of international competition, were 'lean' and extremely cost competitive. Thus, 56 per cent of the respondents regarded globalisation as one of the main forces for change in the payments industry.

A second important force for change and driver of competition, identified by 37 per cent of the respondents, was technology. This included the ability to process cheques electronically, document imaging, smart card technology, Internet banking, and any other technologies which could transform banking or the payments process. Interviewees believed that technological advances would enable newer players to enter the industry without legacy systems 'baggage'. Many expressed the view that organisations would have to align themselves with technological advances in order to remain players in the payments industry.

At present, with the infrastructure for clearing being owned by the 'big-four', interviewees were asked whether this was considered important to the continued existence of the banks. Only 18 per cent of the interviewees answered in the affirmative, with 62 per cent disagreeing. Seventy-five per cent of those interviewed believed that it was highly likely that other organisations that specialised in this type of infrastructure, such as communication and computer specialists would encroach on this area. The possibility of outsourcing the development and maintenance of such an infrastructure to companies such as EDS was also mentioned.

It was argued by 31 per cent of the interviewees, noticeably those from the 'big-four', that control of the infrastructure was far more important than ownership thereof. On the other hand, the SARB employees interviewed believed that neither control nor ownership was necessary (thus aligning with the

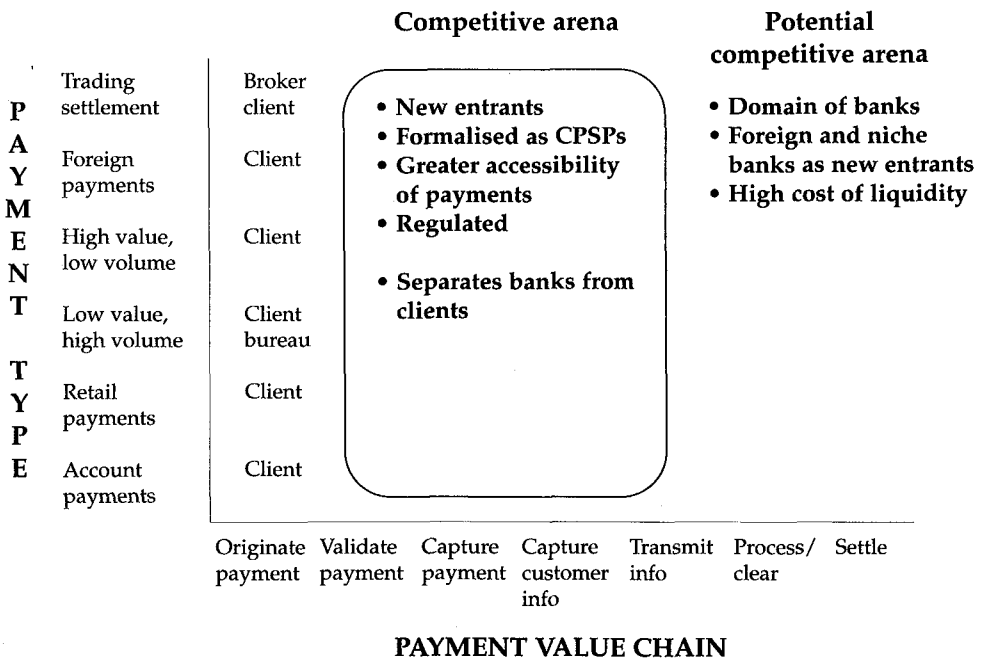

Figure 4. The impact of the NPS on the payment value chain

NPS strategy document that states that Bankserv should be declared a national asset). Thirty-one per cent of respondents said that losing control of clearing was a major threat and 37 per cent replied that they would consider loss of control over clearing as serious. Although it was recognised that there were possibly other organisations that could perform clearing more effectively, these responses show a significant concern on the part of the 'big-four' regarding the issue of control over the clearing stage of the value chain. Owning the infrastructure could be regarded as a strategy to maintain such control. On the other hand, the niche banks, foreign banks and SARB interviewees appeared unconcerned about infrastructure control, reflecting their perceptions of the difficulties associated with gaining access to the infrastructure. Thus, although the clearing stage of the value chain might remain the domain of the 'big-four', it is not certain that the infrastructure will remain under their control.

With regard to increased competition in the potential competitive arena, a number of interviewees thought that becoming a clearing and settling bank could be extremely burdensome because of the intraday liquidity requirements, which could be viewed as a new barrier to entry. Twenty-five per cent of interviewees believed that niche banks would experience liquidity problems, 31 per cent believed that foreign banks would experience liquidity problems, while only 12 per cent thought that the 'big-four' would experience problems. The position of the foreign banks was that the requirement would be particularly burdensome on them because of the restriction on their deposit-taking ability and the fact that they deal in high-value payments. They would therefore require high intraday liquidity that they could not fund through liabilities. Nor would they want to procure liquid assets simply for the purpose of funding an intraday position. Twenty-five per cent of the interviewees argued that it would be difficult for the smaller players to manage their capital exposures.

The necessity for intraday liquidity has therefore added a significant barrier to entry to the clearing and settling stages of the payments value chain. High-value payments will not be able to be made unless sufficient collateral is held. Although 31 per cent of the respondents (including SARB staff) thought that the playing fields would be levelled through the implementation of NPS, the foreign and niche banks did not agree and argued that the scales were tipped even further in favour of the 'big-four'.

Given the liquidity requirements of the new NPS, 31 per cent of the interviewees believed that there would be insufficient liquidity in the payments systems, while 25 per cent thought that the problem could be so severe that payment gridlock could occur (which means that payment processing may come to a standstill). On the other hand, 25 per cent of the respondents argued that there would be no system-wide liquidity problem. SARB officials, in particular, pointed to the fact that they would 'grease' the system every day by making their payments early on. It was also argued by some respondents (18 per cent) that in order to circumvent the intraday liquidity requirements, there would be a reversion to cheque payments. Since these would only be processed overnight or at end-of-day, banks' intraday positions would not be affected. This, however, would defeat one of the purposes of the NPS, namely to reach a quicker settlement for high-value payments.

While liquidity and other infrastructural requirements would seem to protect the domain of the 'big-four' in the clearing and settlement stages of the payments value chain, interviewees were not confident that this position would be sustainable over the longer term. In fact, only 37 per cent and 50 per cent of the respondents argued that clearing and settling respectively were the strict domains of banks. This is an indication that there is a concern that these activities might not be restricted by law to banks indefinitely. If this was to change, banks could be further excluded or their participation minimised in the payments process, or forced to participate only in niche positions where they could add value (e.g. large value, foreign and trading payments). In this regard, some interviewees referred to the Wallis Report (*Electronic Payments International* 1997: April) that has recommended that nonbanks be allowed to hold settlement accounts at the Australian Central Bank.

However, in the mean time, the regulatory aspect of payments would still restrict access to clearing and settling. This was confirmed by the interviewees who argued that very few retailers, for example, would register as banks for the purposes of clearing and settling. Nevertheless, as shown above, there was a view that the formalisation of the position of CPSPs would provide interested organisations with a framework to enter this arena.

Finally, with regard to the ending of financial sponsorship (one of the NPS strategies), 18 per cent of the interviewees felt that it would be better for affected organisations to establish individual settlement accounts at the SARB. However, 56 per cent of the interviewees felt that sponsorship in its new form (i.e. as a corporate customer of a bank that has registered as a clearing and settling bank with the Reserve Bank) would be acceptable as long as the price was right. The foreign banks suggested that the 'big-four' would have difficulty in arriving at the right price as they never had to price the risk element of clearing and settling on someone elseís behalf. This new type of sponsorship would mean that the sponsored bank could not clear and settle in its own name and thus that the sponsoring bank would take on all the risks associated with those payments (including the provision of the necessary intraday liquidity and collateral). These additional risks and liquidity requirements would have to be included in the price charged for such a service.

NPS implications

The results from this study seem to suggest that the implementation of the NPS is unlikely to cause a dramatic change in either industry structure or the nature of competition. However, it is certain to become a force for incremental change in the payments industry towards increased competition and a shift in the make-up of organisations offering payments services. The implications for the 'big-four' banks are significant. As a SARB interviewee suggested, the NPS will force these banks to look more closely at their position in the payments value chain and to acknowledge and respond to the many threats facing them. The following section assesses the post-NPS opportunities and threats facing the 'big-four', based on the results of the study.

Opportunities for the 'big-four'

As a result of the implementation of the NPS, the major banks are faced with a number of avenues worth exploring. The first is the management of payments on an intraday basis and the monitoring of associated risks. This will give the banks a new

understanding of their risk profiles resulting from payment transactions and will enable them to manage real-time flows. Banks that accomplish this will be able to ensure that they are never in a gridlock position, that they can always meet their settlement requirements, and that their cash flow is maximised. A bank will have an advantage if it is able to manage intraday payment flows in such a way as to maximise liquidity and reduce collateral requirements to a minimum.

Another opportunity offered by the NPS will be to take on smaller banks, perhaps foreign or niche banks, as corporate customers. Because sponsorship will no longer be permitted, the smaller banks will either have to register with the Reserve Bank as a clearing and settling bank in order to participate in the settlement process (and, of course, develop the necessary infrastructure) or operate as a corporate customer of another bank. If the 'big-four' are able to offer appropriately priced and flexibly structured product sets and have a good understanding of the risks associated with these products, this has the potential to develop into an extremely profitable business.

Furthermore, banks have the opportunity to develop, proactively, strategic alliances with organisations such as retailers, communication companies and utility companies, which are able to add value in the competitive arena of the value chain. This could become particularly important, since these organisations pose a real threat to the banks in that they could separate the banks from their clients in the important area of payment generation and processing. If the banks enter into these strategic alliances, they would still have access to their client base, more distribution points would be available, economies of scale and synergy could be achieved, and they would still be able to offer expertise in credit vetting, understanding risk, clearing, and settling. Further benefits would be a reduction in branch network costs and the ability to offer banking services to previously 'unbanked' people.

Given the strong possibility that an intraday money market will emerge as a result of the NPS, the 'big-four' could have a huge advantage if they are able to develop applications to charge interest on an intraday basis. As they currently possess the bulk of corporate funding and have relatively large liquid asset stocks, they are unlikely to be borrowers in an intraday market. With the necessary application support, the ëbig-fourí will be able to lend to the smaller players in the settlement system when the latter are short of liquidity. This, however, will require excellent control over the intraday flows of corporate funds to ensure that their clients are not providing intraday liquidity (as it is possible that certain corporate companies with excess cash on an intraday basis may see the opportunities of lending to banks experiencing a cash shortage), and a re-evaluation of their contracts with big corporate customers.

The four major banks currently control the Bankserv group of companies which, in turn, owns the infrastructure used for the clearing and switching of most payment transactions in South Africa. As new organisations decide to clear and settle, it will be important for the 'big-four' to allow access to these new players at a fair price. If they are excluded totally, the new entrants will be forced to find alternative ways of clearing. This could be through an extension of the SAMEX system (a system which the Reserve Bank has made available to all banks for the capturing of interbank settlement instructions by employing a rival company to set up alternative clearing facilities). Neither of these options would bode well for the major banks. It would be far better for them to maintain control of access and run Bankserv at a profit. Should alternative clearing infrastructures be created which undercut Bankserv, even the 'big-four' might prefer to clear more cheaply. This would have the effect of breaking up their 'strategic alliance'.

The 'big-four' also need to consider the interbank lending agreements that they will enter into with each other and with the smaller players. The smaller players may seek to play the larger banks off against each other in these agreements. This will not be possible if the four major banks adopt a common stance in respect of these agreements. Otherwise, this could open up new intra-industry, even segment, rivalry for this business.

The implementation of the NPS will cause liquidity problems for foreign banks and niche banks. This will be an ideal opportunity for the major banks to offer a quick and efficient service that will be more attractive to the smaller competitors. This could even lead to a return of some of the business recently lost to the foreign and niche banks.

An opportunity that should not be missed by the major banks arises from the fact that they are still trusted by many clients. Thus, there will be many people who will prefer to pass their payments, especially high-value payments, through the banks. The banks will be able to capitalise on their perceived trustworthiness and market this aspect to clients. Application security, the banks' sizes, and the duration of their participation in the market will all contribute to a climate of security.

There will be advantage to be gained by being able to offer an RTGS capability to clients by October 1998 when the SARB will implement the RTGS process for corporate clients. There is definitely a market for RTGS where clients wish to be assured of the irrevocability and immediacy of the payment. This will, therefore, be a new product that can be sold and priced appropriately. The first bank to offer it to clients will have a first-mover advantage.

Threats facing the 'big four'

A significant threat to the four major banks is that the SARB will no longer provide a safety net to a bank that cannot meet its obligations. This will put the onus on the banks to monitor their own risk profiles competently and to ensure that they do not overextend themselves.

Another threat faced by the four major banks is that the retailer, insurance, computer, communication and utility companies, in their newly defined roles as CPSPs, could drive a wedge between the banks and their clients. The competitive arena has been identified as the most lucrative section of the payment value chain, and the banks should do everything in their power to maintain a commanding presence there. A possible way to achieve this would be the marketing of innovative payment products through a series of new and existing delivery channels. The banks will have to compete with the low price, process efficiency, and ease of access of the products that other players will provide in this arena.

Another threat that the major banks will face is that of increased competition. Clients will be faced with a broader choice of organisations offering to process their payments. Banks that do not provide excellent service, charge competitive interest rates, and market up-to-date products will not be chosen. This would particularly be the case in the corporate market.

Complacency is also a threat to the four major banks. They could argue that the NPS did not signal any material change to their business and therefore do nothing to address the threats to the profitable business in the competitive arena, nor take advantage of any of the opportunities presented by the NPS. If they were to adopt this attitude, they would lose a significant share of the payments market to those organisations making use of the opportunities presented by the NPS and other forces in the industry.

Conclusion

The implementation of the NPS will not revolutionise the way that payments will be made. In fact, to the man in the street there will probably be no perceived change. However, the impact of the NPS will be far-reaching and important. Apart from the obvious systemic risk reduction and the fact that banks will have to monitor their own liquidity positions more carefully in order to fund their payments, it has focused attention on payments and the threat which the major banks face in this arena. Payments have always been so closely aligned to banks that it could possibly have been considered as only a banking function. However, new players such as smaller niche banks, trading houses, retailers, and foreign banks are poised to enter this arena if they are not already in it. The threat is that these new players will be able to offer payments either more accessibly (in the retail market) or more cheaply and with more sophisticated electronic systems (such as the foreign banks) than the 'big-four'. The major banks will have to stand their ground and provide value-added services to clients so that they are not left with only the clearing and settling part of the value chain. In the corporate market, control of the big corporate customer payments will be of key importance and these clients are already being tempted by the service and systems of the foreign banks and niche banks. In the retail market, accessibility and reliability of payments will see retailers and possibly the utility companies able to offer payment services. The threat in this market could even extend from payments to the issuing of credit.

One of the interesting results will be to see if the onslaught on the payments business by new players and the possible entry of other companies providing infrastructure and clearing will break the strategic group of the 'big-four' or if it will, in fact, strengthen this alliance. If the major banks are not able to secure a definite place for themselves in the earlier parts of the payment value chain, they will find themselves losing their relationship with key clients and will be in danger of losing control of the payment franchise. The NPS has sounded the alarm. If the 'big-four' wish to be in the payment game, they need to provide payments to all their market sectors at the right price, with value-added services and ease of access.

References

Balino, J. T., Johnson, O. E. G. & Sundararajan, V. 1996. 'Payment System Reforms and Monetary Policy', *Finance and Development*, March: 2–5.

Bank for International Settlements. 1997. Real time gross settlement systems. Basle, Switzerland.

The Boston Consulting Group. 1996. The wholesale franchise PSIC study. Boston: The Boston Consulting Group.

Cloete, J. J. 1990. 'The SA Banking Scene', *The South African Banker*, February: 15–17.

De Klerk, E. 1997. Personal interview. August.

Deloitte Touche Tohmatsu International. 1995. The future of retail banking. London: Deloitte Touche Tohmatsu International.

Edgars. Annual Report, 1997: 30.

Electronic Payments International. 1996. Lafferty Publications. May: Issue 108.

Electronic Payments International. 1997. Lafferty Publications. March: Issue 117.

Electronic Payments International. 1997. Lafferty Publications. April: Issue 118.

Electronic Payments International. 1997. Lafferty Publications. June: Issue 120.

Financial Mail. 1997. 'Red Carpets for the Rich', October, 10.

Finance Week. 1997. 'Retailer Take on the Banks', September, 18.

Finance Week. 1997. 'Racing ahead of its Rivals', November, 13.

Grant, R. M. 1995. *Contemporary strategy analysis*. Cambridge, MA.: Blackwell Business.

Melody, W. H. 1996. 'The Role of Payment Systems in an Information Society', *Proceedings of the 14th Payment Systems International Conference*. Switzerland: 23–25 October.

Milgrom, Paul & Roberts, John. 1992. *Economics, organization and management*. Englewood Cliffs, NJ.: Prentice-Hall.

Morgan, D. R. 1996. 'Blueprint For Wholesale Payments', *Proceedings of the 14th Payment Systems International Conference*. Switzerland: 23–25 October.

Payment Systems Worldwide. 1997. Autumn.

Porter, M. E. 1985. *Competitive advantage*. New York: Free Press.

Quinn, B. 1993. 'The UK Approach to controlling Risk in large-value Payment System', *Bank of England Quarterly Bulletin*, November: 530–535.

Rozwadowski, J. 1996. 'Competing in a World of Change – European Perspective', *Proceedings of the 14th Payment Systems International Conference*. Switzerland: 23–25 October.

Saturday Star. 1997. 'Retail Financial Services simply aim to Please', November, 29.

Singer, G. K. L. 1996. 'Key Drivers in Large Value/Wholesale Payment Systems', *Proceedings of the 14th Payment Systems International Conference*. Switzerland: 23–25 October.

South African Reserve Bank. 1995. South African National Payment System Framework and Strategy. Johannesburg: South African Reserve Bank.

South African Reserve Bank. 1997. SAMOS Implementation Schedule.

Stegman, N. 1997. Personal interview. August.

Trumble, P. G. 1995. 'The Feasibility of a National Payment System for South Africa', *Unpublished M Com Research Report*, University of the Witwatersrand.

Vessey, P. 1996. 'Key Factors influencing Retail Payment Systems', *Proceedings of the 14th Payment Systems International Conference*. Switzerland. 23–25 October.

Volschenk, C. 1997. 'SA Attracts more new Banks, says Registrar', *The Star Business Report*, December, 10.