PROCEEDINGS OF THE
PLYMOUTH BUSINESS SCHOOL AND SCHOOL OF
SOCIOLOGY, POLITICS AND LAW

POSTGRADUATE SYMPOSIUM 2006

EDITED BY

Adrian Barton and Jonathan Lean
Contents

Page

1  INTRODUCTION
    Adrian Barton & Jonathan Lean

5  A MODEL OF STRATEGIC HUMAN RESOURCE DEVELOPMENT
    Zhuoran Yang

18 MODELLING THE EDUCATION AND SERVICE ACTIVITIES OF JUNIOR DOCTORS UNDER THE EU WORKING TIME DIRECTIVE
    Sonja Derrick

35 INTERNATIONAL ENTREPRENEURSHIP: CRITICAL FACTORS FOR SUCCESS
    Francisco Serra

54 PROFESSIONAL FOOTBALL SPONSORSHIP - THE RESEARCH PROJECT
    André W. Bühler

75 PROSPECTS FOR THE DEVELOPMENT OF INTERNATIONAL LOGISTICS IN NORTH KOREA
    Kelvin Pang

93 INVESTIGATING ACCIDENTS IN THE MARITIME DOMAIN - THE INFLUENCE OF CREW INTERACTION AND COMMUNICATION AMONG MIXED NATIONALITY CREW ON SAFETY
    Robyn Pyne

103 CREDIT RISK IN EGYPTIAN BANKS
    Hussein Abdou

124 THE OPTIMAL CONTROL THEORY MODEL FOR OPTIMAL DIVIDEND DETERMINATION
    Rumbi Mukonoweshuro

158 FINANCE FOR SHIPPING: PAST FINANCING METHODS
    Jonathan Challacombe

167 KNOWLEDGE TRANSFER BETWEEN EUROPE AND ASIA: IS TURKEY A BRIDGE BETWEEN NETWORKS?
    Dababrata N. Chowdhury

182 THE STRATEGIC MOTIVES FOR INTERNATIONAL JOINT VENTURES AND NETWORKS: A STUDY OF SINO-TAIWANESE JOINT VENTURES
    Chen-Yu, James, Liu & Lynne Butel

197 PARTNERSHIP WORKING - ISSUES OF INFORMATION FLOW
    Greta Squire

205 MCPHERSON ‘One step black’ DIVERSITY, DISCRIMINATION AND PREJUDICE WITHIN THE POLICE SERVICE
Mark Harron

214 A GENDER AWARE ANALYSIS OF PEACEBUILDING - THEORY, POLICY AND PRACTICE.
Annette Elliott

225 A REALISTIC APPROACH IN EVALUATING URBAN REGENERATION PROGRAMMES
Osbourne Glover

235 EVIDENCE IN SOCIAL POLICY: CAN WE SEE PAST THE HIERARCHY OF EVIDENCE?
Mark Pearson

245 MAKING WHOOPPEE: WOMEN AND ALCOHOL, FREEDOM AND RESTRAINT
Patsy Staddon

255 EUROPEAN IDENTITY AND THE EUROPEAN DIMENSION IN ENGLISH AND GERMAN SECONDARY SCHOOLS
Sandra Barkhof

265 OVERCOMING THE MUDDLE AND MESS: THE NEED FOR PRAGMATISM AND FLEXIBILITY IN AN EMPIRICAL STUDY OF INFORMATION SHARING IN CHILDREN’S SERVICES
Sue Richardson

276 THE APPROPRIATE ROLE OF THE STATE WITHIN THE ETHICAL PARADIGM
Alan Sanderson
UNIVERSITY OF PLYMOUTH

PLYMOUTH BUSINESS SCHOOL AND SCHOOL OF SOCIOLOGY, POLITICS AND LAW POSTGRADUATE SYMPOSIUM

Dr. Adrian Barton and Dr. Jonathan Lean

Introduction

The first Plymouth Business School and School of Sociology, Politics and Law Postgraduate Symposium was held on January 14th, 2006. Bringing together research papers from many different subject fields, these proceedings provide a record of this stimulating and lively event.

The business and management papers contained within these proceedings demonstrate the diverse range of PhD projects being undertaken by research students within Plymouth Business School. The papers cover topics relating to international business, shipping and logistics, financial management and human resource management. The nature and focus of studies presented reflect in part the stage that students have reached in their research. Some papers report on findings emerging from primary data collection, whilst the work of recently registered students is focused more on aspects of the literature and initial research design considerations.

Zhouran Yang’s paper reviews literature relating to strategic human resource development (SHRD). Drawing on the work of various authors, a model of SHRD is outlined. In the next stage of his study, the researcher intends to apply the model within UK and Chinese businesses. The paper by Sonja Derrick focuses on human resource management issues within a specific organisational context. The study described uses a system dynamics approach to model the impact of the EU Working Time Directive on the training and service activities of junior doctors in the NHS. Sonja was this year’s recipient of the prize for best symposium presentation in business and management.

Francisco Serra’s paper is one of several presented under the theme of ‘international business and management’. Investigating the factors associated with export performance amongst Portuguese and UK SMEs in the textile industry, the author considers results relating to the impact of firm size and manager characteristics amongst others. Meanwhile, André Bühler
provides an overview of his recently submitted thesis concerning professional football sponsorship in the English Premier League and the German Bundesliga. The significance of relationship marketing within this context is one of the issues highlighted by the author. In his paper, Dabarata Chowdhury provides a descriptive introduction to issues pertaining to his research. His study, currently in its early stages, seeks to explore the role of Turkey in knowledge and resource transfer between Europe and Asia. The paper by James Liu presents results emerging from his research into the strategic motives associated with Sino-Taiwanese joint ventures. The study provides valuable insights into the operation of social and business networks within a unique political context.

Two papers are concerned with topics relating to ‘international shipping and logistics’. Kelvin Pang’s study seeks to assess the role of North Korea in the logistics development of the Korean peninsula. A conceptual model exploring the factors affecting transport and logistics development is proposed and methodological approaches for future testing of the model are outlined. The paper by Robyn Pyne discusses the nature of problems associated with maritime crew communication and the role they play in accidents at sea. It builds a case for further research into the sociolinguistic aspects of shipboard operation in order to limit Human Communicative Error.

The final set of business and management papers are all concerned with issues in finance. Hussein Abdou’s work presents the findings of a pilot study which reveal that Egyptian banks do not utilise statistical techniques in their evaluation of credit risk. Details of a further stage of data collection using an attitudinal questionnaire survey are outlined. Rumbi Mukonoweshuro describes the findings of initial empirical tests relating to a model for optimal dividend determination. The contribution of the research in considering specific industrial and regional contexts is discussed. Finally, Jonathan Challacombe presents a review of the literature on the methods of financing ships up to the post world war two era. The paper represents early research conducted for a project concerning past, present and future trends in ship finance.

In many ways, the Sociology, Politics and Law School papers mirror the range and diversity contained in the PBS papers. Interestingly, the papers offer an overview of the various stages of the PhD process and the manner in which students approach their work. For example, Greta Squire’s work is in its early stages and her paper is literature-based examining the manner in which the incorporation of voluntary sector groups (aimed at helping survivors of domestic violence) into Crime and Disorder Reduction Partnerships has had a marginalising effect on voluntary groups. Greta’s paper is an overview of the current literature and poses some questions
which will form the basis of her empirical work. Likewise, Mark Harron’s work, offers a different focus on aspects of discrimination in police forces, post-McPherson, by examining the Special Constabulary in terms of recruitment and diversity. Again, Mark offers a comprehensive review of the key literature, provides existing statistical information, and concludes by offering suggestions for the future direction of his work. Annette Elliott’s paper provides a fascinating insight into the literature and ideas surrounding gender, peace and war by asking us to reflect on our understandings of war and peace and how a gender aware perspective is essential to gain a holistic picture of attempts to create a peace that is sustainable.

Ossie Glover and Mark Pearson are mid-way through their work and their papers reflect the transition between examining the current literature and collecting empirical data. Ossie’s work is an example of this. His paper provides a discussion of his methodological approach to evaluating the impact of Black and Minority Ethnic groups’ experience of New Deal for Communities Programmes, and the potential this holds for enabling a greater understanding of how, why, and if such programmes work. Mark Pearson’s paper focuses its attention on the use of Evidence Based Medicine in social policy, the construction of a ‘hierarchy of evidence’ and offers an interesting examination of the political economy of research and evaluation.

By contrast, Patsy Staddon and Sandra Barkhof are entering the final stages of their work and they offer a summary of their key findings linked to the theoretical frameworks literature searches provide. Sandra’s paper examines the relationship between European identity and popular support for the European Union and explores which possible means could be employed to create or at least stimulate the development of a European identity. Patsy’s deals with the manner in which society reacts to women and alcohol consumption and argues that women’s drinking habits are viewed as doubly deviant, firstly by contravening social norms concerning alcohol consumption but secondly, and more importantly, running against the social expectations of what a ‘good woman’ should be.

Finally, Sue Richardson and Alan Sanderson are nearing completion of their work, and, although taking completely different approaches, offer an ‘old hand’s’ perspective on their particular work. Alan’s paper asks questions about the role of ethics in the management of public services, which whilst only a peripheral concern to the main thrust of his thesis, demonstrates that research generates at least as many questions as it answers. Sue’s work, on the other hand, provides a warts and all account of the research process and serves us all a timely reminder that good researchers, and by extension, good research, are a product of a mix of a rigorous
methodological approach, solid theoretical underpinnings plus a large degree of flexibility and pragmatism.

The papers below are the full versions of presentations each contributor gave at the Postgraduate Symposium. The rationale behind this event was twofold: to provide a forum for post-graduate students to present their findings in a familiar, supportive and encouraging environment; and to allow students and full time staff to see the diversity and quality of post-graduate research within the two Schools. The event was a success due largely to the quality of the work and the enthusiasm and knowledge of the presenters and seems likely to become an annual event in the University of Plymouth’s calendar. However, as always with events such as these there are people that work especially hard on planning and organising the day. With that in mind, we would like to thank André Büehler, Sue Richardson and Mark Pearson for their enthusiasm and organisational skills and their input into making the event such a success.
A MODEL OF STRATEGIC HUMAN RESOURCE DEVELOPMENT

Zhuoran Yang

Abstract

Strategic Human Resource Development (SHRD) provides a new perspective on managing change for the 21st century. In doing so, it promotes a more enlightened, ethical and skills-focused vision of change management by placing human resources back where they belong - at the forefront of the change agenda. The aim of the research is to explore a model of SHRD which can work effectively in Chinese organizations by considering some important characteristics.

The research is designed in three stages. The first is to review the literatures on strategic human resource development and explore the concept specifically in the context of the work of Garavan (1991) and build up an effective model of SHRD. Then searching for more evidence to support the model from company’s practices and conclude with a new model of SHRD. However, according to the culture differences, the new model will be implemented in Chinese companies for a one year period. By analyzing the data results before and after the model implementation, it will be concluded with positive or negative views of the new model of SHRD.

This research will take four years and will involve several companies in the UK and China.

Introduction

More and more senior managers take a greater interest in the development of their organization’s human resource in order to remain in the competitive marketing. Its may caused by the difficulties in recruiting skilled managers; the need to align the potential of employees with business objectives or a greater emphasis on performance evaluation and management in day-to-day jobs. Senior management is beginning to look inside of organisation and focus on the HRD function to provide solutions to its business problem. According to Yang (2004) research, HRD is best seen as the strategic management of training, development and of management intervention, in the same time to ensuring the full utilisation of the knowledge in detail and skill of individual employees. However, if HRD concerned with the
management of employee long term learning and corporate with business strategic plan, then it will raise to next step-- the strategic human resource development (SHRD).

The concept of strategic HRD (SHRD) has been much explored in the training and development literature of the last decade (Higgs, 1989; Keep, 1989; Noel and Dennehy, 1991; Garavan, 1991; Holden and Livian, 1993; Saggars, 1994; Sloman, 1994, Rainbird, 1995; Garavan et al., 1995; Torraco and Swanson, 1995; Lee, 1996b; Stewart and McGoldrick, 1996; Harrison, 1997; O'Donnell and Garavan, 1997; Garavan et al., 1998), but there are weak evidence to show on what characterises an organisation with a strategic approach to HRD. This paper will reviews the literature on the model of SHRD, especially will focus on the characteristic of SHRD. In order to re-develop a new model of SHRD, Garavan’s model of SHRD (1991) will be used as a comparison study example. Garavan (1991) highlighted nine key characteristics of SHRD was be used as a foundation from which to examine the ways in which the concept of SHRD had changed since 1991.

The aim of this paper is to review and examine the literature of nine characteristics of SHRD and moving towards a new model.

Defining SHRD

Before doing any analysis, it is crucial to review some definitions of SHRD.

Rothwell & Kazanas defined Strategic Human Resource Development (SHRD) is the process of changing an organisation, stakeholders outside it, groups inside it, and people employed by it through planned learning so that they possess the knowledge and skills needed in the future (Rothwell & Kazanas, 1989). In their theory, SHRD should focus on the HRD Effect, in another word, defined as the coordinated learning activities undertaken by HRD practitioners, operating managers, and employees to support business and HR plan. This is a comprehensive, coordinated plan for major learning initiatives by which a firm’s managers intend to meet business and staffing objectives through organised learning.

Garavan et al. (1995) also contends that the term HRD is used in many different contexts and that it concerns a range of widely differing activities, leading to considerable confusion about its use. Despite these difficulties, however, a number of useful definitions of SHRD exist, including the following:

*The strategic management of training, development and of management/professional education intervention, so as to achieve the objectives of the organisation while at the same time ensuring the full utilisation of knowledge in detail and skills of individual employee. It is concerned with the management of employee learning for the long term keeping in mind the explicit corporate and business strategies (Garavan, 1991,p17)*

*Strategic HRD can be viewed as a proactive, system-wide intervention, linked to strategic planning and cultural change. This contrasts with the traditional view of*
Training and development as consisting of reactive, piecemeal interventions in response to specific problems (Beer and Spector, 1989, p25)

Other definitions like Harrison (1997) and Stewart & McGoldrick (1996), stress the need for SHRD to operate within, be linked to and keep in mind, corporate strategy. In another word, HRD should be responsive to corporate strategy. However, it is important to ask whether SHRD can occupy more than the rather reactive role implied by this suggestion. Moreover, it is suggested (O’Donnell and Garavan, 1997; Harrison, 1997; Torraco and Swanson, 1995) that HRD should become more strategically focused, but what, specifically, does being more strategically focused mean? Therefore, the following section will explore the nine SHRD characteristics suggested by Garavan (1991) (Figure 1), each of them will be re-examined in order to check the possibility of working efficiency in the whole SHRD model.

**Figure 1:**

**Key SHRD characteristics**

| 1. integration with organisational missions and goals |
| 2. Top management support |
| 3. environmental scanning |
| 4. HRD plans and policies |
| 5. line manager commitment and involvement |
| 6. existence of complementary HRM activities |
| 7. expanded trainer role |
| 8. recognition of culture |
| 9. emphasis on evaluation |

Source from: Garavan (1991)

**The Characteristics of SHRD**

There is growing literature on the characteristics of strategic HRD activities, some of them have not been updated since they been identified. It is therefore possible to review them again and identify a number of key characteristic.

Before going to the Garavan’s (1991) model, it should be pointed out that Rothwell & Kazanas’ (1989) model about SHRD. They identified seven characteristics of SHRD in order to guild an HRD department in planning its role in the HRD effort. The seven characteristics show in Figure 2.
Figure 2:

The seven characteristics for SHRD

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>clarity the purpose of the HRD Effort</td>
</tr>
<tr>
<td>2.</td>
<td>assess present condition</td>
</tr>
<tr>
<td>3.</td>
<td>scan the external environment</td>
</tr>
<tr>
<td>4.</td>
<td>compare present strengths and weaknesses to future threats and opportunities</td>
</tr>
<tr>
<td>5.</td>
<td>choose a long-term organisational strategy for HRD</td>
</tr>
<tr>
<td>6.</td>
<td>implement organisational strategy for HRD</td>
</tr>
<tr>
<td>7.</td>
<td>evaluate</td>
</tr>
</tbody>
</table>

Source from: Rothwell & Kazanas’ (1989)

Rothwell & Kazanas’ seven characteristics are basic on some assumptions. They did make some good suggestions, for instance, SHRD should have clear long-term goals and scan the external environment before it been implement. However, it is arguable that management support inside the organisation, especially from the top management. This is simply because HRD can not support strategic business planning and long-term HR planning if top management refuse to give them support.

After Rothwell & Kazanas’ seven characteristics model of SHRD, Garavan (1991) developed a new model’s of SHRD with nine characteristics, the new model have concerned with some issues about how HRD become more strategically, but, it is the time to review them one by one in order to check the possibility of working efficiency in SHRD model.

Integration with organisational missions and goals

In order to contributing to corporate goals and awareness of the mission of the organisation, Garavan (1991) argued that the need for integration into business planning is critical for HRD. HRD is viewed as a strategic lever in organisations because it is seen as a means of helping the organisation to implement its business strategies (Garavan, 1998). On Devanna et al’s (1984) HRM model, it’s been suggested that SHRD role should be responsive and reactive. Alternative, a more proactive role may actually be for the real SHRD, moving away from a strategy supporting and implementing role, towards a role where SHRD helps to shape and influence, corporate strategy (McCracken & Wallace, 2000). Furthermore, Torraco and Swanson (1995) suggested that today’s business environment requires that HRD not only support the business strategies of organisations, but that it assumes a pivotal role in the shaping of business strategy (Torraco & Swanson, 1995, p11). For example, an organisation which there are frequent technical innovations, then HRD can have a strategy shaping, rather than simply supporting role. HRD could therefore, at least in theory, play a role in either implementing and or/forming corporate strategy.

Lee (1996) designed a model in relation to management development, also suggested that SHRD takes a proactive role in strategically mature organisations. In
his theory, SHRD helps to shape and formulate corporate strategy for the organisation. Integration with organisational missions and goals, as suggested by Garavan (1991), therefore suggested an implementation role for HRD, but truly strategic HRD should also shape and influence these missions and goals (Legnick-Hall, 1998; Butler, 1988).

Top management support

Strategic HRD must command the support and participation of top management (Burgoyne, K, “Management Development for the Individual and the Organisation”, Personnel Management, June 1988). The essential message is that when management at board level treat HRD as making an important contribution, the HRD function will get a higher profile.

In many cases top management is not committed to training and development activities. This is because of several reasons: (a) they are too busy to consider the benefits of HRD or identify what resources and facilities are available; (b) expenditure on training does not reflect investment in a financial sense but is instead perceived as a cost (King, W, “Advisory Power and Predicament in Organisational Training”, Journal of European Industrial Training, Vol. 13 No.3, 1989, pp 19-31); (c) The concentration of training is focused on the bottom end of the management hierarchy and ignoring the senior level managers (Brown et al, 1989). Lee (1996) considered these issues and pointed out that SHRD support needs to be from the point of view of their operational roles and their own personal development. In another word, according to Harrison (1997), it must be in the form of active, rather than simply passive, support and involvement. Moreover, Torraco and Swanson (1995) suggested that in order to get such support from top managers, HRD professionals must demonstrate their strategic capability, crucially by helping strategic planners to acquire the conceptual, analytical and interpersonal skills they require to do their jobs properly. The same suggestion is made by Garavan et al. (1998) that HRD is often perceived as a means of helping managers to adopt strategic thinking. The role of such managers as “key actors” in HRD (Garavan et al. 1998) suggests that their active leadership of HRD (rather than their passive compliance) is critical. As a key SHRD characteristic, top management support should therefore become top management leadership (McCracken & Wallace, 2000).

Environmental scanning

The strategic HRD function must have reasonable knowledge in order to understanding the external environment. According to Higgs (1989) and Garavan (1991), a Strategic HRD function must have the capacity to analyse the external environment in terms of the opportunities and threats which it presents in the context of both the business and HRD strategy. It gives the advantage to the business when highlights the role of human resources as a key component from a competitive environment. This in turn will present the opportunities to discuss HRD strategic at the highest level and ensure that the first characteristic discussed, that of integration, is achieved. On the other hand, the environment can also act
as a threat to the HRD function. Stiff competition, if not tackled effectively, can reduce profit margins and consequently lead to a reduction in training budgets. This situation clearly pointed to the need for HRD strategy to be well positioned in the overall business context.

However, it is crucial that other senior managers, and not just HRD professionals (Torraco and Swanson, 1995; Sloman, 1993), should be gathering such information. If the company need to do a SWOT or PEST analysis, who takes the responsibility and whether it is done specifically in HRD terms is critical. Truly strategic HRD operates where senior management automatically consider the HRD implications of any changes in the internal or external environment, rather than seeing this as the job of the HRD specialists (Rainbird, 1995; Peery and Salem, 1993). Only in this way, HRD can become properly integrated into the organisation and into the strategic planning process. Therefore, simply scanning of the environment is not good enough. It should be done specifically in HRD terms and by senior management.

HRD policies and plans

It seems policies and plans formulation has in the past been ignored in the training and development literature. However, in more recent times, there is a growing recognition of its importance within the HRD context (Hendry and Pettigrew, 1986). From this literature it is possible to conclude that, for HRD to be strategic in focus, it must formulate plans and policies which flow from, and are integrated with, business plans and policies. In another word, HRD plans can influence, rather than simply react to business plans (Hendry and Pettigrew, 1986). This becomes a very crucial and underdeveloped point. The implication should be a strategic emphasis rather then an operational. Moreover, plans and policies are clearly necessary, but it is also vital that there are HRD strategic, developed by the top management team. Harrison (1997) defined strategy as:

\[
\text{The route that has been chosen for a period of time and from a range of options in order to achieve business goals. It is a guide to action and therefore sets the scene within which policies—including those employee resource (ER) policies relating to the learning and development of people—can be agreed and implemented (Harrison, 1997, p.19).}
\]

It is therefore about the present and future direction of the organisation, from a broad perspective, whereas policy can be seen as the specific route to be followed and the tasks to be undertaken in order to achieve the strategy. Training plans represent the next level down and usually consist of the details of priority training interventions from the point of view of who, how, when and where. For SHRD to thrive, therefore, HRD policies and plans need to be supplemented by HRD strategies.

Line management commitment and involvement

Garavan (1991, 1998) identified the key stakeholders and “actor” in HRD is line managers and Zenger (1985) pointed out that an enthusiastic involvement of line
manager would help HRD function become more strategically-focus and effective. To extend this point is that the role of HRD needs to be clear about the kind of support they expect and desire from other stakeholders or “actor”. Rainbird (1995, p83) regards the trend towards development of HRD responsibilities to line as: “A key indicator of ability to integrate human resource and business strategies.”

It could be argued that such development involves only operational issue for the line managers and the strategic issues always be left to the HRD specialist. What is animate is in order to line managers and HRD specialists work as partnership on both operational and strategic issues. Lee (1996b) and Harrison (1997) identified the needs for shared ownership of HRD, where line managers and HRD staff work in partnership over HRD issues. In reality, unfortunately, it seems not in the same way as they identified because of a variety of complex attitudinal and cultural reasons (Harrison, 1992; Garavan, 1987; Leicester, 1989;). For example, Garavan et al (1998) found that many of the “actors” in HRD either claimed exclusive ownership of HRD strategy or felt that it rested with someone other than themselves. In order to getting more strategic views for HRD function, according to McCracken &Wallace (2000), line managers should drive as well as deliver HR policies and what is crucial for the development of SHRD is that they do this in partnership with HR professionals.

Existence of complementary HRM activities

Garavan (1991) concluded that not only does strategic HRD provide an impetus for a more integrated HRD approach, but it is also necessary to have a close integration between HRD and other HRM activities. It is very important for the development of SHRD, but others (Storey, 1994; Harrison, 1997) question whether in reality such integration exists. What are needed are strategic partnerships between HRD and HRD professionals of the kind implied by O’Donnell and Garavan (1997), where an alliance between HRD strategy and a global-arching HRM strategy is suggested. Without such real alliance and integration between HRM and HRD, to the extent that they are seen as one and the same, there is probably little hope of either having much impact on the achievement of corporate objectives (McCracken &Wallace, 2000).

Expanded trainer role

A strategic HRD function requires a specialist who can be a combination of trainer, provider, consultant, innovator and manager (Garavan, 1991). He also suggested that a strategic HRD function requires who can be innovators and consultants, rather than simply providers or managers of training. It should be involved training, learning, or organisational change issues. The former roles tend to concentrate on meeting the strategic needs of the needs of the individual. As Talbot (1993) suggested that the roles of the trainer could include:

- adaptive (adapting the skills and knowledge of staff to fit existing system);
- adoptive (getting staff to adopt new values or attitudes);
- innovation (informing and influencing organisational change processes)
In order to make a strategic contribution, HRD staffs also need to be leaders, as well as facilitators, of change, despite the possibility of underlying doubts about their own credibility (O'Donnell and Garavan, 1997). They need to be more proactive rather than simply reactive and to see themselves in a central and strategic, rather than peripheral and operational, role (Garavan, et al, 1998). Therefore, to expand the role for trainer is more likely move to the place of an emphasis on strategic. Not only should HRD staff understand their own responsibility of strategic views, also they need for a common understanding across the organisation about the role of HRD staff. The need of such congruence between the trainer's own perception of their role and the expectations and the rest of the organisation (Bennett and Leduchowicz, 1983).

Recognition of culture

A strategic HRD function must be aware of and take account of the relationship between culture and implementing HRD activities. The culture could be inside the organisation or the whole nation culture background. Culture also is seen as an important variable in deciding how HRD should be delivered and evaluated. SHRD should be ready to influencing and changing the culture, rather than simply maintaining, corporate culture. Burack (1991, p.88) gives a strong suggestion as:

*Human resource development has a crucial, challenging role to play in successfully “orchestrating” strategic culture change.*

Recognition of culture, as pointed out by Garavan (1991), may not be enough. Culture is a complex concept, therefore, the influence which SHRD might have in changing corporate culture could be even more difficult to isolate and clarify. However, it has been suggested that on a learning organisation, there would be more likely under the control, because learning, can be both a product of culture as well as a means of transmitting and changing culture, although the interrelationships are complex and poorly understood (Garratt, 1987; Senge, 1990; West, 1994). In other word, it is crucial for HRD function to recognise the potential culture, develop and enhance the HRD activities, corporate with the current culture in order to make the strategic views.

Emphasis on evaluation

Strategic evaluation is a key component of the strategic management process (Garavan, 1991). If the HRD function wants to have a strategic focus then it must evaluate its activities. Johnson and Scholes provide a useful evaluation framework which has read application to the HRD field. They mainly focus in the three categories of criteria: *suitability; feasibility and acceptability*. Furthermore, Zenger and Hurgis point out that strategically-oriented HRD functions spend time evaluating their activities and they use measures that seem most relevant to the management of their own organisation.

However, it been argued strongly by Torraco and Swanson (1995), Harrison (1997) that training also should be results-oriented and that evaluation should take place to addressing whether business needs have been met. This means the evaluation should be controlled in order to make it effective. Therefore, some degree of cost-
effectiveness control should be involved. But effective-cost-control is difficult. If the organisation lacks of training investment calculations, it may lead to a misunderstanding of how it is important in the long-term future of the organisation. Moreover, well cost-effectiveness evaluation can lead to a concentration on achieving quantifiable results within a short timeframe and it is vital not to undermine the importance of HRD as an investment with long-term and less tangible benefits (such as culture change) (Lee, 1996a). Nevertheless, the importance of cost-effectiveness evaluation should not be underestimated (Harrison, 1997).

Model of SHRD

Having examined the nine characteristics, it is vital to stress that in reality the issue are not so easily separated out. As (Baird and Mashoulam, 1998) identified that all components should intricately interlinked and should complement and mutually support one another. If one or more is absent or weak, then this could significantly undermine the development of SHRD. For instance, top management leadership without strategic partnership with line management would mean that HRD was not properly embedded into the organisation and this would represent a significant barrier to the development of a learning culture. The organisation would almost certainly be less strategically mature in HRD terms than one where both senior and line management commitment to HRD were strong. The same could be argued for almost any combination of the characteristics. Therefore, the truly relationship between the SHRD characteristics should be showed as Figure 3. Everyone should combine together to make the contribution, but most importantly, SHRD should have the strong ability to influence other part of the relations, so that SHRD emerges and thrives.
Conclusion

As a result of reviewing the literature, it could be summarised the differences between Garavan’s (1991) description of the characteristics of SHRD and those developed above. Each description has been expanded and enhanced.

The new enhanced version of the nine characteristics suggested that SHRD is about more than “keeping in mind” business strategies (Garavan, 1991,
The role for SHRD should have a much more proactive and influential. This could lead to a different model for SHRD which is detailed in Figure 4.

SHRD, where the role is proactive in both shaping and responding to corporate strategy. SHRD specialists have developed a strategic and innovative role as organisational change consultants (Philips and Show, 1989) leading change. The organisation has become strategically mature in HRD terms and evidence of all nine of the enhanced Garavan (1991) characteristics, as well as a strong learning culture, is in place.

**Figure 4: a Model of SHRD**

- Shaping organisational mission and goals
- Top management/leadership
- Environmental scanning
- HRD plans and policies
- Strategic partnerships-
  - External
  - Internal
  - Line managers
  - HRM
- Change agent/consultant
- Influence culture
- Cost-effective evaluation

*Model of Strategic HRD*

Yang and Badger 2005
References


MODELLING THE EDUCATION AND SERVICE ACTIVITIES OF JUNIOR DOCTORS UNDER THE EU WORKING TIME DIRECTIVE

Sonja Derrick

Abstract

If junior doctors are to work significantly fewer hours in the future, how can they still receive full training and continue to provide necessary levels of service to patients? Historically, excessive hours have been a way of the life for junior doctors, but New Deal regulations, a revised junior doctor contract, and the EU Working Time Directive (EU WTD) are changing this. Additionally, postgraduate medical education is undergoing significant changes under Modernising Medical Careers. This research seeks to use system dynamics as a tool to model the impacts of the re-organisation of the education and service activities of a professional group within an organisational context, by taking junior doctor service and training under the EU WTD as the specific example at hand. The necessarily tailored approach taken to developing the model has yielded important insights into the relationship between junior doctor “training” and “service” and how these are related to the activities they participate in during working hours. Valuable knowledge and lessons were also gathered from the model building process itself, which has involved continuous iterations and modifications in describing this social system.

Background

Historically, excessive hours were integral to the life of a junior doctor; long hours on-call, soaking up knowledge and training alongside other grades in their specialty. Increased recognition of lack of safety for patients and adverse effects on learning and health of junior doctors has led to the introduction of the New Deal working hours regulations and the revised junior doctor contract. Additionally, the European Union Working Time Directive (EU WTD) came into force for junior doctors in the UK in August 2004, requiring a reduction in working hours.¹

Further, postgraduate medical education in the UK is changing under the introduction of Modernising Medical Careers, which sees the restructuring of junior doctor grades and formalising of the curriculum and assessment of training. Postgraduate medical career progression until 2005 can be seen in Figure 1a and a simplified diagram of the new grades proposed under Modernising Medical Careers in Figure 1b.
In an effort to reduce hours, Hospital Trusts have had to introduce significant changes in working patterns for junior doctors: moving from on-call to shift working, increasing numbers of service based doctors, reducing tiers and increasing cross cover, as well as new ways of working and extended roles piloted by the Department of Health. The implementation of Modernising Medical Careers, which aims to restructure and formalise postgraduate medical education will bring further changes from 2005.

Further, it is also feared that as training becomes more intense and more doctors will have to be trained to fill the service gap, then pressures on consultants' time will also increase. Alongside which, it is not only junior doctors that have to comply with hours regulations, but also senior members. In principle, consultants have always fallen under the 48-hour working week imposed by the EUWTD. However, this issue has become more prominent in the recent negotiations of the new consultant contract, and demands continue to increase as hospital admissions rise and growth of service provision flourishes. Non-compliance is not an option,
and failure to train junior medical staff or provide services to patients results in a withdrawal in funding or other financial penalties for NHS Hospital Trusts.

A literature review revealed a significant number of studies into impacts of the EU WTD and its equivalent initiative by the ACGME in the U.S., which reduces resident working weeks to fewer than 80 hours. The research reveals a general concern of potential negative impacts of reduced hours on training and patient care, particularly on the continuity of care, but there are mixed results. Although some use quantified measures of training (e.g. operating hours, patient episodes experienced), there is a trend to use junior doctors and senior medical staff’s perceptions and attitudes in evaluating impacts on experiences of training. However, what is most evident is that the existing research of impacts of reduced hours working on “training” suffers from a lack of definition of what is meant with “training”: is it classroom teaching, on-the-job experience, learning, or all of the above?

Given this lack of direction, it is not surprising that relevant research results to date are inconclusive and contradictory. While some studies report a significantly lower training experience and training satisfaction scores for participants while others purport that either training standards are maintained, or even improved due to better learning and working conditions associated with less tiredness. The literature also refers to adverse impacts on training in relation to loss of continuity of care and that when hours are reduced, there are fewer opportunities for learning, a dilution of experience and “service” takes precedence over “training.”

It is immediately apparent this is a “messy problem” beyond simple optimisation. Reducing junior doctors’ hours involves changes to working patterns, which have very real causally linked impacts, of both a qualitative and quantitative nature. Further, failure to consider the system as a whole is likely to have short and long-term impacts on: training and experience for junior doctors, finance and service provision for the Trust, patient care, and the development of the current and future workforce. Proposed new ways of working need to be evaluated in light of all the potential consequences, beyond simple compliance and costs.

The Context

Plymouth Hospitals NHS Trust, whose main site is at Derriford Hospital in Plymouth, provides acute and specialist care services to approximately 450,000 people in Southeast Cornwall and Southwest Devon, and covering a population of almost 2 million people for some specialist services. With a budget of approximately £250 million, 1300 beds and the busiest A&E department in the South West of England, it employs approximately 6000 people, including approximately 450 junior doctors and 230 consultants. It is a major employer and provider of health services in the southwest peninsula.

Plymouth Hospitals NHS Trust have been working in partnership with the University of Plymouth to analyse the situation they are facing regarding the reconciliation of
training and service needs of junior doctors with the EU WTD. From the outset, it was proposed that systems dynamics modelling be a key approach in the analysis, to be the basis for senior Trust management to understand this complex system, and aid in their strategic planning and decision making. This will allow not only a thorough comprehension of the situation, but also support decision-making by modelling possible future outcomes of proposed scenarios. In particular, it will aid informed strategic decision-making in an industry that is often features short-termism in its decision-making, imposed by year-to-year budgets, meeting constantly changing government targets and constant crisis management.

The aim of this paper is not to document extensively all the activities and outcomes of this 3-year programme of research, but rather to review and highlight some of the major findings, in particular some important insights into the relationship between junior doctor “training” and “service” and how these are related to the activities they participate in during working hours. Further, while the final version of the model and its outputs are not yet available for publication, the process of model building and initial experimentation have yielded some valuable insights into the complex problem to date.

Approach

It was proposed to use system dynamics as an overall method of modelling the impact of the reorganisation of the education and service activities of this professional group (junior doctors) within an organisational context (reduced hours and changing training programmes). Born from Forrester’s Industrial Dynamics (1958), system dynamics (sometimes referred to as business dynamics) is an approach for studying managed systems. The basic principle lies in the assumptions that there is feedback in systems and that a system’s behaviour is directly related to its structure. By representing the structure of a system, be it a physical or social system, with a series of connected stocks and flows with identified relationships, it is possible to concretely model and simulate system behaviour. System dynamics enables not only the understanding of a system’s elements and how they are related and influence behaviour, but once the structure of the model has been accepted, it also allows simulation and prediction of future trends and behaviours in the system. However, while the iconography of system dynamics models are formalised and standardised, there is a wide variety of contexts in which they have been applied, and accordingly a wide variety in approaches adopted in the construction of these models.

While the ill-defined features and relationships of this problem make it ideal for applying an operational research modelling technique, and the link between structure and behaviour in particular warrants a system dynamics approach, it means that the traditional consultancy-style model building approach with key decision makers does not apply until more clarification has been sought. In order to successfully model this “problem”, it has been necessary to design a two pronged approach of research necessary to define and clarify the problem before it could be modelled to the level of detail required by the problem owners.
Early on in the research, both from the literature review and from conversations with problem owners, it became apparent that there was a lack of data and definition concerning both the concepts of “training” and “service” for junior doctors, how these related to what they did at work, as well as how much of their time is spent in training and service provision. Before the model building process could be initiated, it was necessary to bridge these gaps in knowledge, so that they could be accurately represented in the model of this social system.

The Training/ Service Continuum

As described, it was necessary to scope and clarify the concepts of “training” and “service” and question the status quo of the simultaneous nature of the two that was currently assumed in the profession. In particular, with reference to junior doctors it was important to understand: what constitutes training and service? How are they related? Is this a continuum or binary (either training or service)? How do junior doctor activities fall along the training/service continuum and what factors affect this balance?

In addressing these questions, data was collected through questionnaires and focus groups. The entire population of Senior House Officers (a grade of junior doctors) at Plymouth Hospitals NHS Trust (PHNT) were targeted with a questionnaire, yielding quantitative data about training/service balance perceptions and possible independent variables influencing these. Subsequently a sample of SHOs, Consultants and Tutors participated in focus groups, providing qualitative data on
the same matter, gauging balance of training/service activities and identifying factors that affect these. The focus group data helped validate and added depth to the data collected through the questionnaires.

The outcome of this was creation of the training/service continuum diagram, as seen below. This shows how junior doctor activities are placed along the continuum, depicting the inter-quartile range (the peak displays the median, the higher and lower ends of the shapes show the 75% and 25% quartiles). This provides an indication of the level of consensus amongst respondents in their perceptions of the training/service balance in the activities.

Figure 3 - The training/service continuum of junior doctor activities

A number of observations can be noted from these results. Firstly, the more frequently undertaken activities (pictured in the darker shades) are generally towards the service end. Secondly, the activities that are generally supervised by a more senior colleague (the starred activities) are generally placed towards the training end. This indicates that frequency and supervision have a high impact on the perceived training/service balance of an activity. This was confirmed by analysis of the questionnaires responses. Regressions and correlations showed that
just under 70% of variation in training percentage perception could be explained by variations in frequency and supervision across the 28 activities.

Finally, in confirming the experience that there are a wide variety of views on the issues dealt in the training/service/hours conundrum, it is apparent that for some activities, there is a much higher level of consensus than others. For example, there is a high level of consensus about the training/service balance in a SHO teaching session, but a more variable activity in its execution, such as ward round, or discussing patients with colleagues has a high interquartile range of responses, indicating a far lower level of consensus in views.

Complementing these results was a list of factors that affect the training/service balance in the activities, which were arrived at from the focus groups. These included: time available for undertaking the activity, number of patients (workload), type and nature of work in the specialty, the individual trainee and trainers motivation and attitudes, interaction, purpose and focus of the activity, other commitments, experience and competence of the junior doctor. These were all factors regarding the execution of the activities. Some are quantitative features that can be more readily incorporated in the system dynamics model. Others, such as motivation and attitude, are not easily measurable, but this does not make them any less significant. In fact, this problem features a number of “soft issues” that have very definite real and “hard” impacts. This is something system dynamics should be amenable to modelling.

How is junior doctor time spent?

Once it was established what activities junior doctors partake in and how educational or service based these are perceived to be, it was necessary to gauge how much time was spent in these activities, in order to understand what the balance between “training” and “service” is in the working lives of junior doctors. The intention was to incorporate the concept of training/service balance in the systems model, as its importance to the problem stakeholders had become apparent in early discussions.

Analysis of time-use data, collected at PHNT as part of the Medical Workforce Skills Mix Analysis project in 2003, where time spent undertaking each of the 28 activities was multiplied by the relevant training or service percentage, showed that 45% of SHO time is spent “in training” if the “empty” time (i.e. time spent doing other sundry activities such as walking, waiting, time in the doctors’ mess (lounge), etc.) is not accounted for. However, if this time is included then the average time spent “in training” falls to 23% (implying conversely that 76% of SHO time is spent providing service). Detailed results of this time analysis can be found in Derrick et al.


While there are a number of problems with this analysis, mostly deriving from the quality and accuracy of the time-use data, there are two major issues discovered in the time use analysis:

**a.** Across the 24-hr period, there appears to be a relatively large proportion of “slack” or “empty time” in the system (almost 50%): time that could be either reduced (perhaps questioning the need for a doctor to be on duty) or put to better use. Perhaps some of this “empty” time could be spent educationally, either in personal study or skills labs. While no one is (nor should be) 100% productive, the slack identified above is significant.

**b.** A relatively high proportion of aggregate time is spent on tasks such as patient related administration and routine non-complex clinical tasks that could either be made more efficient or in part be supplemented by other members of the team (e.g. doctors’ support workers) thus freeing up more time for service and education.

Inevitably, the validity and reliability of these results could be greatly enhanced by more accurate and detailed data, but interestingly, the 23% figure is similar to the current assumptions made by many clinicians regarding junior doctor training, who have been highlighting that the bulk of NHS service is delivered by the SHO grade, and the 45% figure is close to the target amount of time that SHOs should spend “in training” (according to funding arrangements and Modernising Medical Careers guidelines that propose a 50/50 split). This discrepancy illustrates how statistics and conclusions from studies can be “massaged” to suit, if accurate data is not collected. Although appreciated that this shadowing data is not representative, nor complete, due to the lack of resources, this is the only data of this type currently available for PHNT. It is used here to help provide an indication of the training/service balance for SHOs, and a good starting point for quantifying some of the relationships in the model, that previously haven’t been able to be measured in this way.
Issues arising from the model building process

The model building process that this research underwent is extensively described and discussed elsewhere. It involved extensive and iterative consultation with stakeholders to describe and quantify the system, including not only causal loop diagrams at scoping and interim stages, but also at times required steps backwards to move in a new direction for the resultant model to be valuable to the decision makers.

However, in undergoing this iterative process and consulting with senior Trust members, problem owners and decision makers, one of the main directions the model took, that was not anticipated at the outset of the analysis, was that it had also to encompass the consultants and other senior medical staff that were being affected by changes in the junior grade, rather than simply concentrating on the lives of junior doctors in isolation of the human and physical resources that they operate with.

Additionally, a big debate that evolved during model construction was the issue of granularity and the level of detail in the model. Initially, the system dynamics model included a lot of detail regarding the different grades of junior doctors and their progression through the system in order for the problem owners to identify the system in the format that was unfamiliar to them (the iconography of stocks and flows in system dynamics). However, this quickly resulted in participants being distracted from the longer-term and strategic issues regarding the entire system, because they were concentrating on the detail.

In order to keep attention on the whole system, rather than operational detail, the researcher reverted to developing a causal loop diagram (CLD) which summarised all the major feedback loops and relationships. This still reflects some of the more important day-to-day issues (e.g. bleep policies, doctors’ assistant roles) but also incorporates qualitative elements (motivation, seeking learning opportunities) that had been raised in the earlier focus groups. The CLD has been valued as a good representation of the broad system at the aggregate, overview level.
The causal loop diagram centres on the level of junior doctor experience and knowledge, and it helped all stakeholders get to grips with the key loops at the heart of the problem. Earlier analysis (the training/service continuum) had revealed that one of the major contributors to the level of junior doctor experience and knowledge is the amount of supervised training they receive from senior medical colleagues. In particular, focus has been turned onto supervision, or the lack thereof due to the EU WTD and revised working patterns. The diagram clarifies how the reduction in hours of consultants and the continuing, and even increasing demands on service, impacts on the availability of consultants for junior doctor supervision. The amount of service consultants can provide to patients (be it in the form of clinics, operations, ward rounds, or other activities depending on the specialty) is directly related to the number of hours they work, and their productivity or efficiency. This naturally decreases the residual service (i.e. work that still needs to be done). The less outstanding work there is for consultants to complete, the more time to spend supervising their junior doctors.

Obviously, in addition to having the time, a second factor that has to be their willingness to do so. Individual attitudes and motivation vary from one doctor to the next. However, it is universally appreciated that performing a task while simultaneously engaged in training takes more time than simple completing the task. Some of the doctors estimated that this could reduce their productivity by approximately 30%. Therefore, while supervision increases junior doctors’ knowledge, it decreases consultant productivity, which negatively impacts the
amount of service they can provide, moving the burden of service back onto the juniors. Against this, this supervision will increase the junior doctor knowledge, which means they will become more skilled consultants in the long run, which will increases long-term consultant productivity / efficiency. There are similar complex inter-relations in looking at the work-load, training, service and productivity/efficiency of the junior doctors.

Overall, there are many interrelated issues at hand in this problem, but the development of this CLD was very helpful in focussing attention on the key issues. Translating this into a stock-flow diagram, followed by a set of iterative revisions during consultation meetings with problem stakeholders, has resulted in the current version of the insight model shown in Figure 6. (It is appreciated that this is too small for detailed examination, but this is not the purpose of its exposition.) At the time of writing, this is still work-in-progress and is currently being validated, tested and experimented with to ensure the model accurately depicts both system structure and anticipated behaviour.

Several iterations of model building were also needed due to changing circumstances over the 3-year period of research, as the details of Modernising Medical Careers were specified, and also when it became clearer to the decision makers what aspects of the system required more attention, some of which they had learnt from participating in the research. In particular when stakeholders were asked to quantify the relationships between elements in the system, it became more certain: what it was in the system that was important to model, how dependent the results were on assumptions made and how little is often known about the system that they are living and breathing in.

The issue of granularity in the problem was eventually resolved by developing two parts to the system dynamics model. The first is a spreadsheet model which lists each of the junior doctor activities, their respective typical frequency in which junior doctors participate in them, whether they are likely to be supervised or not, and the perceived training/ service balance. This captures the operational detail and changes to working and training patterns can be input into this spreadsheet. The output of this is an overall training/ service split (in percentages) and number of supervised hours. This feeds into the second part: the system dynamics model, which allows the calculation of supervision needs and consequent impacts on consultant and more senior junior doctor time, as well as productivity and quality of training (based on assumptions made by the problem stakeholders).

Figure 6 - Stock and Flow Diagram of the System Dynamics Model to date (overleaf)
It is not the intention to extensively discuss the system dynamics model in this paper, especially as the development, testing and validation thereof is still in progress. However, it provides an opportunity to highlight some of the findings and insights into the problem found in the model building process and experimentation to date.

Insights to date

What has already become evident from this research and the analysis and models produced to date is that this is a complex system, where system dynamics can be a valuable analysis and decision-making tool, for strategic and operational planning. Further, it is a system, of both quantitative and qualitative elements, whose interrelationships need to be made explicit, because even seemingly “soft” concepts such as junior doctor experience and knowledge or motivation have very real “hard” impacts on current and future service provision.

Once modelling began, it became evident that an ideal consulting-context development process that might comprise qualitative analysis, then simple insight model, then a fully detailed model would not be appropriate. The first modelling efforts were actually focussed on representing the full detail of the progression or ‘hard’ elements of the system. This was desirable so that the medical members of the team, who were the direct owners of the problem, could appreciate how the training system would be captured, that it was a true reflection of actual processes, how the junior doctors would eventually progress to being full doctors, and finally how the consultant level interfaces back with the other grades through their training supervision roles. This is consistent with an observation by Winch that in large consulting projects, especially with a disparate client group, a sub-optimal modelling approach is often needed, and models may have to be over-engineered to gain buy-in from all. The modelling process has also helped the team better understand the interactions with tricky but important detail, like doctors who have technically qualified to proceed to higher grades, but who stay at lower levels either while awaiting internal or external vacancies, but also to reflect what is evident a growing trend of some doctors wanting to stay in more junior posts for lifestyle and/or personal reasons. In terms of gaining maximum output in terms of both training and service in a progressively constrained system, use of these mechanisms are likely to become more important in doctor resource management at the hospital.

Experimentation and simulation with the model has also yielded the following insights into the specifics of the problem at hand:

(i) There will be an increased need for supervision of junior doctors due to projected increase in numbers and the increased amount of structured supervised activities enforced by Modernising Medical Careers. On average this is an hour per doctor per week.

(ii) As long as it is assumed that supervision reduces efficiency and amount of service provision on behalf of senior colleagues, meeting the
supervision demands of an increased number of trainees participating in
more of the structured (and supervised activities) under MMC will have a
relatively negative impact on service provision and on relative quality of
training provided by senior colleagues.

(iii) However, it is of highest interest for the hospital to maintain relatively
high quality of training due to the hypothesised high retention rate both
within the foundation years grade (first two years) and between this
foundation grade and the run through training grade (the next grade) at
Derriford. This is because there is a positive relationship between the
quality of training and attractiveness as an employer, as well as future
levels of efficiency and quality of service provision.

Analysis has shown that solutions that increase the training/ service *perception
balance*, without increasing the number of hours needed for supervision, will have
a relatively lower impact on service provision, while compensating for the loss in
relative quality of training provided by senior colleagues who are increasingly
pushed for time. In particular, this should involve looking at better use of slack
time currently observed in the system.

Therefore, to date, work on the model has suggested that there are a certain
number of areas of manoeuvrability for senior Trust management to make
decisions that may help in the reconciliation of service provision and training
standards within the constrained system:

(i) Who provides supervision to junior doctors? Although in practice a large
proportion of informal supervision comes from other junior doctors of
more senior grades, it may be in the interest of the Trust to emphasise
more formally the amount and type of training to be gained from more
senior colleagues of the junior doctor grades. This may help meet some
of the supervision demands at minimum cost to service provision, due to
the greater proportions of this group of staff. In a similar vein,
especially lower grades of junior doctors will gain valuable supervision
from other more experienced staff groups such as senior nurses and
allied health professionals.

(ii) Effect of supervision on efficiency - while this is largely based on the
assumptions made, if service is planned to account for this efficiency
loss due to meeting training needs, then there may be less of a
deviation from the expected level of service. This would require
significant long-term changes in planning and would be a challenge
within the existing target-oriented framework of NHS governance.

(iii) Alternatively, any technology that may compensate for the reduced
efficiency in service provision incurred by supervision may help maintain
service levels. The costs of technology assisted working could be
investigated against the benefits and savings accrued from an increase
in efficiency. (e.g. revised bleep technologies, PDAs for communication
and paperwork)
Better use of empty time for educational purposes - this would increase the training percentage, without necessarily increasing the demands for structured supervision, especially if this time is used for self-directed learning.

In the long run, better quality training for Foundation Years doctors will increase the efficiency in which they are able to provide a service, due to the higher level of skills that they will have acquired. If PHNT is hypothesised to have a relatively high retention rate (as high as 60%) between the foundation years and run-through-training-grade, there is a strong incentive to maintain relatively high quality of training, not only for long term service provision, but also to compensate for increasing demands on junior doctor time to provide supervision to more junior colleagues. This means that an emphasis on meeting all training demands at the crucial foundation stage, while sacrificing some service at this level early on (which could possibly be taken on by other roles or supported by additional means) would benefit career development and service provision at the run-through-training grade and beyond.

Finally, there is currently no provision for the educational value in the participation in service activities. However, unsupervised and independent working can be as important for career and skills development as supervision. If the balance of this is right and appropriate to the skills level of the doctors, the quality of training may further be enhanced. However, this is difficult to model in a quantified model that treats training and service as ends of a continuum.

More detailed results about the scenarios that are being simulated and the results of sensitivity analyses and testing will be published as they become available.

It is clear that system dynamics is a powerful tool for modelling systems, their components, relationships and changes in complex social problems. This paper has set out to highlight some of the ways in which it has been used as a tool amongst other research techniques to analyse and gain understanding into a complex problem faced by hospital trusts in the UK.

References


INTERNATIONAL ENTREPRENEURSHIP: CRITICAL FACTORS FOR SUCCESS

An Empirical Study Of U.K. And Portuguese SMEs In The Textile Sector

Francisco Serra

Abstract

The purpose of the study was to determine which particular organisational (firms) and managerial (decision-makers’) factors contribute to the fast international expansion of firms. The study analysed firms’ resources and capabilities, decision makers’ objective and subjective characteristics in a sample of 167 Portuguese and 165 U.K. firms in the Portuguese and U.K. textile and clothing industry. Statistical analysis using SPSS was carried out to confirm or reject twelve hypotheses. Results confirmed that the factors influencing internationalisation of firms according to the hypothesis suggested by us. The only exceptions are the level of education, age and international experience of decision-makers.

Introduction

The research in small firms internationalisation has always been concentrated in international entrepreneurship (McDougall and Oviatt, 1977). Literature that encompasses the determinants of a firm’s export behaviour is usually divided into two broad levels (Aaby and Slater, 1989; Ford and Leonidou, 1991). They are factors external to the firm and factors that are internal to the firm (Reid, 1981).

- Factors external to the firm include macro-economic, social, physical, cultural, political aspects. Besides, Aaby and Slater (1988) includes industry characteristics.
- Factors internal to the firm (organizational determinants) include structural and behavioural aspects (Leonidou, 1998). They relate to characteristics of the firm and its products as well as attributes of its decision-makers.

In relation to firms’ determinants - Leonidou (1998) has criticised the failure to test simultaneously the full set of organisational factors against the full range of export dimensions. Concerning managerial determinants - Leonidou (1998) emphasised the absence of agreement among researchers as to what constitute a managerial factor in determining exporting, and what specific export dimensions are influenced by management.
In order to fill this gap in the literature, this research attempts to develop an appropriate model, that can explain which international factors are connected with export performance in this specific sector. Simultaneously, try to provide a clear understanding of the main decision-makers’ and firm characteristics that are associated with export performance in the U.K. and Portuguese textile, clothing and knitwear sector. Finally, develop a profile of decision-makers as well as provide information to managements, firms, trade associations, public policy and future researches.

The Portuguese and U.K. textile-clothing and knitwear industry was chosen as the research context for the following reasons: first, the textile, clothing and knitwear industry because of its significance (21%) of the Portuguese economy and due to the structure of this industry. (ICEP, 2000). Second, it is the most representative sector in terms of Portuguese foreign trade. Third, it was our intention to compare what happens in Portugal with another foreign country, in this case the U.K.. Finally, despite being the most representative sector, it is also the most unfavourable sector in the country (Portuguese Economy Minister, 2003).

The present work is structured in the following way: in the next section is introduced the model. Next, we explore the literature review of the firm and decision makers’ characteristics. Then, the procedures connected with the methodology; the principal findings of the present study are introduced. Finally, the principal conclusions are introduced as well as suggestions for future researches.

Model

The theoretical model below was introduced to investigate and research the firms’ and decision-makers’ characteristics. It shows possible relationships among three groups of variables. These are a firm characteristics, the decision-makers characteristics and propensity to export (see fig.1).
**Fig. 1 Model**
Firm’s Characteristics

### Unique Firm Advantages
- Firm size
- Competitive Advantages
  - Quality products
  - Capacity to develop new products
  - Patents held by the firm
  - Quality Control

### Decision-Makers’ Characteristics

#### Objective Characteristics
- Age
- Level of Education
- Foreign language ability
- International Experience
  - Overseas work experience
  - Foreign residence
  - Citizen foreign countries

#### Subject Characteristics
- Perceived risk in exporting
- Perceived cost in exporting
- Perceived profits in exporting
- Resource commitment

Source: Elaborated by the author
The model is used to show the which association of international entrepreneurship to export performance

**Literature Review and Hypotheses**

Several studies were conducted focusing on export activities and identifying factors influencing firm’s behaviour in an export market.

The process of internationalisation, export literature and entry modes have been investigated by a vast number of studies, emphasising the central role played by managers. As we can see the areas referred to above are fundamental in the process of the internationalisation of SMEs firms. Hence, the purpose of this paper is to consider the two categories of resources and capabilities that influence export behaviour:

1. firm-specific resources and capabilities; and
2. decision makers’ characteristics.

In the following sections, we review the extant literature on export determinants and propose research hypotheses according to this double direction.

**Firm-specific resources and capabilities**

**Firm’s size and export behaviour**

A firm’s size is one of the most important variables in exporting, given that a lack of size can be considered a handicap in exporting for small firms. There are a number of significant studies which found a positive relationship between a firm’s size and its propensity to export (Czinkota and Johnston, 1983; Perkett, 1963; Tookey, 1964, 1983; Hunt and al., 1967; Hirsch and Adar, 1974; Reid, 1982; Burton and Schlegelmilch 1987; Tookey, 1964, 1983, Reid, 1982). Conversely, there are studies showing a strong support for variables (size), but do not show how these variable influence export performance (Miesenbock, 1987-88). Others have shown mixed results (Bilkey, 1978) and finally several others argue that there is no meaningful relationship between size and export behaviour (Bilkey and Tesar, 1975; Abdel-Malek, 1978; Czinkota and Johnston, 1983; Cavusgil, 1984; and Reid, 1982).

Hence, in spite of all the controversy between a firm’s size and it’s propensity to export, it is possible to propose that a relationship truly does exist between a firm size and it’s ability to export:

**H1** - Larger firms are more likely to have good propensity to export.

**Competitive Advantages Product**

**Product quality** is undoubtedly an important factor in entering and remaining in international markets (Daniels and Robles, 1982). Furthermore, a study carried out by Kaynak and Erol (1989) revealed that the major strengths of Turkish
manufacturing and trade house export firms included the areas of product quality, price levels, new product, development capability and an export middlemen network.

Nonetheless, besides the quality of the product, other factors must be considered. Hirsch (1970; and McGuinness (1978) emphasised the importance of the “unique” product giving the company a competitive advantage. Next, product strength is another characteristic that is strongly related to export success (Burton and Schlegelmilch, 1987; Cavusgil and Nevin, 1981; Madsen, 1989; McGuinness and Little, 1981; and Michell, 1979). Finally, product adaptation policy is a very important aspect in export marketing strategy. Therefore, it is proposed that firms with these qualities are positively related to export. Thus, this provides the basis to formulate the following statement:

H 2 - Firms with competitive advantages are more likely to have good propensity to export.

Technology

High-technology firms have a higher propensity to become “born global” in nations with small domestic markets than firms in nations with large domestic markets (cad 718 - pl of 15) Technology intensiveness is consistently related to propensity to export according to some researchers (Cavusgil and Nevin, 1998; McGuiness and Little, 1981; Cavusgil, 1984 a; Cooper and Kleinschmidt, 1985; Daniels and Robles, 1982; Joynt, 1982; Aab and Slater, 1989. However, others such Reid (1986), concluded that there is very little relationship between technology and export performance. Whereas, their studies such as that by Christensen et al (1987) revealed no relationship at all. It seems that the evidence on technology and export performance is definitely mixed. Despite, the mixed evidence between technology and export performance, we are in position to advance the following hypothesis:

H 3 - Firms with technological intensiveness are more likely to have a good international performance.

Decision-makers characteristics

A firm’s decision-makers play an important role in the process of internationalisation, as emphasised by Perlmutter (1969), and also by the concept of psychic distance developed by the Uppsala School (Johanson and Wiedersheim-Paul, 1975). In the 1980s, decision-makers’ international orientation was operationalised through psychic distance, management objective characteristics, management subjective characteristics, and managerial attitude about export (Dichtl et al., 1983).

Furthermore, Pteffer (1983) states that managerial objective characteristics are fundamental in explaining the variances of organisational behaviour. On the other hand, some demographic characteristics help explain a firm’s strategic behaviour because they are a reflection of managerial psychological dimensions (Wiersema
and Bantel, 1992). Hence, taking into account the fact that psychological bases are more difficult to measure (Hambrick and Mason, 1984), this study focussed upon the following managerial objective characteristics:

**Education level.**

There appears to be a general consensus in the literature reviewed that the level of education within senior managers is a crucial determinant of export performance (Langston and Teas, 1976; Reid, 1981; Tumbull and Welham, 1985). Furthermore, in an exploratory study of the elements in major decision making amongst a small sample of Canadian firms doing business abroad, Mayer and Flynn (1973), found an over-representation of university graduates among top managers. However, in the portfolio of cross-cultural studies reported by Dicht et al (1990), the results appear to be inconsistent. In spite of these inconsistencies in the studies of Dicht et al., (1990), the number of authors who regard the education of managers as having a positive impact on export performance appears to be strong enough to advance the following hypothesis:

H 4 - Firms with managers, with high levels of education, are more likely to have a good international performance.

**Manager’s age**

Younger decision-makers tend to be more risk oriented, more associated with policies of corporate growth (Child, 1974; Hart and Mellons, 1970), more internationally minded (Pinney, 1970) and react more positively to an export stimulus, have more energy (Caughey and Chetty, 1994), and are more interested in higher earnings and sales than older decision-makers. Therefore, a hypothesis can be advanced that:

H 5 - Firms with younger managers are more likely to have a good propensity to export

**International Experience**

There are many studies which have examined the importance of managers’ international experience on their firms’ international behaviour (Mayer and Flynn, 1973; Brooks and Rosson, 1982; Ganier, 1982; Ogram, 1982; Angelmar and Pras, 1984; Karafakioglu, 1986; Oviatt and McDougall, 1994; Reuber and Fischer, 1997). Several measures of experience have been posited to predict the propensity for exporting. Among these measures is the manager having lived or worked abroad (Mayer and Flynn, 1973; Simmonds and Smith, 1968; Meisenbock, 1988; Reid, 1981) and the managers’ level of past experience in working on developing international markets for a firm. Hence, we may state that:

H 6 - Firms with managers, who have more international experience, are more likely to have a good international performance.
Foreign language proficiency

According to Turnbull and Welham, (1985) this contributes to international business success in a number of ways. It facilitates social contacts, improves communication to and from markets, assists in understanding the ethos and business practices of a market. Furthermore, individual language skills are believed to be central to effective international marketing (Luostarinen, 1980, Robock and Simmonds, 1983; and Root, 1982), as well as facilitating effective planning and control in markets abroad.

This discussion supports the point of view that there is a positive relationship between linguistic abilities and export performance. Hence, it is possible to hypothesise that:

H 7 - Firms with managers, who have a greater ability to speak foreign languages, are more likely to have good international performance.

Decision-makers - Risk Perception

Despite the fact that psychological bases are more difficult to measure (Hambrick and Mason, 1984), they are no less important. Hence, the managerial subjective characteristics influencing firms propensity to export are reviewed below.

Risk perception and the attitude toward risk in export activities have been found to be important factors in explaining export behaviour Cavusgil and Nevin (1981); Roux (1987); Roy and Simpson (1981); Simpson and Kuyawa (1974). The decision-makers in smaller firms often perceive a higher risk in international activities, due to their lack of appropriate information, unfamiliar conditions in markets, complicated domestic and foreign trade regulations, the absence of trained middle managers for exporting, and the lack of financial resources Cavusgil, (1980); Czinkota, (1982), rather than looking at opportunities the international market can present.

From a survey of managers in 120 manufacturing firms, (Simpson and Kujawa, 1974) it was concluded that firms which developed into exporters were less fearful of the risks involved in doing so than firms which did not become exporters. Thus, from the above research conclusions, it is possible to state that:

H 8 - Firms with managers, who have a greater perceptions about the propensity to take risks, are more likely to have good international performance.

Profitability

Internationalisation involves greater distances, more complexity and new factors (currency exchange rates). All these factors lead to a perception of greater uncertainty and consequently, greater risk, which in turn requires compensation through higher degrees of profitability. A study of 120 Tennessee manufacturing firms declared that the management regarded exporting as a means to obtaining higher profits Simpson, (1973). Moreover, the point of view that decision-makers
in exporting firms expect higher potential profits from exporting than those in non-exporting firms was defended by Simpson and Kujawa, (1974; Roy and Simpson, (1980); Witley, (1980), Cavusgil, (1984b); and Burton and Schlegelmilch, (1987). According to classical economic theory, a firm’s probability in exporting tends to vary directly with the profits it’s management expects from exporting. A believe defended by studies carried out Simpson, (1973); Tesar, (1975); Bilkey and Tesar, (1975).

However, other studies contradict the classic economic theory, as in the studies of 497 Danish, Dutch and Israeli manufacturing firms and 21 Tennessee firms Graueda and Dicer, (1973). Thus, we are able to state that:

H 9 - Firms with managers, who have higher perceptions about profitability, are more likely to have good international performance.

Costs

Cost can be measured in different ways, such as money or economic goods, pain, or disutility. For us, we are only interested in financial costs, being fixed and variable. In determining the price of a product, you have the fixed cost plus variable costs plus profits and tax. However, in international trade there are more additional costs, such as: the cost of shipping products, various tariffs (local taxes - for example VAT), administrative costs (ex. acquiring import and export licenses) and sometimes some physical arrangements.

Thus, we are in position to state that:

H 10 - Firms with managers, who are aware of additional costs, are more likely to have good international performance.

Commitment to Exporting

Commitment level has been heavily studied as a predictor of export initiation (Cavusgil 1984a; Bello and Barksdale 1986; Sullivan and Banerschmidt 1987). There is a positive relationship between the commitment level to exporting and export performance (Gromhang and Lorenzen, 1982; Widershein, Finn and Welch, 1987).

All the above studies conclude that there is a positive relationship between management commitment and the propensity to export (Bello and Barksdale, 1986; Cavusgil 1984a; Cavusgil et al., 1979; Gronhaug and Lorenzen, 1982; Kiriplani and MacIntosh, 1980; Rosson and Ford, 1982 and Sullivan and Bauerschmidt, 1987). Hence, we are in the position to conclude the following:

H 11 - Firms with managers, who have higher levels of commitment, are more likely to have good international performance.

Methodology

A mail survey was the method chosen to collect data because it produced fewer biased responses than in face-to-face interviewes (Dillman, 1978). The sample
members were the person responsible for the exporting in the firm or in case this personagem didn’t exist the person responsible for export decisions.

Sample

The Portuguese population for the research was obtained from Portrade 2001 (Portuguese Export and Import Companies). Consisting of 940 firms, with more than 20 and less than 250 employees, from the textile, clothing and knitwear sector.

Procedure

The questionnaire was developed according to the following process: first, all literature concerning the sector (textile, clothing and knitwear) was reviewed, as well as literature about export behaviour. Second, in order to ensure the fitness of purpose, language, wording, sequence and layout of the questionnaire, a three-stage pilot was adopted. The questionnaire was reviewed by several academics experts from different scientific areas, as well as by two expert managers. The reviewed version was sent to the person responsible for exporting in twenty five firms. The last version was e‐mailed directly to the person responsible for exporting in the company.

To encourage participation a stamped addressed envelope was enclosed with the questionnaire, a covering letter and a sponsorship letter. A second wave of questionnaires were sent to all non-respondents two months after the initial e‐mailing.

Following a stratified sampling procedure, the questionnaire was sent to 940 firms in Portugal and 800 in U.K. The percentage of valid response by Portuguese firms was around 27 %, while amongst U.K. firms it was 25%. The response rate obtained in the present study can be considered acceptable if we consider the results obtained in similar studies (Morgan, 1995; Algahtani, 1996; Oklahoma et al, Cad 75; Koh, 1990; And Samill & Walters, 1990).

The T-test was the method statistic employed to measure any potential non‐response bias and was carried out in both countries separately. It was developed according to what is advocated by Armstrong and Overton (1977). No significant differences were found between the early and latest respondents.

Operationalization of Independent Variables

The independent variables were divided into two broad groups. Those associated with a firms’ characteristics, and those associated with a decision-makers’ characteristics.

In the “firm” group (firms’ characteristics), the principal focus was in the following variables: company-size, competitive advantages and technology orientation. Those associated with decision - makers’ characteristics - The decision-makers’ characteristics (second group) were divided into two groups. The first connected
with objective characteristics and the second with the subjective characteristics of
the decision-maker.
Among the objective characteristics were referred the following: age; education;
foreign language ability; international experience were all considered.
Perceived attractiveness in exporting: perceived levels of risk cost, profits and
resources commitment were the subjective variables considered.

Each conceptual variable is measured by multiple indicators. The purpose of this
procedure is for the development of measures to improve the validity and
reliability of the study (Churchill, 1979 and Peter, 1979).

The operationalisation of the dependent variables was through the propensity to
export.
The two most commonly used dimensions are the rate of growth in export sales
and the percentage of total sales accounted for by exports. In our specific case we
considered the percentage of total sales accounted for by export.

Results

Several analytical techniques were applied to the data. firms’ characteristics -
Hence, analysis of variance (Anova) test was used to evaluate the differences in
the level of firm size (H1), competitive advantage (H2), and technology orientation
(H3). However, before running the ANOVA test, Levene’s test of homogeneity of
variances was carried out to establish whether or not the homogeneity assumptions
of ANOVA were violated or not. The results obtained on this test were: H1 (p =
0.311); H2 (p = 0.251); H3 (p = 0.03)

As we can see in the hypothesis H3 (p = 0.03), the homogeneity assumptions
were violated (p < 0.05). Hence, in this hypothesis it was not possible to run the
ANOVA test. Alternatively non-parametric tests were used. On the remaining
hypothesis because the assumptions were not violated. Hence, the results obtained
with ANOVA were: H1 (p = 0.00), H2 (p = 0.00) the p value for homogeneity
of variances was always inferior to 0.05 (p < 0.05). This means that there is a
significant difference between the variances. This is what happened with H1 and
H2. Hence, the Tukey and Scheffe test were chosen to determine which group
differ significantly from the others. The former two tests provided evidence that
there are clearly significant differences among the groups.
H.1 - Tukey test = 0.203 and Scheffe = 0.235. In fact, from these results, i the
Anova test was carried out in both Portugal and the U.K., however the results
come only from Portugal, showing that size is responsible for the propensity to
export. H.2 - Firms that have competitive advantages and technology intensiveness
(H.3) have more propensity to export.

Decision-makers’ objective characteristics

The analysis of variance (Anova) test was used to evaluate the differences in the
ages of decision-makers (H4), the level of education (H5), and the number of
languages spoken (H6). However, before running the ANOVA test, Levene’s test of
homogeneity of variances was conducted out to establish whether or not the
homogeneity assumptions of ANOVA were violated or not. The results obtained on
this test were: H4 (p = 0.918); H5 (p = 0.102); H6 (p = 0.000). As we can see in the hypothesis (H6) the homogeneity assumptions were violated (p < 0.05). Hence, in this hypothesis it was not possible to run ANOVA. Alternatively non-parametric tests were run on the remaining hypothesis because the assumptions carried on the Anova test were not violated. Hence, the results obtained with ANOVA were: H4 (p = 0.274), H5 (p = 0.193).

T-Test or Mann-Whitney were the tests used to carry out the Hypothesis seven (international experience) and eight (nationally decision-maker). However, before running out the T-Test as equally run out the Levenes’ test. The Levenes’ test for H7 has a probability greater than 0.05, in this case 0.375. Hence, the T-test was carried out.

For decision-makers’ nationality (H8), the value obtained for Levenes’ Test was (p = 0.39), substantially superior to 0.05. Relatively to independent sample tests, for equal t value, there is a difference (p < 0.05), in this case (p = 0.000).

From the results obtained we can affirm the following: H4 - age. The Anova test, 0.274 (p > 0.05). We can see that the age of decision-makers does not impact on the propensity to export. H5 - p=0.193. The level of education does not affect their propensity to export. H6 - After running the Tukey and Scheffe tests, we can conclude there are significative differences between those who have the ability to speak foreign languages and those who do not, in terms of their contribution to export. H7 - (p = 0.375). This means, those who have lived abroad are better at propensity to export. H8 - (p = 0.000), although nationality does affect the propensity to export.

Decision-makers’ subjective characteristics

Again, the analysis of variance (Anova) test was used to evaluate the differences in the level of risk perception (H9), perception of costs (H10) and finally profitability (H11). However, before running the ANOVA test, Levenes’ test of homogeneity of variances was used to establish whether or not the homogeneity assumptions of ANOVA were violated or not. The results obtained on this test were: H9 (p = 0.293); H10 (p = 0.058); H11 (p = 0.260). Because the assumptions were not violated, the Anova test was utilised. Hence, the results obtained with ANOVA were: H9 (p = 0.293), H10 (p = 0.289) and H11 (p = 0.153). Due to these results, we can affirm that decision makers’ who have more risk perceptions, a perception of costs and higher perceptions about export profitability do not have more propensity to export.

Principal Components Analyses Method

The variables from planning, promotional international, and involvement in export departments (question 25, 26 and 28) were subjected to a factor analysis with varimax rotation with Kaiser normalisation. Hence, the Kaiser-Meyer-Olkin (KMO) test and Bartlett test of sphericity were carried out initially with 18 variables. Latent Root Criterion was the method chosen to determine the number of factors to be extracted. Five factors were extracted which represent 61% of the total variance explained among the variables. Finally, Varimax orthogonal Rotation was carried out. Finally, all the test were repeated but with only 17 variables. The new results obtained were: KMO (0.753): Bartlett’s (0.000); Latent Root Criterion (
62%); Unrotated Component Matrix (five factors and 17 variables) and finally orthogonal (varimax) rotation confirmed five factors extracted on unrotated solution. The five factors being: Decision-making; planning; public relations; advertising and promotion and human resources.

Conclusions

There are two categories of internal determinants that explain the propensity of firms to export. They are a firms’ characteristics and the objective and subjective decision-makers' characteristics.

A firms’ characteristics are positively related to the firms' size, competitive advantages and technology intensiveness.

Our findings conclude that firm size is a determinant in the propensity to export. Hence, the conclusions arrived at in this research are in accordance with previous research evidence, such as the studies of (Hunt et al. 1967; Hirsch and Adar, 1974; Reid, 1982; Czinkota and Johnston, 1983; Perkett, 1963; Tookey, 1983; and Burton and Schlegelmilch, 1987).

Equally, competitiveness advantages of products are determinants in the propensity to export of these firms. Daniels and Robles, (1982) and Kaynak and Erol (1989) had arrived at the same conclusion.

Regarding technology intensiveness we came to the same conclusion as Cooper and Kleinschmidt (1985); Daniels and Robles (1982); Joynt (1982) and Aab and Slater (1989). It does influence the propensity to export.

Concerning managerial objective characteristics – managers with high levels of education do not have any extra influence in their propensity to export. Neither does a managers' age have not significant influence in the propensity to export as suggested by Manalova et al. (2002). Similar conclusion was reached with variable international experience. However, decision makers who have higher ability to speak foreign languages do not have higher probability to export. These findings are consistent with the works carried out by (Hunt, 1969; Weinrauch et al., 1975; Khaun, 1975; Bilkey, 1978; Reid, 1981; Turnbull and Cunningham, 1981; Turnbull and Welham, 1985; Schlegelmilch, 1986; Roux, 1987; Schlegelmilch and Crook, 1988; Holzmuller and Kasper, 1990 and finally with Enderwick and Akoorie, 1994).

Subjective decision-makers’ characteristics - firms with managers, who have higher perceptions about the propensity to take risks, of the profitability and perceptions about the costs, all have a better propensity to export.

Limitations to this research First of all studies of this kind are extremely rare concerning this area and any kind of generalisation could be potentially misleading. Second, it is extremely difficulty to validate the questions about decision-makers' because they are extremely subjective / difficult. Third, this study is static in nature and it is determined by the absence of a longitudinal analysis. Finally, the conclusions arrived at in this study are limited only to countries of Portugal and U.K.

Direction for future researchers In conceptual terms the framework developed in this study studied a large amount of variables (see model). However, these
variables could be extended to include other areas in future researches. Second, empirically this study could be replicated in other international businesses as well as in other countries. Third, it would be extremely positive if future studies looked at this problem under longitudinal studies. Finally, it would be interesting to study the interaction between management demographics and its attitude, perception and behaviour as suggested by Kundu and Katz (2003).

Bibliography


Simmonds, K. and Smith, H., 1968. The first export order: a marketing innovation. *British Journal of Marketing*, 2, (Summer), 93-100.


PROFESSIONAL FOOTBALL SPONSORSHIP - THE RESEARCH PROJECT

André W. Bühler

Abstract

This research project, which has been kicked off in March 2003 at the University of Plymouth Business School and which is about to be completed three years later, looks at professional football sponsorship from three different perspectives: the clubs’ perspective (sponsorship as an income stream for professional football clubs), the sponsors’ perspective (sponsorship as a marketing tool for companies), and a joint perspective (sponsorship as an inter-organisational relationship between professional football clubs and their sponsors). The English Premier League and the German Bundesliga serve as the context of research owing to their extraordinary reputation in commercial terms. A combination of qualitative and quantitative research methods is used in order to answer the research questions which derived from an extensive literature review. The findings of this research project provide both theoretical and practical implications for professional football clubs, sponsoring companies and also contribute to the literature on sports marketing and relationship marketing.

1. Introduction

In March 2003 a PhD-research project was kicked off at the University of Plymouth Business School under the working title ‘A Comparative Investigation of Sponsorship and Related Commercial Aspects of Professional Football in England and Germany’. Nearly three years later, the research project is about to be completed under the title “Professional Football Sponsorship in the English Premier League and the German Bundesliga”. This research project examines professional football sponsorship from three different perspectives:

1. Sponsorship as an income stream for professional football clubs
2. Sponsorship as a marketing tool for companies
3. Sponsorship as an inter-organisational relationship between professional football clubs and their sponsors

The English Premier League as well as the German Bundesliga serve as the context of research owing to their extraordinary reputation in commercial terms compared to other European football leagues. Various reasons speak well for choosing an Anglo-German perspective: First, football is regarded as the national game, both in
England and in Germany (Pepels, 2001; Key Note, 2002). Second, the English Premier League and the German Bundesliga are not only ranked within the top five European football leagues from a point of view of performance on the field\(^1\), but are also widely considered to be the most businesslike football leagues in the world (Süßmilch, 2002; Rachmann, 2002). For example, the German licensing model is widely recognised as a beneficial model and acts as a prototype for the UEFA-licensing\(^2\) scheme. Furthermore, Deloitte (2005, p. 11) notes that ‘the “big five” leagues differ widely in their operating performance with only the English Premiership and German Bundesliga clubs recording profits in 2003/04.’ In addition, both the English and the German top league record the largest growth rates (+14% and +12% from 1999/2000 to 2003/04 respectively) in comparison to other European football leagues according to Deloitte (2005). The English Premier League qualifies for this research in view of the fact that the English clubs are the European benchmark in terms of television income, merchandising revenues and matchday income. The German Bundesliga qualifies in view of the fact German clubs generate the largest income from sponsorship. In addition, the sponsorship market in Germany is highly developed.

A combined approach of qualitative and quantitative research methods served as the methodological foundation of this research. In a qualitative research phase, which took place from September 2003 till August 2004, seventeen representatives of English Premier League and German Bundesliga clubs and their sponsors, as well as English and German sponsorship experts, were interviewed. The insights of these face-to-face in-depth interviews plus findings of the literature review led to the generation of seven principal research propositions (PRPs) and six hypotheses, covering the three perspectives of sponsorship in question. Following a comprehensive content analysis of more than 500 websites and 106 Premier League and Bundesliga games on video tape in order to identify as many sponsors as possible, a quantitative questionnaire survey was carried out in April/May 2005. Non-parametric tests were carried out to test the PRPs, whereas factor analyses as well as a multiple regression analysis were used in order to test the hypotheses.

This paper introduces the above research project in greater detail. First, a brief review of the literature and a rational for choosing the three perspectives of sponsorship in question will be provided. The paper then moves on to the structure of the thesis before dealing with the methodology on which the primary research is based on in greater detail. The paper concludes with some findings of the primary research process.

---

\(^1\) According to the latest UEFA-ranking in 2005, England takes second and Germany takes fifth place. The basis for the UEFA rankings is the performance of teams in the European Cups during a five-year period according to www.uefa.com

\(^2\) The UEFA licensing regulations define the minimum quality of standards in five main criteria categories as follows: sporting, infrastructure, personnel and administration, legal and financial. These criteria must be fulfilled in order for a club to be admitted to any of the UEFA Club competitions as from the 2004-05 season onwards. In addition, these national regulations regulate the process to be applied by the licensor when assessing their respective clubs.

\(^3\) i.e. the English Premier League, the French Ligue 1, the German Bundesliga, the Italian Serie A, and the Spanish Primera Liga.
2. Sponsorship as a subject of research
This section provides some reasons for choosing the three perspectives of football sponsorship under scrutiny. First, literature and previous studies on sponsorship in general will be briefly reviewed before dealing with the literature on (professional) football sponsorship.

2.1 Sponsorship in general
There are currently more than 300 books in English and around 130 books in German on sponsorship available. Within the last years, numerous research papers have been written on sponsorship. A search enquiry at the Deutsche Bibliothek revealed the publication of 104 doctoral theses on sponsorship between 1987 and 2005 in Germany, covering various areas. For example, some of them deal with legal aspects of sponsorship (e.g. Irle, 2001; Wegner, 2001; Höltkemeier, 2004), some with psychological aspects (e.g. Gloger, 1998; Wagner, 1993; Erdtmann, 1988) and some others cover marketing aspects (e.g. Dinkel, 2002a; Vogt, 2000). Surprisingly, the British Education Index records only three published theses on sponsorship. However, although sponsorship has been researched in the areas of arts, media or education, most studies focus on sports sponsorship.

Walliser (2003, p. 6) analysed and compared more than 230 papers on sponsorship. He notes that 'sponsorship may be one of very few areas which has attracted more academic interest in Europe – particularly in Ireland, France and Germany – than in North America or other parts of the world.' Sponsorship as a subject of research was discovered at the end of the 1970s or beginning of the 1980s respectively. The first books and reports about sponsorship have been published in the United States, followed by some German publications, primarily the works of Bruhn (1987) or Hermanns (1987) which are still the core books on sponsorship in Germany. The same is true for Sleight (1989) in the UK. In the early editions of the ‘marketing bibles’ sponsorship was just a peripheral phenomenon. With the increasing meaning of sponsorship and the transformation from ‘a small-scale activity in a limited number of industrialised countries to a major global industry’, research interest in sponsorship increased as well (Walliser, 2003, p. 5). Most research papers deal with the measurement of sponsorship effects, relate to managerial aspects or the nature of sponsorship, or investigate sponsorship strategies and counter-strategies according to Walliser (2003). Olkkonen, Tikkanen and Alajoutsijärvi (2000, p. 13) argue that ‘the bulk of existing sponsorship research is very “managerially” oriented, with strong emphasis on the sponsoring company’s viewpoint’.

Dinkel (2002b) also notes that most research on sponsorship has been done from the sponsor’s point of view and that only a few consider the other side of the deals, the sponsees. Therefore, this research project takes both sides of the sponsorship dyad into consideration by examining professional football sponsorship as an income stream for professional football

4 Revealed by a search inquiry of the English and German Amazon website in December 2005.
clubs, as a marketing tool for sponsors and finally as an inter-organisational relationship between sponsors and sponsees.

2.2. (Professional) football sponsorship

Although professional football is used as a popular example in a lot of papers on general sponsorship, professional football sponsorship as a subject of research is not as common as assumed. This subsection looks at to what extent previous research examined (professional) football sponsorship as an income stream, marketing tool and inter-organisational relationship.

Research on sponsorship as an income stream for professional football clubs is mainly limited to reports dealing with the finance of football, for example reports from Deloitte (2004, 2005), Key Note (2002), the WGZ-Bank (2002) or Ernst & Young (2004) as mentioned above. In addition, reports published by the Deutsche Fußball Liga (DFL, 2004) or the FA Premier League (2004) also deal with sponsorship as an income source. Most of these reports describe the current situation of professional football in commercial terms by providing figures relating to the key income streams of professional football clubs but neglect an in-depth analysis of sponsorship and other key income streams and how they relate to each other. Therefore this study will examine sponsorship as an income stream for professional football clubs in greater detail.

With regard to football sponsorship as a marketing tool, more studies can be found. Sengle (1989), for example, describes sponsorship as a marketing tool on the basis of German football. The benefit of her research is the fact that she was one of the first researchers who analysed the football sponsorship situation. The limitation, however, is the solely focus on shirt sponsorship. In addition, the sponsorship market as well as the importance of sponsorship for professional football clubs has been changed dramatically since the publication of her PhD-thesis. However, a more recent paper comes from Köster (2003), who describes sponsorship in the context of professional football in Germany. She may take both the clubs’ and the sponsors’ perspective into consideration, but mainly focus on the management process of the sponsorship activities. Another study on issues relating to football sponsorship as a marketing tool (such as objectives of football sponsors or evaluation of sponsorships) comes form Thwaites (1995) who examined sponsorship programmes in English professional football. Other examples come from Rosson (2001) and Wilcox, Andrews and Longmuir (2001) who used case studies in order to assess football sponsorship as a marketing tool. In addition, some market research reports, such as Sponsor Visions (Pilot Group, 2004, 2005) publish survey results relating to objectives of football sponsors. However, it is quite surprising that professional football sponsorship as a marketing tool is not as popular as a subject of research given the popularity of the sport and the fact that most papers on sponsorship deal with managerial issues and mainly cover the sponsor’s point of view. In view of the limited papers on the subject it is felt that further research is necessary. Therefore, this research project examines professional football sponsorship as a marketing tool in greater detail.
Hardly any research has been done on professional football sponsorship as an inter-organisational relationship. Few exceptions come from Bembennek and Meier (2003) and Farely and Quester (2003, 2005). However, both exceptions include major limitations. Bembennek and Meier, one the one hand, may investigate the relationship between sponsors and sponsees in the first and second football Bundesliga in Germany, but also included handball and basketball in their study. Therefore, the results of their study cannot be referred to professional football sponsorship alone. Farely and Quester, on the other hand, may examine the football sponsor – football club relationship as a business-to-business relationship, but in the context of Australian Rules Football. Therefore, the only notable exception comes from Chadwick (2004) who looked at determinants of commitment in the English professional football club/shirt sponsorship dyad in his PhD-thesis. Consequently, further research is needed in the area of professional football club – sponsor relationships.

In summary, research on sponsorship as an income stream for professional football clubs, as a marketing tool for companies and as an inter-organisational relationship is very limited despite the popularity of both the sport and the medium. It is therefore felt necessary to examine all three perspectives in greater detail by including both sides of the sponsorship dyad and therefore contribute to existing knowledge. As a matter of fact, no research to date has ever incorporated all three perspectives and therefore this study is also seen as a foundation for further research in the area of professional football sponsorship.

3 The nature and structure of the thesis
The nature of the research subject is not only reflected in the numbers of chapters (this thesis on professional football sponsorship is composed of eleven chapters - just as every football team is composed of eleven players) but also in the structure of the thesis. The three perspectives under scrutiny run through the thesis such as a main thread. Figure 1 illustrates the structure of the thesis.
The opening chapter introduces the nature of the thesis and provides a general review of the literature on football and sponsorship. It also deals with sports marketing as a research subject and therefore leads over to the second chapter.

Professional football sponsorship belongs to the research area of sports marketing and therefore this research is a sports marketing thesis. Sports marketing differs from traditional forms of marketing in numerous ways. For example, the sports industry, the sports product and the sports consumers have unique characteristics (van Heerden, 2001; Bühler, 2005) Therefore, Chapter 2 of the thesis introduces the context of professional football sponsorship, i.e. the business of football. First, a number of reasons are be presented why football can be regarded as a business.
Then the markets and the market players are described before an examination of the unique characteristics of the football business takes place. The chapter then moves on to a description of the product provided by professional football clubs. The third section of Chapter 2 deals with the customers of professional football clubs with special emphasis on the supporters.

Chapter 3 examines professional football sponsorship from the football clubs’ point of view by looking at the nature of sponsorship as an income stream for professional football clubs. The chapter is divided into two parts. The first part looks at the key income streams for professional football clubs in general and investigates television income, matchday income and merchandising revenues in greater detail. Each subsection gives a general overview of the importance of the respective income stream before looking at the specific situation in the English Premier League on the one hand, and the German Bundesliga on the other. Each subsection concludes with a comparison between both leagues and an outlook regarding the further grow of the three income streams in question. The second part focus on sponsorship as the fourth main income stream of professional football clubs. First, a general overview of the importance of sponsorship as well as an analysis of football clubs’ sponsor pools is presented. The second part then moves on to an in-depth analysis of the sponsorship situation in the English Premier League and the German Bundesliga, including subsections on shirt sponsors, kit suppliers, commercial partners and smaller sponsors. The information provided is mainly based on a comprehensive content analysis of more than 500 clubs’ and sponsors’ websites as well as 106 televised football games. The section on sponsorship in English and German football concludes with a comparison between the English Premier League and the German Bundesliga. The next section then compares sponsorship with the other key income streams of professional football clubs and also explains how all four revenue sources interrelate.

Chapter 4 examines professional football sponsorship from the football sponsors’ point of view by dealing with (professional football) sponsorship as a marketing tool for companies. This chapter, too, is divided into two parts. The first part investigates the nature of sponsorship in general terms starting with a review and critical appreciation of previous definitions of sponsorship on the one hand and sports sponsorship on the other hand. The various types of sponsorship and the role of sponsorship within the marketing mix are also explained. The second part of Chapter 4 then explicitly focuses on professional football sponsorship. In view of the fact that a specific definition of professional football sponsorship does not exist to date, the following definition was developed:

*Professional football sponsorship is a business-related partnership between a sponsor and a sponsee based on reciprocity in the context of the football business. The sponsor provides financial or non-financial resources directly to the sponsee and receives a predefined service in return in order to fulfil various sponsorship objectives.*

The chapter then moves on with an examination of the question why companies invest in football sponsorship. The final sections of this second part then deal with
objectives of football sponsors and advantages as well as disadvantages of professional football sponsorship.

Chapter 5 brings both sides of the sponsorship dyad together by examining professional football sponsorship as an inter-organisational relationship. The chapter starts with reasons why professional football sponsorship should be examined from a relational perspective. The second section then describes the rise of relationship marketing, the nature of business-to-business relationships and finally conceptualises relationship quality in general terms and in the context of sponsorship relationships. The third section of Chapter 5 finally focus on the relationship between professional football clubs and their sponsors by dealing with the ‘perfect’ relationship between sponsors and sponsees and describing the ‘real’ picture in English and German professional football sponsorship.

Chapter 2 - 5 all conclude with implications for the primary research phase and therefore provide the following research questions in five main areas:

- **Football as a business**
  - What are the main characteristics of the football business?
  - Is football a ‘big’ business or not?

- **Sponsorship as an income stream for professional football clubs**
  - What are the reasons for the differences in sponsorship between the English Premier League and the German Bundesliga?
  - How could the gap in sponsorship between the English Premier League and the German Bundesliga progress according to the protagonists?
  - How important is sponsorship as an income stream for professional football clubs?
  - How could the importance of sponsorship as an income stream for professional football clubs develop?

- **Football sponsorship as a marketing tool**
  - What kind of objectives do football sponsors have?
  - Why do companies invest in football sponsorship?
  - What are the reasons for successful sponsorship deals?
  - What are the current trends in the area of football sponsorship?
  - What can be improved in the area of football sponsorship?
  - Can sponsorship be measured, and what do companies actually measure?

- **Relationship between clubs and sponsors**
  - How would an ideal relationship between clubs and sponsors look?
  - How do clubs and sponsors actually perceive their relationship?
  - What are the dimensions of relationships?

- **Anglo-German differences**
  - What are the general differences between the English Premier League and the German Bundesliga in commercial terms?
Chapter 6 kicks off the primary research phase by introducing the methodology on which the primary research is based on. The chapter starts with a discussion of various research methods for business researchers, followed by the rational for choosing the appropriate research strategy for this study, i.e. a combination of qualitative research and quantitative research methods. Then the data collection methods are described in greater detail. Finally, special emphasis is addressed to the issue of anonymity and confidentiality in the fourth section.

Chapter 7 presents the results of the qualitative research phase in four sections. The nature of qualitative in-depth interviews is to generate an understanding of the research subject in question. Therefore, crucial statement of clubs’ and sponsors’ representatives as well as sponsorship experts have been compared and contrasted. The first section deals with comments and statements made about the business of football in the one hand and about sponsorship as an income stream for professional football clubs on the other hand. The second section analyses the qualitative data relating to professional football sponsorship as a marketing tool. The third section then presents statements referring to the third perspective of sponsorship under scrutiny, namely the relationship between professional football clubs and their sponsors. Finally, the fourth section highlights Anglo-German differences.

Chapter 8 discusses the findings of the qualitative research phase and consequently generates research propositions and hypotheses for the quantitative research phase. Hypotheses refer to statements where a correlation between two variables is drawn, while PRPs relate to more general issues and assumptions which primarily describe phenomena and/or situations. The differentiation makes sense when it comes to analysis. Whereas the principal research propositions are mainly described by means of frequency tables or cross-tabs, hypotheses are tested in a more sophisticated way using sound statistical techniques such as multivariate analysis methods. Therefore, the qualitative data has been linked with the findings of the first five chapters in order to build the following principle research propositions (PRPs) and hypotheses (Table 1).
**PRINCIPAL RESEARCH PROPOSITIONS**

PRP-1: *The Premier League and Bundesliga clubs rate football as a big business in terms of public perception and media coverage, but as a small business in terms of turnover.*

PRP-2: *German clubs rate the current importance of sponsorship more highly than their English counterparts, whereas English Premier League clubs rate the future importance of sponsorship more highly than German Bundesliga clubs.*

PRP-3: *Clubs see sponsorship both as an income stream and as an opportunity to build the brand of the club and to network with their sponsors.*

PRP-4: *Commercial reasons for companies to go into football sponsorship prevail over personal reasons.*

PRP-5: *The World Cup 2006 is an important reason for German sponsors to sponsor Bundesliga clubs, but does not play an important motivating role for sponsors in England.*

PRP-6: *Brand awareness and image transfer prevail over other sponsorship objectives.*

PRP-7: *The majority of clubs/sponsors perceive relationship quality as being important for the success of their sponsorship, but then only the minority of sponsors/clubs measure or even manage it.*

**HYPOTHESES**

H-1: *Commitment positively influences the quality of the relationship between professional football clubs and their sponsors.*

H-2: *Trust positively influences the quality of the relationship between professional football clubs and their sponsors.*

H-3: *Satisfaction positively influences the quality of the relationship between professional football clubs and their sponsors.*

H-4: *Mutual understanding positively influences the quality of the relationship between professional football clubs and their sponsors.*

H-5: *Cooperation between both parties positively influences the quality of the relationship between professional football clubs and their sponsors.*

H-6: *Communication positively influences the quality of the relationship between professional*
Table 1: Principal Research Propositions and Hypotheses of the Research

Chapter 9 presents the results of the quantitative questionnaire survey according to the three perspectives of sponsorship under scrutiny and the respective PRPs and/or hypotheses. Chapter 9 also provides a description of statistical techniques used for analysis. Depending on the results, PRPs and hypotheses were accepted or rejected. The final section then summarises the quantitative results.

Chapter 10 discusses the results of the quantitative questionnaire survey and completes triangulation by linking, comparing and contrasting the quantitative findings with qualitative findings and findings generated from the literature review.

Chapter 11 finally summarises the findings of this study, draws conclusions, and provides implications for professional football clubs, for sponsors as well as for further research. Limitations of this research are also addressed.

4 The methodology
The above research questions were generated from the first five chapters which were based on an extensive literature review including previous studies and reports. After discussing and evaluating various research methods and data collection techniques the decision was made to use a combination of qualitative and quantitative research methods as illustrated by Figure 2.
I. LITERATURE REVIEW
of previous studies, reports, books, newspaper and magazine articles in order to identify possible issues for interview phase.

Results serve as the basis for describing the sponsorship situation in the English Premier League and the German Bundesliga during the 2004/2005 season.

II. QUALITATIVE RESEARCH
semi-structured face-to-face in-depth interviews

Collection of ideas and opinions reflecting sponsorship and related commercial aspects of professional football in England and Germany

building up hypotheses & principle research propositions (PRPs)

6 clubs’ representatives
6 sponsorship specialists
5 sponsors’ representatives

III. CONTENT ANALYSIS
of 106 televised English Premier League and German Bundesliga games and more than 500 websites in order to get a broader picture of the sponsorship situation in the English Premier League and the German Bundesliga and to generate a sample and provide addresses for the questionnaire survey

IV. QUANTITATIVE RESEARCH
Structured questionnaire survey
cross-sectional snapshot 2004/2005 season

20 English Premier League clubs
18 Bundesliga clubs
163 sponsors Premiere League
307 sponsors Bundesliga

confirmation/rejection of hypotheses and PRPs + subsequent analysis

Figure 2: graphic representation of the data collection process
In order to triangulate data, the findings of the qualitative and quantitative research phase as well as the findings of the literature review are compared as illustrated in Figure 3.

![Figure 3: triangulation of findings](image)

### 5 Summary of the findings

This section presents the main findings of the qualitative and quantitative research. Table 2 summarises the findings of the qualitative in-depth interviews with representatives of English Premier League and German Bundesliga clubs, sponsors of these clubs and sponsorship experts in both countries.

<table>
<thead>
<tr>
<th>Sponsorship as an income source of the football business</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Football as a Business</strong></td>
</tr>
<tr>
<td><em>Can football be seen as a business?</em></td>
</tr>
<tr>
<td>→ overall agreement that football is a business nowadays</td>
</tr>
<tr>
<td>→ in terms of turnover it is not a big business</td>
</tr>
<tr>
<td>→ football is a part of the entertainment business with a highly emotionalised product</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>What are the main characteristics of the football business?</th>
</tr>
</thead>
<tbody>
<tr>
<td>→ It is a weird business</td>
</tr>
<tr>
<td>→ In terms of perception it is a big business. The main characteristic is the public</td>
</tr>
</tbody>
</table>
interest and the media attraction
- the role of the fans has been named as another characteristic
- sponsors and experts emphasized the negative aspects of the business such as unhealthy financing, unprofessional and insufficient management structures
- football clubs have relatively small marketing budgets compared to other companies

Sponsorship as an income stream for professional football clubs
How important is sponsorship for the clubs?
- sponsorship is different at any club. The sponsorship levels differ widely between big and smaller clubs because a lot of companies want to be associated with the big companies
- difference between England and Germany in terms of significance: in England third income stream, in Germany second income stream
- Sponsorship is important because it can help to increase the image of the club or to grow the club international. The nature of sponsorship as a revenue stream is important since it is a high-margin and pre-paid income stream

What is the future potential of sponsorship?
- the overall outlook is optimistic with one exception who assess a consolidation of sponsorship on a lower level
- desire of Premier League clubs to increase their sponsorship revenues

Football sponsorship as a marketing tool
What kind of objectives do sponsors have?
- objectives differ from sponsor to sponsor. Sponsors tend to have multiple objectives
- brand awareness and image transfer are named most frequently. Other objectives are staff motivation, improving business links or meeting the social responsibility
- economical objectives become more and more important

Why do companies invest in football sponsorship?
- football provides passion and image
- football provides wide audiences and different target groups
- football delivers opportunities which help companies to reach their objectives on a national or international level

What are the reasons for successful sponsorship deals?
- linking sponsorship with other promotional tools, which is also a question of budget
- concentrating on fans and having a reason for being there
- looking for scandal free clubs

What are the current trends in the area of football sponsorship?
- a shift from sweetheart deals to business/related deals
- Sponsorship as a network opportunity
- long-term partnerships and joint communication campaigns of clubs and sponsors
What can be improved in the area of football sponsorship?
- getting rid of the clutter and providing more exclusivity
- enabling more networking between sponsors
- pushing social sponsorship as a part of football sponsorship
- solving the problem of player access
- being more imaginative

Can sponsorship be measured and what do companies measure?
- sponsorship is measurable to some extent
- sponsors measure above all media exposure and brand awareness

Sponsorship as a relationship between professional football clubs and their sponsors

Relationship between clubs and sponsors

How would an ideal relationship between clubs and sponsors look?
- a mutual business-related partnership, where both clubs and sponsors benefit from

How do clubs and sponsors actually perceive their relationship?
- both use positive adjectives to describe their relationship. Sponsors use the term ‘partnership’ more often
- both emphasize the business dimension of their relationship
- the length of the sponsorship and the loyalty of sponsors seem to be a determinant of the relationship quality

Which factors contribute to relationship in the football sponsorship dyad?
- trust and commitment as established RQ dimensions have been mentioned.
- satisfaction as another established RQ dimension has not been mentioned.
- with communication, mutual understanding and cooperation three other dimensions have been revealed as being important for the relationship quality of the sponsorship dyad.

Anglo-German differences

What are the general differences between the English Premier League and the German Bundesliga?
- English Premier League clubs publish data and figures more open than their German counterparts
- English Premier League clubs benefit more from globalisation than the German Bundesliga clubs
- The difference in sponsorship terms is reflected in its importance as an income stream

What are the reasons for the differences in sponsorship between the English Premier League and the German Bundesliga clubs?
- German Bundesliga clubs depend stronger on revenues from sponsorship than the English Premier League clubs
The World Cup 2006 in Germany is a real driver for sponsorship in the Bundesliga. A couple of other reasons and assumptions have been identified, such as the size and attractiveness of the German market and the clutter in English Premier League stadiums.

How will the difference in sponsorship between the English Premier League and the German Bundesliga progress?

The gap will be closed because the English Premier League clubs are catching up and there might be a consolidation of sponsorship revenues in Germany.

Table 2: findings of the qualitative interviews

Table 3 summarises the results of the quantitative questionnaire survey relating to the principle research propositions and hypotheses.

<table>
<thead>
<tr>
<th>PRPs hypotheses</th>
<th>Statistical method(s) used</th>
<th>Results</th>
<th>conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRP-1</td>
<td>Descriptive statistics</td>
<td>The participating clubs perceive football as a big business in terms of public perception and media coverage but only a medium-sized business in terms of clubs’ annual turnover.</td>
<td>PRP-1 was partially confirmed.</td>
</tr>
<tr>
<td>PRP-2</td>
<td>Descriptive statistics, Mann-Whitney U test, Kolmogorov-Smirnov Z test</td>
<td>English clubs rate the current importance of sponsorship as an income stream for their club as well as the future importance of sponsorship revenues significantly more highly than German clubs.</td>
<td>PRP-2 was partially confirmed.</td>
</tr>
<tr>
<td>PRP-3</td>
<td>Descriptive statistics, Mann-Whitney U test, Kolmogorov-Smirnov Z test</td>
<td>The participating clubs appreciated the importance of sponsorship as brand building and networking opportunity although sponsorship is mainly perceived as an important source of income.</td>
<td>PRP-3 was confirmed.</td>
</tr>
<tr>
<td>PRP-4</td>
<td>Descriptive statistics, Mann-Whitney U test, Kolmogorov-Smirnov Z test</td>
<td>The results show that the vast majority of English and German sponsors sponsor their respective clubs for pure commercial reasons. Only a minority admitted that personal reasons also play a role.</td>
<td>PRP-4 was confirmed.</td>
</tr>
<tr>
<td>PRP-5</td>
<td>Descriptive statistics, Mann-Whitney U test, Kolmogorov-Smirnov Z test</td>
<td>The forthcoming World Cup 2006 in Germany plays an important role only for the minority of sponsors, although a statistically significant difference between English and German statements exists.</td>
<td>PRP-5 was partially confirmed.</td>
</tr>
<tr>
<td>PRP-6</td>
<td>Descriptive statistics, Mann-Whitney U test</td>
<td>Objectives relating to awareness and image are placed at the top of the</td>
<td>PRP-6 was confirmed.</td>
</tr>
</tbody>
</table>
### Table 3: Summary of findings regarding the principle research propositions and hypotheses

<table>
<thead>
<tr>
<th>Proposition (PRP)</th>
<th>Test(s)</th>
<th>Description</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRP-7</strong></td>
<td>Kolmogorov-Smirnov Z test, Descriptive statistics, Kruskal-Wallis test, Mann-Whitney U test</td>
<td>Objective ranking indicating a clear preference.</td>
<td>PRP-7 was partially confirmed</td>
</tr>
<tr>
<td><strong>H-1</strong></td>
<td></td>
<td>Commitment does not make a significant unique contribution to the prediction of relationship quality.</td>
<td>H-1 was rejected</td>
</tr>
<tr>
<td><strong>H-2</strong></td>
<td></td>
<td>Trust makes a significant unique contribution to the prediction of relationship quality.</td>
<td>H-2 was confirmed</td>
</tr>
<tr>
<td><strong>H-3</strong></td>
<td>Standard Multiple Regression Analysis, Hierarchical Multiple Regression Analysis (Validation)</td>
<td>Satisfaction makes a significant unique contribution to the prediction of relationship quality.</td>
<td>H-3 was confirmed</td>
</tr>
<tr>
<td><strong>H-4</strong></td>
<td></td>
<td>Mutual understanding does not make a significant unique contribution to the prediction of relationship quality.</td>
<td>H-4 was rejected</td>
</tr>
<tr>
<td><strong>H-5</strong></td>
<td></td>
<td>Cooperation makes a significant unique contribution to the prediction of relationship quality.</td>
<td>H-5 was confirmed</td>
</tr>
<tr>
<td><strong>H-6</strong></td>
<td></td>
<td>Communication does not make a significant unique contribution to the prediction of relationship quality.</td>
<td>H-5 was rejected</td>
</tr>
</tbody>
</table>

The results relating to the hypothesis testing were quite surprising with three confirmed hypotheses and three hypotheses which had to be rejected. Therefore, trust, satisfaction and cooperation have a positive impact on relationship quality, whereas commitment, mutual understanding and communication seem not to have any impact on the relationship between professional football clubs and their sponsors in the English Premier League and the German Bundesliga. As these findings contradict results of previous studies and findings of the qualitative interviews, the decision was made to investigate further by using a principal component analysis in order to identify underlying dimensions. Indeed, three factors were revealed incorporating all six dimensions. The factors which were labelled as ‘relationship compatibility’, ‘long-term perspective’ and ‘collaborative behaviour’ all make a statistically significant contribution to relationship quality as confirmed by a second multiple regression analysis. Finally, all statistical tests were rerun for English and German responses separately in order to detect any statistical differences between them.
The key findings of the quantitative questionnaire survey are as follows:

- Premier League and Bundesliga clubs perceive football as ‘big business’ in terms of public perception and media coverage, but only as ‘medium-sized’ business in terms of clubs’ annual turnover. This qualifies the common opinion that football was ‘big business’ in purely financial terms.

- Concerning sponsorship as an income stream, English Premier League clubs rate its current importance more highly than their German counterparts and expect a significant increase in the foreseeable future whereas German Bundesliga clubs doubt that the importance of sponsorship as an income stream will significantly increase in the next five years.

- The vast majority of English and German sponsors are into football sponsorship for pure commercial reasons. Only a handful sponsors stated that personal reasons play a role in their decision to sponsor a football club.

- ‘Increasing public awareness’ and ‘enhancing the image of the company/brand’ are the top commercial objectives of English and German sponsors, followed by various other commercial objectives.

- The majority of clubs and sponsors think that maintaining a good relationship quality is important for the success of the sponsorship as a whole. Consequently most of the clubs and sponsors manage their relationship proactively. In contrast, only the minority evaluates the quality of relationship between them and their sponsorship partner.

- A good relationship quality between professional football clubs and their respective sponsors is mainly determined by three factors: ‘relationship compatibility’ (involving a sense of understanding and fairness in dealing with each other), ‘long-term perspective’ (involving mutual commitment) and ‘collaborative behaviour’ (involving cooperation and communication).

6 Summary
This research examines professional football sponsorship from three different perspectives and therefore contrasts strongly with other research in this area. In order to answer the various research questions a sound methodology was incorporated, making use of qualitative and quantitative research methods. The results of this research contribute to existing knowledge by confirming the results of previous studies on the one hand and adding new insights on the other hand. The findings provide both theoretical and managerial implications and therefore serve as a sound empirical foundation for professional football clubs, sponsoring companies and sports marketing academics.
References


Ernst & Young (2004) *Bälle, Tore und Finanzen: Wege aus dem finanziellen Abseits*, Essen, Ernst & Young AG


WGZ-Bank (2002) *FC Euro AG, Düsseldorf, WGZ-Bank eG*

PROSPECTS FOR THE DEVELOPMENT OF INTERNATIONAL LOGISTICS IN NORTH KOREA

Kelvin Pang
Centre for International Shipping and Logistics, University of Plymouth Business School

Abstract

North Korea is one of the last remaining centrally-planned economies but currently has little interaction with the world economy and minimal logistics infrastructure and activities. However, there are prospects for the future development of international logistics and transport infrastructure. With impending reforms and reunification, regional stability will improve and North Korea offers an affordable labour force with a market of 24 million people, abundant underdeveloped natural resources and the potential to become a strategic distribution location in Northeast Asia. This paper presents the ongoing research into the role of North on transport and logistics development in the Korean Peninsula. The first stage of the research incorporates a comprehensive review of academic and trade literature. In the second stage, a conceptual model is developed and tested using a range of mixed methods. Finally, a scenario model will be created to derive composite policies for transport and logistics development in North Korea.

1. Introduction

This paper forms part of the ongoing research into the role of North Korea on transport and logistics development in the Korean Peninsula. North Korea is one of the last remaining centrally-planned economies but currently has little interaction with the world economy and minimal logistics infrastructure and activities. Transportation is considered not as part of the social infrastructure, but as an element of production, the purpose of which is to take care of the demand derived from other economic activities. Transport problems include deficiencies in infrastructure, poor maintenance, inefficient institutional structures and imbalances in modal shares (Oh, 2001). Recently, North Korea embarked on a series of reforms which includes the opening up of the country which led to an increase in inter-Korean changes. A historic summit in June 2000 between leaders of both Koreas resulted in landmark changes in the Korean Peninsula that might lead to potential reunification. The Gaesong Industrial Complex, which houses many South Korean companies, was established together with the ongoing reconstruction of the Trans-Korean Railways (TKR). Analysis of North Korea has been confined to political and economical aspects. This study has particular interest because it is unquestionably unique, in an area where there has been little if any previous research, and of topical importance since transport and logistics is regarded as one of the key factors in economic development of the country.
The first stage of the research incorporates a comprehensive review of academic and trade literature both within and external to North Korea to identify key factors and themes. In the second stage, a conceptual model is developed and tested using a range of mixed methods, namely, (1) qualitative interviews to gather opinions of North Korea experts; (2) quantitative surveys with South Korean SMEs who are the major investors in North Korea. Finally, these expert opinions will be used as inputs to create a scenario model to derive composite policies for transport and logistics development in North Korea.

2. **Overview of North Korea**

North Korea is strategically located in Northeast Asia bordering South Korea, China and the Russian Far East. North Korea’s topography underpins transport and logistical consideration creating difficulties for an efficient transport network which is centred on roads and railways. North Korea’s transport network is concentrated on the east and west coast consisting of single-tracked railways and unpaved roads. This prevents an effective and efficient flow of transport connectivity and accessibility.

The lack of data has prevented in-depth research into North Korea. It is important to point out that research into North Korea has tended to focus on the political and economic aspects. Due to the nature of the country, very limited research has been conducted into transport and logistical developments in the country. It is imperative to mention that there has been more descriptive than empirical research. Academics have used the historical approach comparing China and Vietnam with North Korea (Harvie, 1992; Herold, 1996; Lee, 1996; and McMillan 1996). ‘Poverty trap’ and ‘big push’ theories have been used to explain economic development and regime change in North Korea (Lee, 2000). Managerial approach (Park, 2005) and marketing concepts (Choe, et al., 2005) have been used to analyse North Korea’s economic strategies. Statistical approaches such as the ‘CGE models’ (Noland et al., 2000), Social Accounting Matrix (Lee and Deok, 2004), and ‘IMAGE Model’ (Meade, 1997) have been employed to estimate the deficiency in data.

2.1 **Historical and Economic Evolution**

Colonial and military influences shaped North Korea’s supply chains in WWII until Korea gained liberation from Japanese colonial rule. After World War II, Korea was split with the north coming under communist domination while a republic, led by the USA, was set up in the south. The South was exposed to market mechanisms but the North was influenced by command economy with emphasis on heavy industry aided by subsidies from China and the Former Soviet Union (FSU) (Hwang 1993). Table 2.1 compares the economies.
Table 2.1: Comparison of the Korean economies.

<table>
<thead>
<tr>
<th>Year</th>
<th>GNI (US$ 100million)</th>
<th>Per Capita Income (US$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>South</td>
<td>North</td>
</tr>
<tr>
<td>1953</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>1960</td>
<td>19</td>
<td>35</td>
</tr>
<tr>
<td>1970</td>
<td>78</td>
<td>90</td>
</tr>
<tr>
<td>1980</td>
<td>627</td>
<td>434</td>
</tr>
<tr>
<td>1990</td>
<td>2635</td>
<td>572</td>
</tr>
<tr>
<td>2000</td>
<td>5096</td>
<td>168</td>
</tr>
<tr>
<td>2003</td>
<td>6086</td>
<td>184</td>
</tr>
</tbody>
</table>

Adapted from Hwang (1993) and Bank of Korea

In the early years after the Korean War, the North outgrew the South with extensive resources, basic infrastructure inherited from the Japanese occupation and an internally focused economy chosen by the Stalinist state (Choe et al., 2003). From the 1970s North Korea became victim of its own success. The old equipment, including trucks and railways, that was used successfully during the early days of heavy industrialisation became obsolete and dilapidated. North Korea had to borrow heavily from its neighbours (French, 2005). The situation reversed as the South benefited from successive five-year plans focused on industrialisation and attracting foreign direct investment with South Korea's GNI trebling in the 1970s. In the areas of logistics, the concept of international supply chains have developed in the South, integrating the country with global systems. This can be seen with the development of ports and shipping in South Korea. On the contrary, logistics concepts are eschewed in the North as a result of a command-economy which placed little emphasis on development of transport and logistics infrastructure. This research aims to provide strategies for its redevelopment.

After the collapse of the FSU and Eastern Bloc in 1990, North Korea's economy went into a decline. According to Kang (2000), besides production problems that arise from inadequate work incentives, fertilizers, and working tractors, there was also the problem of distribution. No trains were available for transport. The only alternative was the country's aging fleet of run-down trucks, which kept breaking down on the unpaved roads. Rice that was needed in the cities sat rotting in the countryside, while manufactured goods the country people needed never left the city.

2.2 Overview of Transport and Logistics in North Korea

North Korea's economy grew steadily after the Korean War in 1953 but stagnated since the mid-1970s. Currently, logistics concepts are eschewed with transportation considered not as part of the social infrastructure, but as an element of production, the purpose of which is to take care of the demand derived from other economic activities. Transport problems include deficiencies in infrastructure, poor maintenance, inefficient institutional structures and imbalances in modal shares (Oh, 2001). Logistics infrastructure in North Korea is archaic and its transport system was developed mainly for military purposes.
Transport infrastructure development was impeded by mountainous terrain, creating weak links between the western and eastern regions and a national transport network concentrated along the east and west coasts. A rail oriented network ensued, believed to be more efficient in mountainous terrain, but road, maritime and air transportation were underdeveloped (Oh, 2001). This section reviews the present conditions of the transport and logistics infrastructure in North Korea. Due to its autarky economy, North Korea stopped publishing external statistics and information from the early 1960s. This posed problems in conducting any reliable research on the country. Data in the following sections are obtained from South Korea and international institutions, which is considered the most reliable information available.

2.2.1 Railways

Rail forms a major part of the transport network in North Korea, linking all major cities and accounting for 90% of cargo and 62% of passenger movements (Library of Congress 2005). With 98 percent of the network being single track and dilapidated, it cannot be sufficiently utilised. The conditions of the rail beds, linear rails and rail tracks are so poor that the railways are known for their slow speed and low capacity (Kim et al. 2001, p. 28). In June 2000, the historic summit between the leaders of North and South Korea resulted in the idea of reconnecting the Trans-Korea Railway (TKR). Four TKR lines were constructed during the Japanese colonial period: Kyongui Line (Seoul-Shinuiju), Gyeongwon Line (Seoul-Wonsan), Mt. Gumgang Line (Seoul-Mt. Gumgang) and Donghae North Line (Samcheok-Anyun). Both Korean governments agreed to make the reconnection of the Kyongui Line the highest priority. Connecting the South-North transport network can turn both sides’ weaknesses into strengths from which both can benefit (Yang 2002). According to Ahn (2002), if the railway and highway between the two Koreas are connected, logistics costs will be reduced by 75 percent of the current level and shipping period shortened by 80 percent. In addition, South Korean companies will enjoy the ability to ship goods from Pusan to Paris, which reduces shipment times of goods and materials from South Korea to Europe by as much as 15 days. Presently, container shipments to and from Europe require nearly 30 days (Choe, 2005). There will be reduction in transportation charges for shipping goods by land rather than sea. It is estimated that transportation charges between both states will be reduced to US$250/TEU compared to the current rate of US$1,000/TEU (McMullan, 2000).

This project has form part of the catalyst leading to an increase in inter-Korean exchanges. The reconnection of the TKR brings benefits including, reduced tension between both Koreas, lower logistics costs, increased direct trade in the region of US$400 million, standardised facilities in both Koreas and the promotion of inter-Korean economic cooperation (Yang 2002). Reconnection of the TKR will not only establish a new transportation network that covers the Korean Peninsula but also the whole of Northeast Asia, creating the world’s largest overland transportation route linking Asia to Europe. Plans to connect the TKR and Trans-Siberian Railways (TSR) to replace partial maritime shipments between South Korea and the Russian Far East would permit cargo to be transported by rail from South Korea to Europe (Tsuji 2005). North Korea will serve as a transit point between the Trans-China Railway in the west and the TSR in the north-east. Investment is essential to
rehabilitate disconnected transportation routes, construct double track lines and modernise communication systems and equipment to enhance operational efficiency.

2.2.2 Roads

The development of roads in North Korea started after the Korean War. However, the closed economy, a lack of financial resources and mountainous terrain in most of the country constrained the road network. The ratio of paved roads is only 6.7 percent, which in turns lowers the efficiency of roads. Unpaved roads are covered with gravel or crushed stones, or have dirt surfaces and are poorly maintained. As part of the June 2000 summit, the rehabilitation of roads was one of the suggestions to improve inter-Korean exchanges. Presently, the Panmunjeon-Gaesong route is attracting considerable interest from both sides due to the establishment of the Kaesong Industrial Complex. North Korea needs to take steps to reconstruct the disconnected route to improve the efficiency of roads. As North Korea’s trade with China increases, the Dandong-Shinuiju and Wonjeong-Rajin routes will play an important role. In the future, Shinuiju will be an important location for the creation of a special economic zone, similar to Kaesong. Therefore, it is important that the government provides adequate road access around the surrounding areas.

2.2.3 Ports

As North Korea traditionally only trades internationally with other socialist countries there has been little need for investment in seaports and given the availability of rail and land transportation with China and Russia this has further reduced the need for seaports. Due to a lack of investment in ports, regular power failures result in slow unloading of ships and few ports can handle 40ft containers as there are no cranes available (French, 2005). The main problems for North Korea’s maritime trade include the country’s collection of exorbitant port-entry fees and poor quality of inland transportation within North Korea (Ahn, 2002). Other bottlenecks identified include the lack of freight cargo base; lack of funds for infrastructure development; inadequate shipping and port capacity; lack of container handling capacity; and an inadequate maritime fleet. North Korea should seek to lower its port charges thereby attracting more cargoes through its ports. In addition, as ports provide a gateway to world markets and opportunities for access to trade with other countries, the maritime sector will be important in North Korea’s future economic development and transition to a market economy.

2.3 Special Economic Zones (SEZs)

Establishing new SEZs is regarded as the most important policy measure for opening the North Korean economy (Lee, 2004). In September 2002, North Korea established two SEZs - Shinuiju and Gaesong. The government expects its SEZs to assist in acquisition of hard currencies, which will facilitate trade with other countries (Jung et al., 2003). Gaesong Industrial Complex - dubbed the ‘Peace Project’, was established as a joint project between both Koreas. Gaesong integrates South Korean capital and technology with North Korean land and labour
to produce reciprocal economic profits (Lee, 2005). When the ongoing restoration of the Trans-Korean Railway (TKR) is completed, land routes between China and South Korea will open and consequently, Gaesong could become a major transportation hub and export-oriented production base for North Korea (Lee, 2004). South Korean SMEs under business associations such as the Korean Federation of Small and Medium Businesses and the Korean Federation of Textile Industries have expressed strong interest in moving their production base to Gaesong (Lee, 2004) with 1,300 applicants received (Kim, 2003). About 300 to 400 SMEs are scheduled to move into the park when the first phase is completed in 2006. The claimed benefits of Gaesong are:

- Close proximity to Seoul, easy for transportation of raw materials and electricity (Lee, 2004);
- Easy access to cheap land (Park, 2004);
- Strategic location (Choe et al., 2005);
- Affordable and educated labour force (Choe 2003);
- Same language and culture (Park, 2004); and
- Beneficial effect on economic growth of North Korea (Lee, 2004)

Jung et al.’s (2003) study showed that SEZs will most likely play a critical role as one method of stabilising the country and thus the Korean Peninsula. They reviewed and analysed surveys of South Korean businesses that were conducted by various South Korean governmental and non-governmental organisations. They concluded that poor quality of North Korean logistics and infrastructure support impairs business opportunities in Gaesong. This research intends to follow on their premise by looking at the transport priorities needed.

There are many issues yet to be resolved in order for Gaesong to be successful and for North Korea to induce foreign investment. Social infrastructure such as electricity, harbours, railways and communication systems is very poor and not expected to improve in the near future (Oh, 2002). One survey showed that one of the most serious problems faced by South Korean businesspersons is that North Korea partners do not understand business practices (Park, 2005). Lingering political and military instability, lack of a legal and institutional framework, the North’s lack of purchasing power and product variety and difficulty in finding markets for North Korean products poses a problem as well (Hoon, 2001). Payment methods and logistical distribution systems between both Koreas are still primitive thus contributing to the high transport costs (Dong, 2004). Lee (2004) points out that there are several challenges and future tasks for North Korea. These include facilitating the TKR restoration and infrastructure building. Another major issue is the high transport cost. In the case of textile products processed in North Korea and sold to the South, transportation accounts for 40 percent of total production costs. The cost of transporting a 20 foot container from Incheon port to Nampo is four times higher than the cost of shipping the same container to China (Babson 2001).
3. Research Methodology

The central research question is as follows:

“What is the role of ‘reformed’ North Korea on transport and logistics development in the Korean Peninsula and Northeast Asia?”

Very limited research has been conducted into the transport and logistical development in North Korea. There is a need to appraise the present transport and logistics conditions in the country. In addition, it is important to examine the future economic and political impact of the inter-Korean exchanges on transport and development links between North and South Korea.

North Korea has the potential to play a pivotal role in Northeast Asia with the development of its transport and logistics infrastructure. With the impending reconnection of the Trans-Korean Railway, exporters to Europe can exploit the role of the Trans-Siberian Railway. Exports from China to Japan can use the transport facilities in North Korea, which now has to be re-routed through South Korea. Research has been conducted widely in Europe and North America into multimodal transport connectivity. However, very little equivalent work has been carried out in Northeast Asia.

In the foreseeable future, South Korean investors will be the only companies to invest heavily in the North. South Korean investors have the added incentive to invest in the North due to a number of factors such as relatively cheaper labour and land. Therefore, this research seeks to survey the South Korean investors about their opinions on factors that will determine their investment in North Korea, which has implications on the priorities of transport developments and introduction of logistics concepts.

The research and especially the data collection process will be guided by the following questions:

1. What are the strategies and priorities in planning transport development in North Korea?
2. What are the types of logistics concepts that should be introduced in North Korea?
3. What are the short, medium and long term strategies for development of transport infrastructure?
4. How is North Korea going to finance the development projects?
5. What is North Korea’s role within the context of transport/accessibility in Northeast Asia?
6. What do potential and present South Korean investors think about investing in North Korea?
7. What are the implications for transport and logistics of the impending reunification on the Korean Peninsula?
After considering the literature and research questions, this study will employ a mixed methods approach. Mixed methods studies are those that combine the qualitative and quantitative approaches into the research methodology of a single study or multi-phased study (Tashakkori and Teddlie, 1998). This research employs a two-stage sequential design strategy. The first stage is characterised by an initial phase of qualitative data collection and analysis, which is followed by a phase of quantitative data collection and analysis. The choices of qualitative Interviews, quantitative questionnaires and scenario planning have been made on the basis of research objectives, the type of data require, the research methodology adopted and the extent of resources available.

3.1 Conceptual Model

A preliminary conceptual model for the research is presented in Figure 3.1. The model does not seek to test or explain casual relationships. However, it does aim to generate factors that could influence the development of transport and logistics developments in ‘reformed’ North Korea. Then these factors will be facilitated into the qualitative research to generate further hypotheses for the quantitative study. These factors are drawn from mainly from (1) the PESTLIED model, which is used for assessing an organisation external environment (Harding, 1998), (2) MacCarthy and Atthirawong (2003)’s study into factors affecting international location decisions and (3) Roe’s (2003) work into shipping policy making. These contexts provide a framework to analyse the possible strategies for transport and logistics development.

The model identified sub-elements which are critical to the development process. Political stability between both Koreas and in the region is crucial for North Korea to induce much needed foreign capital to rehabilitate its dilapidated transport modes (Babson, 2001; Lee, 2004 and Lim, 1997). Currently, North Korea is not a member of international financial institutions, which prevents them from securing loans. The reconnection of the TKR is the first step in reconstruction (Oh, 2001 and Tsuji, 2005). North Korea has to embrace new modern technologies (Lee, 2004 and Oh, 2001). Economic reforms in the form of SEZs will enhance regional and local development with implications on transport linkages and accessibility (Lim, 1997). Transport cost factors will also be an important determinant (Babson, 2001; Ahn, 2002 and Jung et al., 2005). New laws are needed for international payment methods (Dong, 2001). Managers need to be educated in transport planning policies and logistics concepts (Lee 2004 and Oh, 2005). The need to find new markets for North Korean products, regional and/or local, plays an important role in transport development (Hoon, 2001; Kim, 2003 and Park, 2004). A pool of skilled, educated workforce is required for understanding operating new technologies (Babson, 2001; Jung et al., 2003 and Lee, 2005). Organisational factors centre on logistics linkages and accessibility as well as the logistics distribution system (Dong, 2001).

3.2 Qualitative Interviews and Analysis

Consideration was given to the different types of qualitative techniques available. After weighing up time and cost factors, it was decided that interviews would be the best method to be employed. Detailed consideration is given to the
identification of eligible academic and industry experts on North Korea. Once selected, all experts will be sent a preliminary letter stating a description of the research, how the respondent was chosen, time expectation and confidentiality assurance. Unstructured pilot interviews will be conducted with experts depending on the geographical region where the expert is located and also considering the time constraint. Following pilot interviews, the pool of experts to be interviewed will be extended. All statements will be subsequently analysed using computer-assisted qualitative data analysis software

3.3 Quantitative Surveys and Analysis

From the assumptions derived, a quantitative questionnaire will be conducted with South Korean SMEs in the second stage. This study will be carried out in South Korea in conjunction with a research institute yet to be determined at this stage. The duration of this study is expected to take six months. The purpose of collecting data from South Korea is due to the autarkic nature of the North where it is virtually impossible to interview North Koreans or conduct case studies with North Korea companies. As such, the next best option is to speak to South Koreans where there are many similarities in terms of language and culture. Prior to departure to South Korea, a pilot survey will be conducted to make necessary changes to the structure, wording or content of the questionnaire. Thereafter, the questionnaire will be administered to the CEOs and managers of South Korean SMEs using the data base from the Korean Federation of Small and Medium Business and Korean International Trade Association (KITA). The respondents will be contacted prior to the sending out of the questionnaire to seek their approval. The number of SMEs to survey is yet to be determined at this stage. However, sampling methods will be employed to determine the size of the sample. Data collected will be analysis using SPSS for windows. The advantage of SPSS over Minitab is its ability to manage variables collected from multiple answer questions and arrange these into set for analysis purposes (David and Sutton, 2004). Preliminary analysis of the data collected will consist of frequency and cross tabulation tables to provide a basic understanding of the data. Thereafter, multivariate techniques will be used for advance analysis of the data.

3.4 Scenario Planning

After analysing the quantitative data, the next stage of the study turns to scenario planning. In future-oriented strategic planning, when making sense of complex events, presenting alternative images for the future instead of extrapolating current trends from the present is very important (Van der Heijden, 1997). Using scenarios is researching the future by recognising the warning signs and the drama that is unfolding one can avoid surprises, adapt and act effectively. Ultimately, the end result of scenario-planning is not a more accurate picture of tomorrow but better decisions about the future (Schwarz, 1998). Scenarios can be categorised into two major groups: exploratory scenarios that start from the past and present trends and lead to likely futures, and anticipatory or normative scenarios that are built on the basis of alternative visions of the future (Godet, 2001).
A scenario is a tool that describes pictures of the future world within a specific framework and under specified assumptions. It is a description of a hypothetical development or future states of the city. The scenario approach involves a description of two or more scenarios designed to compare and examine alternative futures (CEC, 1994). Scenario building allows new perspectives to be brought to bear on these issues, particularly in situations of high uncertainty such as the next 20 years. Consideration of alternative scenarios, reflecting the major uncertainties encountered in moving in particular directions, can tell us what might happen (RAND, 1997) and give an insight into possible ways to avoid adverse outcomes (such as urban unsustainability and low quality of life) that might occur in the future.

Eight main stages in the process can be distinguished (Stead and Bannister, 2003). Although some stages are reliant on one or more previous stages in the process, the link between each stage does not form a simple sequence. There are interactions and feedback loops between many of the main stages. Upon developing different scenarios packages, they will be validated and assessed by North Korea experts, resulting in composite policies for North Korea.
Figure 3.1 Conceptual Framework

Historical factors influencing transport and logistics policies in ‘Autarky’ North Korea

Communism → Political → Juche (Self-sufficiency) → Economic → Heavy industries

Factors influencing transport and logistics developments in ‘Reformed’ North Korea and Northeast Asia

Political → Technical → Economic → Legal → Managerial → Spatial → Social → Organisational

Reunification with South Korea → Rehabilitation of transport infrastructure → Market reforms - Special Economic Zones → New laws for logistical practices → Knowledge of logistics concept → Regional and local markets → Pool of skilled, educated labour force → Logistics linkages and accessibility

Strategies for transport and logistics development in ‘Reformed’ North Korea (Main issues)

Regional stability and investments → New technologies → Regional and local development → Payment methods → Transport planning → New markets for N. Korean products → Introduction of business practices → Logistics distribution system

Qualitative Interviews (with academic and industry experts) – facilitates Quantitative research

Quantitative questionnaire (South Korean Small and Medium Enterprises)

Scenario Planning Model

Develop composite transport and logistics policies for North Korea

Source: Author (2005)
4. Conclusion

In its attempt to analyse the role of North Korea on transport and logistics in the Korean Peninsula, this paper is an illustration of the progress of research carried out to date. The first stage of the research examined the background of all aspects of North Korea by means of a literature appraisal. The second stage involved the development of a conceptual model to explore the factors that influencing transport and logistics developments in ‘reformed’ North Korea and Northeast Asia. The third stage encompassed the design of the system by which the conceptual model will be tested in the real world, through triangulation methods of qualitative interviews, quantitative surveys and scenario planning. Results derived will be used as composite policies for North Korea.

There is a great degree of uncertainty in the Korean Peninsula especially with North Korea’s nuclear weapons issue. It is difficult to comment on the type of changes will North Korea experience in the future. A lot depends on its willingness to continue to reform its economy and maintain stability in the region. Recent experiments with market reforms presented a last resort for survival, with the transition to a market economy resulting in changes including the creation of special economic zones, presenting a new impetus for logistics. However, with logistics concepts often eschewed, no efficient network for effective information and product flows and transport infrastructure decrepit due to underinvestment, there is huge potential for

**Figure 4.3 Eight stages in scenario planning process**

Source: Author (2005)
developing logistics facilities and systems. All transport modes are undergoing development which will enhance connectivity and special economic zones are attracting investment from South Korea companies. There is potential for future logistics development. Impending reunification, reforms and increasing inter-Korean trade present potential for corporate investment, with SEZs revealing South Korean companies keen to invest. Although many South Korean companies relocated to China because of cheaper labour and land, an improving logistical environment will help to attract them back to North Korea, promising improved regional stability, an affordable labour force, a market of 24 million people and abundant underdeveloped natural resources and a strategic distribution location in Northeast Asia.
References


Choe, S.T. (2003), North Korea Moving from Isolation to an Open Market Economy: Is it Time to Invest or Continue Observing?, *Competitiveness Review*, 13, 2, pp. 60-69.


Yonhap News (2005), Growing Number of Local Firms Ask Whether to Relocate Strategic Goods to North Korea, Yonhap News, 26 April 2005.
INVESTIGATING ACCIDENTS IN THE MARITIME DOMAIN - THE INFLUENCE OF CREW INTERACTION AND COMMUNICATION AMONG MIXED NATIONALITY CREW ON SAFETY

Robyn Pyne

Abstract

The Human Element, among a group of maritime human factors terms, is used widely within the commercial maritime environment to incorporate many issues being addressed by maritime research institutions and industry bodies. In accident investigation losses are attributed frequently to Human Error, which itself has a variety of identifiable categories e.g. Human Communicative Error. This paper describes an analysis of a number of maritime accident reports in which a failure of effective crew communication played a central role in the causal chain. The analysis is based on an aviation industry accident investigation taxonomy, ADREP 2000 Taxonomy (IACO, 2004). The paper discloses, accentuates and exemplifies the structure of problems related to maritime crew communication and problems related to different cultures and languages. It also explores the value, contribution and limitations of formalised taxonomies and analysis systems to maritime training authorities, when they are applied as tools in the analysis of accident reports. It also looks to highlight the need for further research into the sociolinguistic aspects of shipboard operation within the shipboard society not just during times of restricted manoeuvrability, but during everyday working and social communications in pursuit of recommendations to aid the reduction of occurrences of Human Communicative Error.

1. Introduction

1.1 Human Error, communication and multi-cultural crewing issues

Research into the pros and cons of multicultural crews is attracting increasing interest from across the widespread maritime research community and industry alike. The issues arising from the most recent of research publications are those of onboard culture and language as it is commonplace on ocean going vessels engaged in world trade that staff will frequently come across crews and other industry groups whose members come from a number of countries and related cultures (Horck, 2004). Given that this paper is proposing that the interaction of multicultural crews onboard has an effect on English language ability, these concerns regarding culture - the pattern of activities, values and artefacts of a society - are of utmost importance and relevance.
It is evident from studies of maritime accident reports where, in the chain of causation, ‘human error’ has been identified as a significant factor, that it is possible to break down this category further and to identify that crew communication failure has played a key role in the incident (European Union-DG VII 1999). Poor communication between crewmembers from the same culture who are speaking the same language can, through misunderstandings and mistakes, be a threat to the overall safety of a vessel and pose an additional threat if one considers the risk of subsequent pollution. If one adds the additional variables of crews using English as a second language and the cultural differences which may be experienced, then the odds of miscommunication may be increased. In turn, it is recognised that maritime safety can be enhanced by the improvement of crew communication, facilitated through training in the use of Standard Marine Communication Phrases (SMCP) (Hughes, 2000).

Navigators will work in a bridge team when required to do so in order to ensure the safe operation of a vessel. During such times e.g. when manoeuvrability is restricted, miscommunication can cost lives, loss of cargo, and pollution. Hutchins (1996) writes about Learning to Navigate and describes navigation as an activity normally carried out by an individual, working alone, who is joined by other deck crew when the computational requirements of the task are assumed to exceed the capabilities of one Navigator. Examples of such a time include times of restricted manoeuvrability, when a ship enters or leaves port or when there is heavy congestion of maritime traffic. At these times it is common for a Pilot to be onboard and it is through the presence of a Pilot that the most comprehensive crew communication research project to date, the MARCOM Project, collected some of its data. Data were collected from the Pilot’s own experiences of mixed nationality crews’ use of English and Maritime English and from observations made during periods of pilotage, when a researcher joined the vessel with the Pilot.

1.2 Accident Investigation and Human Communicative Error

Much of the work carried out in the field of marine accident investigation reflects what is underway and has already been done in the aviation industry, which is generally viewed by the marine industry as being superior in terms of its advancement in research and policy making and its overall safety record. It is for this reason that a taxonomy used in the investigation of accidents in the aviation industry - the ECCAIRS ADREP 2000 taxonomy of explanatory factors - has been applied in the analysis of maritime accidents in search of a further breakdown of Human Communicative Error (HCE) into relevant categories. HCE can be defined as Human Error which occurs as a result of a failure in communication, be it ship to ship, ship to shore or intra-ship.

In order to illustrate the failures in communication and the role of formal taxonomies, this paper describes a number of accident reports in which HCE has been cited as a factor in the chain of causation.
1.3 Objectives

The main issues surrounding culture and the work context are discussed by means of a literature review in Section 2 of this paper. Sections 3 contains any references to methodology and features concepts suitable for use in analysis of marine accidents. Section 4 incorporates a brief summary of each accident and the analysis of secondary data from accident reports with the ultimate aim of breaking down the HCE involved and classifying it using the ADREP 2000 Taxonomy.

Given that this paper has been written with its firm foundations set in the hypothesis that the interaction of multicultural crews onboard has an effect on the ability of the crew to communicate effectively, the concerns regarding culture and HCE are of utmost importance and relevance and are integrated into the investigation and analysis.

Conclusions have been drawn from the secondary data analysis contained within section 4, which incorporates a brief summary of the issues surrounding multicultural crew communication and why the ADREP 2000 Taxonomy is a suitable classification for maritime accidents.

2 Culture and the Work Context

2.2 Intercultural Communication

Intercultural communicative competence is vital because people’s communication styles are inherently culturally bound (Hofstede, 1997). The STCW convention and its amendment in 1995 theoretically dissolve such divides as value and culture in training and education, but in reality these factors still thrive. Discussing culture is a sensitive issue and to avoid slipping into stereotyping this paper will, in referring to a particular nationality, be referring to a bell curve of a culture. In this bell curve the people in the centre represent the mainstream, which should be interpreted with an awareness that cultures are made up of individuals whose behaviour can vary greatly, (Hutchins, 1996). The 1995 major revision of the STCW Convention saw the first inclusion of specific requirements for English Language certification. Communication lapses identified as Human Error in the causal chain of accidents have led to the use of English as the common language under the revised STCW Convention 1995.

Setting aside the further issues of loneliness, stress and fatigue brought about by the reduction in numbers as a result of technological advances, the 80% of the world’s merchant ships which have become multilingual and multi ethnic in terms of crew composition (European Union-DG VII 1999) are facing the reality of those very concerns relating to the Human Element as outlined by Horck (2004), Hughes (2000), (Hofstede, 1997) and Hutchins (1996).
If there are also problems of communication contributing to a lack of mutual confidences, suspicions and misunderstandings, then the opportunities for human errors leading to dangers to the ship, the people on board and the environment, are greatly increased

(European Union-DG VII 1999).

2.3 Potential Limitations

As with any study of activity in any context - the circumstances in which an event occurs e.g. an occurrence of an accident - the primary debate surrounding the observation of human behaviour, specifically that of cultural interaction among mixed nationality crews in relation to its affect on English language competence, for scientific analysis, centres around one fundamental question; how can one analyse and interpret data that record and describe human behaviour and discourse? (Engestrom, 1996)

The written accounts of marine accidents are recorded in official Accident Reports and any analysis of data in this paper must come from these written accounts. The ability of these reports to record the relevant information for discourse and communication analysis is limited, as they are usually based on log books, charts and witness accounts and are retrospective. One difference between shipping and aviation is the use of the black box flight recorder which records intra-cockpit speech and external verbal communications. Such devices are available in shipping, but their use is limited and is not enforced by statute, therefore data from this source is not readily available to marine accident investigators.

3 Methodology

3.1 Concepts suitable for analysis of HCE in marine accidents

There are two principal approaches to exploring the issues relating to culture and language issues that were identified in the literature review. The first is a post accident approach and involves the analyses of maritime accidents and indications of cultural issues and HCE in safety critical situations. These analyses are based on accident reports and the technical terminologies contained within them reflect the fact that investigators are often navigators or engineers. This approach suits analyses which incorporate the use of formal taxonomies and has the inclination to fall within the quantitative bracket of research as its results are often suitable for manipulation using statistical programmes in order to generate frequency descriptives and to establish statistical associations. The ADREP 2000 taxonomy suggests a terminology based on psychological concepts, rather than the technological ones used by the navigators or engineers who usually investigate marine accidents, which are suitable for description of the human element and human factors involved in an accident. This paper promotes the belief that the development of the marine accident investigation process to include Industrial Psychologists and Human Factors specialists could aid the inclusion of such
psychological terminology into accident reports and would subsequently make this information source more suitable for analysis using the ADREP 2000 taxonomy.

The second of these approaches to exploring issues relating to culture and language in the marine environment is a qualitative one, which studies culture and language in everyday life. By studying crew communication and behaviour and by interviewing the crew onboard multicultural vessels the researcher can observe the issues relating to culture and language in non-safety critical situations and inform how this behaviour and discourse may relate to or influence communication in safety critical situations. This paper utilises the former approach.

3.2 Accident classification from taxonomy

A quick browse through maritime accident reports from USA, Canada, UK, Australia and Denmark shows several examples of accidents in which the human element and culture or language related problems played an important role in the causality of the accident. The following examples can be mentioned:

- Problems related to multi-cultural crews (e.g. the Bunga Teretai Satu accident and the Scandinavian Star)
- Problems related to different cultures/languages among crew and pilot (e.g. the Bright Field accident)
- Problems related to different cultures/languages among crew and passengers on passenger vessels (e.g. the Skagerak accident and the Scandinavian Star accident)
- Problems related to different cultures/languages with respect to external communication, VHF communication with other vessels (e.g. the Royal Majesty accident)

4. Results and Discussion

4.1 Problems related to multi-cultural crews

The grounding of the Malaysian flag container ship Bunga Teratai Satu on the Great Barrier Reef occurred when a waypoint alteration was not made. The significant act identified by the accident investigators to have caused the waypoint alteration to be missed was the telephone call made by the Pakistani Mate and his wife to their family. The mate had developed a practice of asking the AB from Myanmar to plot the ship’s position from the GPS every hour when the ship was in open waters. As the Pilot had left the ship and they were out of the compulsory protection zone, but not onto their next waypoint, the AB assumed his role and proceeded to plot this position.

The position plotted was adjacent to waypoint 34 on the ship’s passage plan, where the ship’s course was due to be altered. According to the AB, he kept expecting the mate to come back into the wheelhouse to alter course. But the mate did not re-enter the wheelhouse until about 0715, whereupon he and his wife proceeded to make some coffee at the sink at the port side of the wheelhouse. At around 0717,
after making coffee, the mate went to the chart table and checked the 0700 position. He looked over the chart table console and told the AB that he had made a mistake in plotting the position. Shortly afterwards and in desperation he told the AB to ‘change to hand steering’ and shortly after the vessel was aground.

The AB was obviously an intelligent young person with some six years seagoing experience. He had learnt to plot GPS positions but was not familiar with chart symbols or issues such as scale, or time/distance estimations. He did not realise the ship was standing into danger. He resumed his lookout duties assuming that the mate would make the appropriate alteration in due time. Such an attitude reflects a large ‘power distance’ according to Hofstede (1997) and in their account the accident investigators noted that there existed a strict hierarchy between the Pakistani senior officers and the Malaysian, Indonesian and Myanmese junior officers and crew.

It was important in the national culture of the crew that the AB - although he knew that something was wrong - did not question the decisions of his superior. The ADREP 2000 taxonomy classifies this problem as follows:

202030000 Cultural issues.

*Factors related to cultural issues, e.g. crew mix, authority gradient, cultural issues and crew resource management.*

It is less likely that this problem could occur with an all Scandinavian crew because there are few differences in culture, nor can this culture be characterised by a high ‘power distance’ (Hofstede, 1997). The appropriate ADREP 2000 category is:

501010100 The interface between humans in relation to communication between crew members.

*Factors related to the interface between humans in relation to communication between crew members.*

4.2 Problems related to different cultures/languages among crew and between the crew and the Pilot

The communication problems in the Bright Field case are quite similar to the problems in the Bunga Teretai Satu case in relation to Hofstede’s cultural dimensions and the phenomena he describes as ‘power distance’. While the Bunga Teretai Satu case was a matter of communication between crewmembers from the same culture, the Bright Field case illustrates a situation with a crew and a pilot from different cultures: American and Chinese. The word “no” is a very impolite word to the Chinese. It is therefore the cultural practice of Chinese crews that they always answer “yes” - especially to an authority such as a pilot - even though they are well aware that the correct answer is “no”. This is in fact a very extreme example of concealing communication according to a model by Metze and Nystrup. The correct ADREP 2000 taxonomy classification of the problem is again:
202030000 Cultural issues.
Factors related to cultural issues, e.g. crew mix, authority gradient, cultural issues and crew resource management.

Further, the pilot suffered from a lack of information due to the fact that he was not able to understand the communication between the engine room and the bridge, which was in Chinese. He was prevented from recognition of the engine problems from the bridge/engine communication and he therefore suffered from information deprivation. The correct ADREP 2000 taxonomy classification of this problem is:

501010500 The interface between humans in relation to language.
Factors related to the interface between humans in relation to the use of a particular language, e.g. English in a French speaking area.

4.3 Problems related to different cultures/languages among crew and passengers on passenger vessels

The ferry Skagerak foundered in heavy weather in 1966 on route between Norway and Denmark. The passengers and the crew were all saved due to a remarkable effort from the crew as well as from the vessels and helicopters engaged in the search and rescue operation. The mustering of the passengers was not done using loudspeakers. A member of the crew knocked on the door to every cabin and asked the passengers in Norwegian or Danish to don their lifejackets and go to the mustering stations as quickly as possible. A couple of French speaking passengers did not understand the instructions given and assumed that the crewmember talked about the arrival. They therefore dressed carefully and prepared for the arrival and went to the passenger area where they found the other passengers dressed in pyjamas and lifejackets. Although the situation now can be considered amusing - the passengers were in fact saved - it is evident that the problems with the communication between the crew and the passengers could have had fatal consequences.

The ADREP 2000 categories of explanatory factors for this particular example from the foundering of Skagerak are:

501010500 The interface between humans in relation to language
501011100 The interface between humans in relation to aural/oral interpretation/misinterpretation
Factors related to the interface between humans in relation to interpretation/misinterpretation associated with communication.
The ferry Scandinavian Star burned out completely on a voyage from Norway to Denmark in 1990. A lot of the passengers and crewmembers died in the fire, and the accident was considered to be one of the worst passenger ferry disasters ever in European waters. Witness testimonies express problems related to crew-passenger communication and crew-crew communication due to different languages. The captain even complained about the poor English language skills of the crew in a telex to the shipowner before the accident occurred. The appropriate ADREP 2000 category describing this problem is:

501010500 The interface between humans in relation to language.

4.4 Problems related to different cultures/languages with respect to external communication, VHF communication with other vessels

The grounding of Royal Majesty on Rose and Crown Shoal near Nantucket, Massachusetts on 10th June 1995 is a very complicated case with a number of human factors issues. The issue to be used in this analysis is the communication between M/V Royal Majesty and a group of Portuguese fishing boats and the ship to ship communication between the fishing boats on VHF channel 16 a short time before the grounding. The M/V Royal Majesty was off route due to a malfunction of navigation equipment on the bridge, but the crew were unaware of this malfunction due to false indications from the navigation equipment. At a certain point, the crews on board a group of Portuguese fishing boats realised that M/V Royal Majesty was heading towards danger and tried to call it on channel 16. Because they called a vessel on a certain position, and the crew on board M/V Royal Majesty was convinced that they were in another position, the crew on M/V Royal Majesty did not respond to the call - the call was made in English. The call in English did not indicate any danger, but the ship to ship communication within the group of Portuguese fishing vessels did indeed indicate danger, but this communication was in Portuguese and was not understood by the crew on board M/V Royal Majesty. There is a possibility, that the crew would have paid attention to it, had the communication been in English and there is a further possibility that the crew might had been alerted that their vessel was off course. Again the appropriate ADREP 2000 category describing this problem is:

501010500 The interface between humans in relation to language.

5 Conclusion

The issues surrounding intercultural crews and the potential for HCE were identified in the literature review and examples have been given from maritime accidents of problems related to communication, language and culture. These examples have been analysed successfully using the ADREP2000 taxonomy and the psychological model of professional communication formulated by Metze and Nystrup.
The need for clear verbal communications between parties in the commercial marine environment is multi-faceted as the ship is the working environment, learning environment and social environment for its personnel. The multinational crew must interact and communicate in a common language to maintain ‘social harmony’ in an off duty context and in their everyday ‘teamwork’ to ensure effective day to day operation. The most commonly recognised failure occurs with the level of understanding of English between ship to ship and/or ship to shore under conditions of restricted manoeuvrability, or when under critically congested circumstances where little time or space can be afforded for mistakes to be made. It is apparent from the examples of accidents that seafarers cannot be expected to communicate in a variety of languages, using some English of non standard origin and then be expected – under the stressful circumstances of restricted manoeuvrability or emergency operations – to revert to using a standard marine vocabulary and to remove any redundant language.

It is possible to minimise the amount of accidents directly related to poor communication by improving crew communication through training and education of the crew, improved procedures for communication, better selection of personnel and improved design of maritime equipment and technology including means for communication (e.g. telephones, VHF, radios etc.). Any effort of improvement of crew communication should be based on fundamental knowledge about the dynamics of crew interaction and communication as can be obtained from analyses of maritime accidents using psychological terms and concepts as exemplified in this paper.
6. References


CREDIT RISK IN EGYPTIAN BANKS

Hussein Abdou

Abstract

This research aims to identify the currently used techniques in evaluating credit risk in Egypt's banking sector, then to evaluate these techniques and to develop model(s) to evaluate credit risk in Egyptian banks. In a pilot study, the researcher conducted informal interviews with key personnel in three of the Egyptian banks, in order to evaluate credit risk policies.

The pilot study showed that these banks do not use any of the statistical techniques in the evaluation process. So, the researcher expects that there will be differences between the existing techniques and the proposed techniques in terms of the results of the evaluation process, which will be reflected in the prediction ability in the future, after the implementation of the proposed techniques. At the final stage of the research an investigation will be conducted, by means of a questionnaire survey of Egyptian banks, into attitudinal aspects of implementation of sophisticated techniques. Logistic regression will be run in order to attempt to establish any significant link between the degree of predisposition to implementation and other banks characteristics, for example, size and number of branches.

Introduction

Nowadays, credit risks have become one of the most important financial topics of interest, especially, in the banking sector. The role of credit risks has changed dramatically over the last ten decades, from passive automation to a strategic device.

Risks are uncertainties resulting in adverse variations of profitability or in losses. In the banking universe, there are a large number of risks. Most are well known. However, there has been a significant extension of focus, from the traditional qualitative risk assessment towards the quantitative assessment of risks, due to both evolving risk practices and strong regulatory incentives. The different risks need careful definition to provide sound bases serving for quantitative measures of risk. As a result, risk definitions have gained precision over the years.

Banking risks are defined as adverse impact on profitability of several distinct sources of uncertainty (Figure 1).

Credit risk is the first of all risks (see: figure 1) in terms of importance. Credit risk is critical since the default of a small number of important customers can generate large losses, potentially leading to insolvency, (Bessis, 2003).
Credit risk is the potential loss in the event of default of a borrower, or in the event of deterioration in credit standing.

In this research the term of “default” will be used a lot. Default risk is the probability of the event of default. There are several possible definitions of “default”:
- missing a payment obligation for a few days,
- missing a payment obligation for more than 90 days,
- filing for bankruptcy,
- restructuring imposed by lenders,
- breaking a covenant triggering a cross-default for all lenders to the same entity.

It depends on the default definition, (Bessis, 2003).

Modern banks in all the world need to reduce their risks. Risk may be defined as the volatility or standard deviation (the square root of the variance) of net cash flows of the firm, or, if the company is very large, a unit within it. In a profit-maximising bank, a unit could be the whole bank, a branch or a division. Furthermore, risks specific to the business of banking are:

- Credit
- Counterparty
- Liquidity or funding risk
- Settlement or payments risk
- Market or price risk, which includes
  - currency risk
- interest rate risk
- Capital or gearing risk
- Operational risk
- Sovereign and political risk, (Heffernan, 2005).

This research will focus on the first type of risks, which is credit risk. Banks accept risk in order to earn profits. Credit risk is the risk to earnings and capital that an obligor may fail to meet the terms of any contract with the bank. It is usually associated with loans and investments, but it can also arise in connection with derivatives, foreign exchange, and other extensions of bank credit. Although banks fail for many reasons, the single most important reason is bad loans. Banks, of course, don’t make bad loans. They make loans that go bad. At the time the loans were made the decisions seemed correct. However, unforeseen changes in economic conditions, and other factors such as interest rate shocks, changes in tax laws, and so on, have resulted in credit problems. Credit risk is the primary cause of bank failures, and it is the most visible risk facing bank managers, (Gup and Kolari, 2005).

Therefore, the process of credit risk evaluation has the interest of many researchers nowadays. [Recently, bankers have come to realise that banking operations affect and are affected by the natural environment and that consequently the banks might have an important role to play in helping to raise environmental standards. Although the environment presents significant risks to banks, in particular environmental credit risk, it also perhaps presents profitable opportunities, (Thompson, 1998)].

Banks must evaluate the ability of their customers to repay their financial obligations according to the agreement established between the respective parties. This evaluation process can be carried out using a credit scoring model, which is one of the most important and critical models in this area. Regarding consumer credit risk, people today take credit scoring for granted. It seems obvious that models to assess creditworthiness should be used and yet it was not that long ago that more widespread acceptance was achieved, (Bailey, 2001). Credit scoring techniques assess the risk in lending to a particular consumer. One sometimes hears it said that credit scoring assesses the creditworthiness of the consumer, but this is an unfortunate turn of phase. Creditworthiness is not an
attribute of individuals like height or weight or even income. It is an assessment by a lender of a borrower and reflects both the circumstances and the lender’s view of the likely future economic scenarios.

**Credit scoring** was one of the earliest financial risk management tools developed. Its use by U.S. retailers and mail-order firms in the 1950s is contemporary with the early applications of portfolio analysis to manage and diversify the risk inherent in investment portfolios. Also, credit scoring could claim to be the grandfather of data mining because it was one of the earliest uses of data on consumer behaviour.

**Credit scoring** is the set of decision models and their underlying techniques that aid lenders in the granting of consumer credit. These techniques assess, and therefore help to decide, who will get credit, how much credit they should get, and what operational strategies will enhance the profitability of the borrowers to the lenders [see for more details: (Long, 1973) and (Thomas, et al., 2002)].

**Regarding the history of credit scoring,** this is only 50 years old. During the 1930s, some mail-order companies had introduced numerical scoring systems to try to overcome the inconsistencies in credit decisions across credit analysts. Credit scoring is essentially a way to identify different groups in a population when one cannot see a single characteristic that defines the group but only related ones. The first approach to solving this problem of identifying groups in a population was introduced in statistics by Fisher, 1936.

In the 1940s, Durand was the first to recognize that one could use the same techniques to discriminate between good and bad loans. It did not take long after the war ended for some folks to connect the automation of credit decisions and the classification techniques being developed in statistics and to see the benefits of using statistically derived models in lending decisions. The first consultancy was formed in San Francisco by Bill Fair and Earl Isaac in the early 1950s. During 1960s, the first arrival of credit cards made the banks and other credit card issuers realize the usefulness of credit scoring. The event that ensured the complete acceptance of credit scoring was the passage of the Equal Credit Opportunity Acts and subsequent amendments in the U.S. in the 1970s.

In the 1980s, the success of credit scoring in credit cards meant that banks started using scoring for other products, such as personal loans, while in the last few years, scoring has been used for home loans and small business loans. During the 1990s, growth in direct marketing led to the use of scorecards to improve the response rate to advertising campaigns. In fact, this had been one of the earliest uses in the 1950s.

At present, the banks’ emphasis is on changing the objectives from trying to minimize the chance a customer will default on one particular product to looking at how the firm can maximize the profit it can make from that consumer, (Thomas, et al., 2002).

**Judgemental techniques versus credit scoring techniques:** The overall idea of the application of credit assessment is to compare the characteristics of an applicant with those of past applicants whose repayment history has been observed. If an applicant’s characteristics are sufficiently like those of people who have been
granted credit, and have subsequently defaulted, the application will normally be rejected. If the applicant’s profile is sufficiently similar to those of people who have not defaulted, the application will normally be granted. In general, there are two techniques which can be used:

- loan officer’s subjective assessment, and
- credit scoring (the main interest of this research).

Comparison between a judgemental system and a credit scoring system: was discussed by Altman (1981) as follows:

In a judgemental risk evaluation system, each credit application and the information contained therein are evaluated individually by an employee of the creditor. The success of a judgemental system depends on the experience and common sense of the credit evaluator. Otherwise, in a credit scoring system: some creditors have used their historical experience with debtors to derive a quantitative model for the segregation of acceptable and unacceptable credit applications. With a credit scoring system, a credit application is processed mechanically and all credit decisions are made consistently. The scoring system is based on the addition or subtraction of a statistically derived number of points relating to the applicant’s credit score on the basis of responses given to a set of predictor variables, such as time on a job or the number of credit sources used. Given a statistically derived cutoff credit score, a creditor can thus segregate the acceptable from the unacceptable credit applicants.

On the other hand, credit scoring has been criticized because statistical problems with the data used to derive the model frequently violate the assumptions of the statistical technique used to derive the points, Altman (1981). It is also pointed out that some of the variables used in a credit scoring system may have the effect of social discrimination, although to the statistician the variable may appear to be neutral. Finally, the credit scoring model is derived by analysis of the characteristics of the customers who were once granted credit by the creditor for whom the system is derived. The characteristics of the part of the population to which the credit grantor has not granted credit are not directly considered. Thus the scoring system may provide biased results when it is applied to new credit applicants.

Despite the criticism of credit scoring models, these models can be regarded as one of the most successful models used in the field of business and finance.

Benefits of credit scoring: credit scoring techniques have a number of benefits relative to judgemental techniques for both the lender and borrower:

Credit scoring requires less information to make a decision than does judgemental assessment, because credit scoring models have been estimated to include only those variables which are (statistically) significantly correlated with repayment performance; whereas judgemental decisions, prima facie, have no statistical significance and so no variable reduction methods have been used. Credit scoring models attempt to correct the bias that would result from considering the repayment histories of only accepted applications and not all applications. They do this by inferring how rejected applications would have behaved if they have been accepted. Judgemental methods are usually based on only the characteristics of persons who were accepted, and who subsequently defaulted. Credit scoring models take into account the characteristics of good as well as bad payers, whereas, judgemental methods are generally biased towards perceptions of bad payers only.
Scoring models are built on much larger samples than a loan analyst can remember. Scoring models can be seen to include explicitly only legally acceptable variables whereas it is not so easy to ensure that such variables are ignored by loan officer. Credit scoring models demonstrate the correlation between the variables included and repayment behaviour, whereas this correlation cannot be demonstrated in the case of judgemental methods because many of the characteristics which a loan officer may use are not objectively measured. A credit scoring model includes a large number of an applicant’s characteristics simultaneously, including their interactions whereas a loan officer’s mind cannot do this, the task is too complex. A further important benefit is that given the same data, variables and statistical methodology two statisticians would come to essentially the same weights: they are independent of the person performing the analysis. This is highly unlikely to be so in the case of judgemental methods, (Crook, 1996).

Criticisms of credit scoring:
Credit scores use any characteristic of an applicant regardless of whether one can put forward a clear causal relationship to likely repayment. Furthermore, economic factors sometimes are not included. Another criticism sometimes made of scoring models is that an individual may have the characteristics which make her/him more similar to a bad than a good payer, but may have these purely by chance. Statistically the scoring model is incomplete: it omits some variables, which taken with the others, would predict that you will repay. But unless a scoring model has every possible variable in it which predicts for every single person it will misclassify some people. Another criticism which is sometimes made of scoring models is of indirect discrimination, [for more details see, (Crook, 1996)].

Furthermore, these models are not fail-safe. They are only as good as the original specification, and one limitation is that the data are historical. Though the original discriminant analysis may have produced a Z score which was fairly accurate, unless it is frequently updated, either the variables or weights, assumed to be constant over time, make it less accurate. This problem can be minimised if the bank keeps records of its type-I and type-II errors, and acts to implement a new model to address any necessary changes. A more difficult problem is that the model used imposes a binary outcome: either the borrower defaults or does not default. In fact, there is a range of possible outcomes, from a delay in interest payments to non-payment of interest, to outright default on principal and interest. Often the borrower announces a problem with payments, and the loan terms can be renegotiated. These different outcomes can be included, but only two at a time, (Heffernan, 2005).

By looking at how such credit scoring works, an applicant for credit is evaluated in a credit scoring system by simply summing the points received on the various application characteristics to arrive at a total score. This score may be treated in a number of ways depending on the system design. In the single cut-off method, the applicant’s total score is compared to a single cut-off point score. If this score exceeds the cut-off, credit is granted; otherwise the applicant is rejected. More complex systems are based on a two-stage process. For example, the applicant’s total score may be compared to two cut-off points. If the score exceeds the higher cut-off, credit is awarded automatically, while if it falls below the lower cut-off,
credit is automatically denied. If the score is between the two cut-offs, credit history information is obtained, scored, and the points added to the total score obtained from information in the application form (some other approaches re-evaluate the applicant according to their requirements). If this new score is above a new higher cut-off, credit is awarded; if not, credit is denied, (Capon, 1982).

Altman and Haldeman (1995) have suggested steps to follow for corporate credit scoring models as follows:
- Applying primary client-data to credit scoring models
- Testing a credit scoring model
  - Definition of risk
  - Model development
  - Test of time
  - Stability
  - Public versus private company data
  - Probability of failure
  - Credibility
  - Model support
  - Pilot testing
- Using a supplemental system:
  - Firm-capital market approach (i.e. using systematic beta risk)
  - Firm-econometric approach.

So, the main thrust of this research is to develop a model (a credit scoring model) to evaluate credit risk in the banking sector in Egypt. The investigative stage of this research is to conduct informal interviews with key personnel in Egyptian banks in order to evaluate credit policies and to establish potential key factors. Results so far revealed that most banks in Egypt use judgemental techniques in their evaluation process.

By reviewing all previous lines of enquiry, in addition to the results of the pilot study that exhibited that all banks in Egypt, with which the researcher made informal interviews do not use such a model to evaluate their credit risks. The researcher is considering developing an integrated model for both consumers and small businesses to evaluate credit risks in the banking sector in Egypt, especially since credit scoring models have undergone a noticeable success in different environments in USA and Europe, taking into account all requirements for the proposed model according to the nature of the Egyptian environment.

The general idea of a credit scoring model rests on identifying some independent variables, which typically reflect the financial position of the client, using a linear equation to estimate the ability of the client to repay. The dependent variable would represent a total score. Depending on this total score, the decision-maker takes the logical decision, as to whether to grant or reject the applicant, (Motawa, 2001).

In addition, the researcher has thought about using some advanced techniques to build the proposed model, for example Excel and VBA (Visual Basic for Application), to enable the creditor to use the model in such a simple way saving effort and time.
of the credit process, [for more information see: (Sengupta, 2004) and (Jackson and Staunton, 2001)].

Literature review

Relevant literature in this research is classified into three groups:

The first group deals with credit scoring models for commercial loans/small business loans.

The second group relates to credit scoring models for consumer loans/personal loans, including credit cards and housing loans.

The third group deals with credit scoring models in general, including a comparison between different techniques used in building credit scoring models.

The first group: credit scoring models for small business/commercial loans

By reviewing studies in this group, it is clear that most of these studies used financial ratios in building the credit scoring models [see for more details: (Orgler, 1970); (Falbo, 1991); (Leonard, 1992); (Frame, et al., 2001); (Emel, et al., 2003); (Liang, 2003); and (Cramer, 2004)]. Most of this group handled the scoring models using different types of ratios and almost all the main ratios were repeated in all models with some slight differences.

While the objectives of these studies seem to be different, ostensibly, all of them led to the same result, which is to develop a credit scoring model for small business loans/commercial loans.

However, using the advanced computer programmes in building the scoring systems helped to find out and control all the selected variables in this group, (Tsaih, et al., 2004).

Some of these studies develop credit scoring models to evaluate existing loans [see: (Orgler, 1970); (Kumar and Motwani, 1999); and (Cramer, 2004)] while other studies develop scoring models to see the effect of the total discriminatory function (either discriminant or multi-discriminant) on the quality of the scoring model [see: (Falbo, 1991) and (Liang, 2003)].

One study tried to find the best model in a specific environment in Croatia, i.e., (Zekic-Susac, et al., 2004), had specific features not found in other environments. Only one study studied the effect of income-area on the scoring model and then on the small business lending activities, for both low-and moderate-income areas and moderate-and high-income areas census tract, credit scoring is generally associated with increased small business lending activities. Besides, credit scoring is associated with increased “lending out-of-market”, but often with decreased “in-market lending”, (Frame, et al., 2004).

Some of these studies used traditional statistical techniques in building the credit scoring models [see: (Orgler, 1970); (Falbo, 1991); (Leonard, 1992); (Frame, et al., 2001); (Emel, et al., 2003); (Frame, et al., 2004); and (Tsaih, et al., 2004)] except, (Zekic-Susac, et al., 2004), who depended on neural networks, comparing them with more than one conventional statistical technique. The credit scoring model can be considered a tool for allocating the time of bank officers and bank examiners in the review and evaluation of existing commercial loans, (Orgler, 1970), and easily altered to cope with changes in the business environment, (Tsaih, et al., 2004).
Statistical association measures showed that the neural network models are better representations of data than logistic regression and CART, (Zekic-Susac, et al., 2004), while discriminant analysis, in general, has better classification ability but worse prediction ability, whereas logistic regression has a relatively better prediction capability, (Liang, 2003).

Thus, it can be said that information technology reduces information costs and asymmetry between borrowers and lenders, using a simultaneous equation model, (Frame, et al., 2001). Some results suggest that applications of credit scoring techniques to areas of credit other than consumer credit may prove to be quite fruitful, (Leonard, 1992).

Finally, all studies in this group that compare credit scoring models using different statistical techniques find no essential differences between them [see: (Emel, et al., 2003), (Liang, 2003); (Cramer, 2004); and (Zekic-Susac, et al., 2004)].

The second group: credit scoring models for consumer/personal loans

By having a look at this group, it was found that some of these studies focused on consumer loans [see: (Orgler, 1971); (Steenackers and Goovaerts, 1989); (Kim and Sohn, 2004); and (Sarlija, et al., 2004)]. Whilst other studies focused on credit scoring models for credit cards [see: (Boyes, et al., 1989); (Greene, 1998); (Banasik, et al., 2001); and (Lee, et al., 2002)]. One study investigated housing loans as a type of consumer loan, (Lee and Chen, 2005). Another study discussed the credit scoring efficiency and the development of a scoring effectiveness measurement tool using some criteria for the evaluation process on consumer loan, (Leonard, 1995).

Besides these, two studies from the consumer loans focused on existing consumer loans [see: (Orgler, 1971) and (Kim and Sohn, 2004)].

In this group, like the first group, the objectives of these studies seem to be different but all of them gave rise to the same result which is to develop a credit scoring model for either consumer loans or credit cards.

Most of these studies depended on conventional statistical techniques, such as discriminant analysis, regression analysis, Probit analysis and logistic regression in building the scoring models [see: (Orgler, 1971); (Boyes, et al., 1989); (Steenackers and Goovaerts, 1989); (Greene, 1998); (Banasik, et al., 2001); and (Sarlija, et al., 2004)]. By contrast the others depended on modern/advanced statistical techniques, such as neural networks and hybrid models, in building the scoring models [see: (Lee, et al., 2002); (Kim and Sohn, 2004); and (Lee and Chen, 2005)] including two studies which made a comparison between traditional and advanced statistical techniques [see: (Lee, et al., 2002) and (Lee and Chen, 2005)].

Most studies in this group depended on case studies either of a bank [see: (Leonard, 1995); (Banasik, et al., 2001); (Lee, et al., 2002); and (Lee and Chen, 2005)] or financial institution, (Boyes, et al., 1989), or credit company [see: (Steenackers and Goovaerts, 1989) and (Greene, 1998)]. However, just one study depended on two banks, (Orgler, 1971).

Almost all studies used similar variables, with some differences regarding the nature of data or in the number of variables from eight variables in (Sarlija, et al., 2004) to twenty five variables in (Banasik, et al., 2001). Some other studies used grouped variables [see for example: (Greene, 1998) and (Boyes, et al., 1989)].
Most studies in this group used two groups of customer credits either “good” or “bad” [see: (Orgler, 1971); (Boyes, et al., 1989); (Banasik, et al., 2001); (Lee, et al., 2002); and (Kim and Sohn, 2004)]. While two studies used three groups of customer credits, one of them used either “good” or “bad” or “refused” , (Steenackers and Goovaerts, 1989), and another study used either “good” or “poor” or “bad”, (Sarlija, et al., 2004).

The credit scoring models are designed for practical use. Hence, it is important to determine if they are economically justified. The cost of operating a credit scoring system, once the model has been developed, is quite small, especially in banks that manage their consumer loan portfolio with a computer. While it is relatively easy to estimate this cost, it is very difficult to predict the benefits resulting from the use of a credit scoring model, (Orgler, 1971).

When the scoring model includes both effectiveness and efficiency, the scoring quality index, which includes some criteria for use in the evaluation process, will provide a direct assessment of performance, (Leonard, 1995). Furthermore, accounting for the selection nature of the historical data will be important. Acceptance/rejection is based on extremely simple easily explained and justified criteria, (Greene, 1998).

The neural network model has the highest average correct classification rate in comparison with discriminant analysis and logistic regression, although results are very close, approaches. Besides, the designed hybrid model not only has better accuracies, but also has the lowest type II error associated with high misclassification costs, (Lee, et al., 2002). Almost the same results were found by (Lee and Chen, 2005) which is that the hybrid model has the best average correct classification rate in comparison with discriminant analysis, logistic regression, MARS and BPN, but with very small differences. Besides, the hybrid model in this study still has the lowest not only type II error but also type I error in comparison with the other statistical techniques.

The third group: credit scoring models in general
By reviewing this group of studies, it was found that this group differs from the other two groups in terms of purpose and nature of studies5.
Most of these studies focused on comparing different statistical techniques used in credit scoring [see: (Guillen and Artis, 1992); (Desai, et al., 1996); (Hand and Henley, 1997); (Baesens, et al., 2003); (Chen and Huang, 2003); and (Ong, et al., 2005)]. Such comparisons were between prevailing/popular statistical techniques and most recent statistical techniques in some cases [see: (Baesens, et al., 2003) and (Ong, et al., 2005)].
Most studies in this group discussed the scoring models, taking into account the two classes of clients which are “good” and “bad” [see: (Guillen and Artis, 1992); (Guillen and Artis, 1992); (Desai, et al., 1996); (Hand and Henley, 1997); (Baesens, et al., 2003); (Chen and Huang, 2003); and (Ong, et al., 2005)].

This group of studies discussed credit scoring in general, taking into account the limitations and restrictions that may be faced in building scoring models and the accuracy of the models with a comparison between different statistical techniques used in building the scoring models regardless of the nature of these techniques, either recent, conventional, or even unpopular techniques.
One study focused on a credit union and the effect of new statistical techniques in such scoring models, (Desai, et al., 1996). Another study focused on credit scoring performance using the Total Quality Management (TQM) concept, (Leonard, 1996). Some other studies reviewed different statistical techniques that are used in scoring models [see: (Hand and Henley, 1997) and (McGrath, 2003)]. The genetic algorithm, as one of the most recent techniques used in the scoring models, used twice in this group of studies [see: (Desai, et al., 1996) and (Ong, et al., 2005)]. Some of these studies discussed some unpopular techniques in scoring models, such as, The Rorschach Comprehensive System and The Kolmogrov-Smirnov Curve [see: (McGrath, 2003) and (Yang, et al., 2004)].

The possibility of establishing a two step procedure for prediction has been stressed. In fact, this combines the use of discriminant analysis and modelization and it leads to a good global classification rate keeping the number of bad clients accepted as good well below 10%, (Guillen and Artis, 1992).

In an effort to follow TQM principles to the credit scoring industry, the training and continuous improvement programmes must be goal-oriented, and in order to assist in the establishment of these goals, benchmarking must be introduced. For business in general and the credit scoring industry in particular, the need for training and education is becoming more and more apparent, (Leonard, 1996).

In general, there is no overall best statistical technique method used in building credit scoring models, what is best depends on the details of the problem, the data structure, the characteristics used, the extent to which it is possible to separate the classes by using those characteristics, and the objective of the classification, (Hand and Henley, 1997). All studies that made a comparison between different techniques found that, first, most recent/advanced statistical techniques are better than the traditional ones; second, there is no apparent difference between different statistical techniques in terms of the percentage of correct classification or hit rate.

And sometimes this depends on the original group that is used to compute the correct classification (depending on bad or good and bad together), (Desai, et al., 1996).

When considering both correct classification percentage and area under the curve performance measures, it can be concluded that the non-linear LS-SVM and NN classifiers yield a very good performance. However, the more simple classification techniques, such as linear discriminant analysis and logistic regression, also have a very good performance, which is in majority of the cases not statistically different from that of the LS-SVM and NN classifiers, (Baesens, et al., 2003).

Unfortunately, there is no evidence concerning the degree to which the practice of psychological assessment is flawed by inaccuracies in test administration and scoring. No standardized observational measures are immune to potential misuse,
and there is an ethical obligation to ensure that test administration and scoring are conducted fairly and accurately, (McGrath, 2003).

But, in some cases, there is no better approach than classifying randomly, given that one knows the ratio of good to bad in the whole population, (Yang, et al., 2004). Finally, genetic programming outperforms other models. Furthermore, artificial neural networks and logistic regression can also provide the satisfied solutions and can be an other alternative. Compared with other popular models, genetic programming is more suitable for credit scoring models. The discriminant function, which is derived by genetic programming, can provide a better forecasting performance than the other algorithms, (Ong, et al., 2005).

Lastly, by reviewing the previous discussion in the three groups, the researcher intends to develop two or three models. One will be for the small business, the second for the personal loans and the last for credit card purposes, at the same time three statistical techniques will be used: discriminant analysis, logistic regression and one of the neural networks models as some of the most popular and successful techniques apparent in the literature, especially in such a new environment that has specific features.

The chosen environment will be the Egyptian banking sector, in which no studies have investigated the use of sophisticated statistical appraisal techniques in credit scoring. Indeed, from the review of literature to date, no studies were found in Egypt, or in the rest of the Middle-East countries, except for one study in Qatar, (Al Amari, 2002). Therefore, the researcher intends to cover that gap, which was found in the Egyptian banking sector.

Methodology

The research importance can be summarized as follow:

- There are no Arabic studies (to the extent of the researcher’s knowledge) dealt with statistical techniques in the field of evaluating credit risks using a model such as a credit scoring model.
- This research contributes to clarifying the concepts related to credit risk evaluation.
- This research contributes to determining the requirements of credit scoring in the field of credit risk evaluation.
- This research contributes to developing an integrated theoretical framework to credit risk evaluation for those interested in this important field.
- This research contributes to providing the decision makers in the banking sector with a suitable methodology for making logical credit decisions.

---

6 The number of models that will be developed depends on the availability of the data, in other words if the data are available for all the three models, then three models will be developed; otherwise, the researcher will restrict the breadth of the study.

7 One PhD thesis was found, which focused on the role of credit scoring in credit evaluation in Qatar.
Therefore, the main concern of this research is that “there is no integrated model for evaluating credit risk in banking sector in Egypt despite the rapid increase in funds-size invested as credit granted by these banks”.

So, the objectives of current research are to:
- Identify the currently used techniques in evaluating credit risk in Egyptian banks then, evaluate these techniques.
- Develop a proposed model to evaluate credit risk in Egyptian banks.

A credit scoring model is the proposed model to evaluate credit risk taking into account all relevant factors in the Egyptian environment, which clearly appeared from the pilot study.

The hypothesis of this study:
The core hypothesis of this research is that: “H₀: there is no significant difference in the effectiveness of alternative credit scoring techniques applied to Egyptian banks”.

The proposed models:
The researcher will depend on both traditional/conventional and modern/recent statistical techniques used to build credit scoring models which are:
- Discriminant analysis
- Logistic regression
- Neural networks, as follow:

Discriminant Analysis:
Discriminant analysis is a statistical technique which allows the researcher to study the differences between two or more groups of objects which respect to several variables simultaneously, (Klecka, 1980). The following figure (2) exhibits the relationship between groups and discriminatory variables:
Logistic Regression:

Regression methods have become an integral component of any data analysis concerned with describing the relationship between a response variable and one or more explanatory variables.

It is important to understand that the goal of an analysis using logistic regression is the same as that of any model building technique used in statistics.

What distinguishes a logistic regression model from the linear regression model is that the outcome variable in logistic regression is binary or dichotomous.

This difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. Once this difference is accounted for, the methods employed in an analysis using logistic regression follow the same general principles used in linear regression, (Hosmer and Lemeshow, 1989). If a simple linear regression equation model using weighted least squares violates the constraints on the model or does not properly fit the data, it can be said that it may need to use a curvilinear model. The simple logistic regression model can easily be extended to two or more independent variables. Of course, the more variables, the harder it is to get multiple observations at all levels of all variables. Therefore, most logistic regressions with more than one independent variable are done using the maximum likelihood method, (Freund and William, 1998).

On theoretical grounds it might be supposed that logistic regression is a more appropriate statistical tool than linear regression, given that two discrete classes “good” and “bad” have been defined, (Hand and Henley, 1997).
Neural Networks:
The study of Neural Networks (NNs) is an attempt to understand the functionality of the brain. In particular it is of interest to define an alternative “artificial” computational form that attempts to mimic the brain’s operation in one or a number of ways. In the last few years interest in the field of NNs has increased considerably, due partly to a number of significant break-throughs in research on network types and operational characteristics, but also because of some distinct advances in the power of computer hardware which is readily available for net implementation.

Neural network elements:
Neural networks, in general, consist of a number of simple node elements, which are connected together to form either a single layer or multiple layers. The relative strengths of the input connections and also of the connections between layers are then decided as the network learns its specific task(s).

The basic node elements employed in NNs differ in terms of the type of network considered. The following figure shows one commonly encountered model. In this, the inputs to the node take the form of data items either from the real world or from other network elements, possibly from the outputs of nodes in a previous layer.

![Basic neuron model](image-url)

Figure (3): Basic neuron model
Source: (Irwin et al., 1995), p. 3.

The output of the node element is found as a function of the summed weighted strength inputs.
Neural networks are adaptive nonlinear systems that adjust their parameters automatically in order to minimise a performance criterion. This obviously links them to adaptive and optimal control.

The proposed variables of this study:
The researcher figured out all proposed variables from two sources: first, are the literature reviews, second, is the pilot study that the researcher conducted as informal interviews with the key persons in Egyptian banks in order to make out the current used variables which differ from the literature variables in a part of it. The researcher intends to use both types of variables in building the proposed scoring models, either in the proposed consumer credit scoring model or the small business credit scoring model.
Variables classified into three sections as follows:
First, some variables which appropriate to personal loans, including credit cards include: name, age, time at the present address, telephone, owning a credit card, etc.
Second, another group of variables of this type which are appropriate to small business scoring models and the most suitable variables, as the literature has stressed, are financial ratios, such as: gross profit, current ratio, inventory turnover, gross profit margin, cash ratio etc.
Third, all new variables, which were not mentioned in the previous two types and the pilot study, differ from the usual variables used in the literature such as: good personal reputation of the client, guarantee, investigation report, field visit of the client etc. [For more details about all study variables see for example: (Steenackers and Goovaerts, 1989); (Emel, et al., 2003); (Liang, 2003); (Sarlija, et al., 2004); (Zekic-Susac, et al., 2004)].

Data collection of this study:
The research sample comprises all banks working in Egypt, which the researcher will focus only on the main head offices (a total of 59 banks). The research will be limited to two samples, first, commercial banks, which are 28 banks, second, business and investment banks, which are 31 banks.
Data collection sources:
- Egyptian banks (the historical data, specially related to credit processes);
- Datastream; and
- Bankscope.

and perhaps also a questionnaire, because the proposed models may include some qualitative data.
Data can be described as qualitative or quantitative. Qualitative data are concerned with qualities and non-numerical characteristics, whilst quantitative data are all data that are collected in numerical form. And qualitative data can be classified as discrete or continuous. Discrete data can take only one of a range of distinct values; for example, the number of employees. On the other hand, continuous data can take any value within a given range, such as time or length (Collis and Hussey, 2003)].

Pilot study:
The researcher conducted informal interviews with the credit department managers in three banks which are regarded as the top banks in Egypt, in order to evaluate

---

8 Egyptian banking system structure: (1) Commercial Banks (28 banks): Public Sector Banks (4 banks) and Private and Joint Venture Banks (24 banks). (2) Business and Investment Banks (31 banks): Private and Joint Venture Banks (11 banks) and Branches for Foreign Banks – Off-Shore Banks - (20 banks). In addition to Specialized Banks (3 banks): The Egyptian Industrial Development Bank, The Arab Egyptian Real Estate Bank and Principal Bank for Development and Agriculture Credit. Egyptian banks abroad are not included, also two banks established under private laws and are not registered with CBE; namely, Arab International Bank, and Nasser Social Bank.
credit policies, and to establish potential key factors, during the period between 22/12/03 and 05/02/04, and the interviews took the shape of open questions with the credit department managers.

This pilot study so far revealed that those banks in Egypt used judgemental techniques in the evaluation to the credit process. Furthermore, this process depends almost exclusively on the decision of the evaluator (credit department team, especially the department manager).

Conclusion

As it mentioned before this research aims to identify the currently used techniques in evaluating credit risk in Egypt’s banking sector, then to evaluate these techniques and to develop model(s) to evaluate credit risk in Egyptian banks. In a pilot study, the researcher conducted informal interviews with key personnel in three of the Egyptian banks, in order to evaluate credit risk policies.

Expected results in this research can be classified into two sections as follows: First, the expected results relate to expected statistical techniques: The pilot study showed that these banks in Egypt do not use any of the statistical techniques in the evaluation process. So, the researcher expects that there will be differences between the existing techniques and the proposed techniques in terms of the results of the evaluation process, which will be reflected in the prediction ability in the future, after the implementation of the proposed techniques.

Second, the implementation of the research: At the final stage of the research an investigation will be conducted, by means of a questionnaire survey of Egyptian banks, into attitudinal aspects of implementation of sophisticated techniques. Logistic regression will be run in order to attempt to establish any significant link between the degree of predisposition to implementation and other banks characteristics, for example, size and number of branches.

References


9 The informal interviews took place in the following three banks: Egyptian National Bank, Misr Bank and Misr International Bank (MIBank).
10 Especially the governmental banks, because the researcher could not accomplish enough meetings with the credit manager in the private banks (just MIBank).
11 Except one of the three banks seemed to use, in a part of its evaluation, a statistical technique which related to a CRR scoring sheet.
12 Discriminant analysis and NNs may be run beside the logistic regression.


Emel, A. B. et al. 2003. A credit scoring approach for the commercial banking sector. Socio-Economic Planning Sciences, 37, 103-123.


THE OPTIMAL CONTROL THEORY MODEL FOR OPTIMAL DIVIDEND DETERMINATION

Rumbi Mukonoweshuro

1.1 Abstract

Inspired by previous empirical evidence, which succeeded in finding a general formula for the optimal amounts to pay out as dividends to the companies’ shareholders, this PhD study for the first time in literature utilises the preliminary conclusions presented by Dalton and Pointon (1999), to device tailor made models which are specific to particular individual industrial sectors and regions in the world. This is in order to enhance the dividend prediction process. Dalton and Pointon (1999)'s preliminary results demonstrated that dividend policies varied between countries, each country has unique dividend policy factors that are applicable to its environment.

The key aim of this PhD research project is to integrate and synthesize knowledge on dividend policy; in turn this will improve the dividend determination process.

The new derived dividend strategy models produced by the PhD will possess the ability to determine the average strategic dividend payout pertinent to the organisations.

Currently the initial empirical tests required by this PhD study are being conducted and the findings will be presented at the postgraduate symposium.

1.2 Introduction

As indicated above work is being conducted to establish the effectiveness and efficiency of the optimum control theory framework that is presented in literature. The work mentioned above constitutes ‘Part A’ of a three stage PhD study. The PhD study is being conducted in three stages as follows:

1. ‘Part A’ tests in depth the accuracy of the control theory framework for dividend determination as presented in literature. This section highlights the key weaknesses of the control theory framework with regards to dividend determination and establishes the fundamental strengths of the framework. This segment will indicate vital dividend factors which should be incorporated to enhance the framework.

‘Part A’ is in response to the call made by Davidson (1980), when he called for empirical examination of the optimal control theory framework to establish if
the framework is a good explainer of intertemporal divided change. Therefore, this paper is vital, because it is in line with work that has been identified as essential for the dividend determination process.

2. Part B of the PhD study generates and presents new dividend models for determining optimum dividends applicable to the current global environment.

3. Empirical examination of new models generated in ‘Part B’ of the PhD research study will be tested in ‘Part C’ of the study. Therefore the third and final segment of this study will test the effectiveness and efficiency of the new models.

The key motivation for the PhD being pursued is to bring dividend determining models line with the changes occurring across the globe. For example research by Hofstead over the years has clearly shown the existence of cultural differences globally, and this affects dividend policy. Dalton and Pointon (1999) have proved that dividend policies are varied in the world. In recent years organisations have become large and complicated therefore it is important for the dividend models to move in step with these changes.

This paper starts by reviewing relevant literature on dividend policy. The optimal control theory framework for intertemporal dividend change is presented. The methodology used to conduct this study is detailed. Results, recommendations and conclusions of the project are presented.

2.1 Literature Review

A number of researchers have advanced theories that determine concurrently the optimum amount to pay out as dividend and the appropriate times when payments should be made to shareholders. Some researchers have provided empirical evidence regarding determinants of individual firms’ dividend policies. For example, Cadenillas et al (2003) have successfully modelled dividend policy as a stochastic control problem. However, according to, Collins, Saxena and Wansley, (1996) the optimal dividend policy issue still remains unresolved, despite all the research that has been done on this particular subject area.

Around the middle of the previous 19th century, Miller and Modigliani (1961) claimed that dividend policy was irrelevant to share valuation in a perfect tax-free capital market place, but the Miller-Modigliani assumptions are violated in the real world’s financial markets. The research that followed Miller and Modigliani (1961), has examined various real-life market imperfections in order to explain the relevance of dividend policy and to identify the optimum dividend policy.

If we consider practising managers seeking assistance on dividend policy, Davidson (1980) states that Modiglian and Miller’s (MM) (1961)’s irrelevant viewpoint will not provide any significant support. This is because as stated above (MM) concluded that in a perfect market place in this world, where all the information is known, dividends do not affect the share valuation. Therefore, in the real imperfect world dividends are relevant to share price valuation. In order to assist practising managers, effective and efficient dividend determination tools are fundamental in enabling wise decisions to be attained. This is why study is necessary and relevant,
because it evaluates one of the key currently used dividend determination process and builds on that to produce enhanced models.

Literature has identified several factors which influence the dividend policy decision. The following have been acknowledged in literature as key factors which influence dividend policy in an imperfect real-world environment and the factors include:

1. Stability of dividends: Management tend to prefer stable dividends and avoid risking an increase in dividends, that may have to be reduced in the future, Lintner (1956), Rutterford, (1994), Gill and Green (1993).
2. Dividend payout ratio: Adedeji (1998), found a positive relationship between dividend payout ratios and debt financing in the UK, this was mainly because firms with larger proportions of debt financing are more financially risky and investors should demand higher dividend. Khan and Jain, (1996) identify dividend payout ratio as one of the key factors influencing the dividend policy decisions.
3. Legal taxes and contractual factors: The treatment of dividend tax and the capital gains may be important in dividend policy effects on share prices. Chui, Strong and Cadle (1992) discovered that there were some significant tax effects between periods, when account was taken of an insufficiency of taxable profits to benefit from existing tax complexities. Therefore the hypothesis that tax coefficients were similar across three different tax regimes was rejected.
4. Stockholders' considerations (investor's preferences).
5. Inflation: Undoubtedly inflation has a significant influence on dividend policy.
6. Earnings retention ratio: According to Adedeji (1998), retentions tend to be used to finance investment and dividend payout ratios are negatively associated with capital investment in the UK. Myers(1998)'s pecking order theory supports Adedeji (1998)'s pecking order theory that, firstly firms prefer to finance investment from retained funds, secondly, if funds are inadequate, debt finance is considered and finally, as a last resort external equity finance from issue of shares is considered.
7. Level of earnings (profit): Firms that exhibit a strong record of dividends and earnings, subsequent dividend reductions tend to be associated with a loss reported in the organisation's financial statements.
9. Actual vs. target capital structure
10. Agency Costs and Floatation costs: Dewenter and Warther (1998), concluded that firms that experienced dividend omissions or resumed dividends following an omission of dividend payments, the share price effect was smaller for Japanese firms than US firms. Dewenter and Warther (1998) argued that in Japan there was a close connection between managers and investors hence the investors have an active role in the management of the organisation hence resulting in a lower agency costs. US firms tend to have a higher degree of managers not acting fully in the shareholder's best interests hence high agency cost. Jensen (1986) defines the meaning of agency costs.
11. Signalling, Benartzi, Michaely and Thaler (1997) concluded that firms
that increases their dividend are less likely to find that their future
earnings are reduced. This confirms Lintner’s study (1956).

12. Cultural effects: The study carried out by Dalton and Pointon (1995),
gives some evidence that cultural differences do provide some
explanation of the varying importance of factors for different countries.

13. Takeovers: Dickerson, Gibson and Tsakalotos (1998), discovered that
firms that pay out larger dividends are less likely to be taken over.

With regards to dividend strategy there is a trade-off between the desire to pay a
dividend as high as possible and the need to maintain liquidity in order to reduce the
default risk, Cadenillas, Sarkar and Zapatero (2003). Therefore the models that
derive the optimal dividend policy have to keep a balance between these two
factors.

An excellent survey of stochastic models for the optimal dividend policy can be
found in the work carried out by Taksar (2000). Asmussen and Taksar (1997), and,
Højgaard and Taksar (1999) study the optimal dividend policy of a company that
tries to maximize the expected value of the dividends received by the stockholders.
According to Cadenillas and Haussmann (1994) many studies have modelled optimum
dividend policy as a stochastic singular control problem, while other studies have
applied the mixed classical-singular stochastic control problem. Jeanblanc-Picque
and Shiryaev, (1995) and Cadenillas, Sarkar and Zapatero (2003), model the optimal
dividend strategy as a stochastic impulse control problem specific to a company.
With regards to this study, the optimal control theory modelling technique will be
used to devise dividend strategy models specific to two distinctive industrial sectors
(telecommunications and financial services sectors) and to four individual countries (USA, UK, Japan and Malaysia).

This study models optimal dividend policy models through application of a
sophisticated optimal control theory technique.

Comparison of the effectiveness, efficiency and usefulness of the new dividend
policy models produced by this study be critically examined an conclusions drawn.

1.1 The control theory framework (The model outlined)

Two key managerial policy functions used in the dividend model are:

1. \(U(D(t))\) = Dividend utility function

2. \(W(A(T))\) = weighting function for terminal liquid assets [weighted
stock of liquid assets at time \(T\)]

Assuming no external injection of capital the dividend control model using the above
mentioned functions can be represented symbolically as:
\[
\text{Max} \left[ \int_0^T U(D(t)) e^{-\rho t} \, dt + W(A(T)) e^{-\rho T} \right]
\]

Managerial time preference rate \( \rho = \text{rho} = \text{cost of capital} = ke \)

\[ \text{Dividend utility function} \]

\[ \text{Weighting function for terminal liquid assets} \]

**NB: The above model is subject to the following conditions:**

\[ \dot{A}(t) = \delta A(t) + Y(t) - D(t) \] \hspace{1cm} \text{equation of motion} \hspace{1cm} \text{................................................(2)}

\[ A(0) = A_0 \] \hspace{1cm} \text{initial asset position} \hspace{1cm} \text{..................................................(3)}

\[ D(t) = 0 \] \hspace{1cm} \text{non-negativity of dividends} \hspace{1cm} \text{..............................................(4)}

\[ A(T) = A_T \] \hspace{1cm} \text{terminal liquid assets} \hspace{1cm} \text{..................................................(5)}

The adaptive expectation hypothesis presents dividends depending on long run liquid assets and not long run earnings. The picture at the end of time \( T \) has the following condition.

\[ U'(D(T)) = W'(A(T)) \] \hspace{1cm} \text{transversality condition} \hspace{1cm} \text{.................................................(6)}

The transversality condition is the point when the utility function of dividends crosses the terminal line.

At the terminal point:

The gradient of the dividend utility function = The gradient of the weighting function of liquid assets.
When \( t = 0 \)

Change in \( A(t) \) = \( \dot{A}(t) = \frac{dA(t)}{dt} \)

Considering time value: when \( t = 1 \): it can be estimated that

Change in \( A(t) \) = \( \dot{A}(t) \approx \frac{dA(t)}{dt} \) \( U(D(t)) \) = concave graph

The above condition holds because of the relationship which exists between dividends and long run liquid assets as in the adaptive expectations hypothesis.
Therefore:

\[ U' (D(T)) = W' (A(T)) \] \( \ldots \) \ldots \ldots \ldots . \text{transversality condition} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots (6) \]

\textbf{Using the optimal control theory to find the path for dividends through time (t) \ldots (7)}

Finding a stream of liquid assets and dividends will utilize the literature on dynamic optimization as follows:

The literature on optimal control theory gives us the equation \( \frac{dU}{dC} = U' C(t) = Y_0 e^{(\rho - \delta)t} \) which says:

\[ \frac{dU}{dC} = U' C(t) = Y_0 e^{(\rho - \delta)t} \]

Therefore:

\[ 'd^2 U = d \frac{dU}{dC} = d \frac{dU}{dt} . d(t) \]

\[ (dC)^2 \frac{dC}{dt} \]

Considering:

\[ 'd / dt . \frac{dU}{dC} = d / dt . Y_0 e^{(\rho - \delta)t} = (\rho - \delta) Y_0 e^{(\rho - \delta)t} \]

but:

\[ \frac{d}{dt} . \frac{dU}{dC} . dt / dC = \left[ \frac{d}{dt} . \frac{dU}{dC} \right] / dC / dt \]

Therefore:

\[ \left[ \frac{d}{dt} . \frac{dU}{dC} \right] / dC / dt = \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / (dC / dt) = d^2 U / (dC)^2 \]
Change $C$ to $D$

Therefore the above becomes:

$$\left[ \frac{d}{dt} \left( \frac{dU}{dD} \right) \right] / \frac{dD}{dt} \ = \ \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / (dD/dt) = d^2 U / (dD)^2$$

and

$$d^2 U / (dC)^2 = \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / (dC / dt)$$

So

$$d^2 U / (dD)^2 = \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / (dD / dt)$$

Since:

$$d^2 U / (dD)^2 = \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / (dD / dt)$$

Therefore:

$$\frac{dD}{d(t)} = \left[ (\rho - \delta) Y_0 e^{(\rho - \delta)t} \right] / [d^2 U / (dD)^2]$$

Now let:

$$\frac{dU}{dD} = U_D$$

and
\[ \frac{d^2U}{(dD)^2} = U_D \]

But:
\[ \frac{dU}{dD} = Y_0 e^{(p-\delta)t} = U_D \]

Therefore;
\[ \frac{dD}{d(t)} = \frac{(\rho - \delta) Y_0 e^{(p-\delta)t}}{d^2U/(dD)^2} = U_D (\rho - \delta) / U_{DD} = D(t) \]

\[ \frac{dD}{d(t)} \text{ is the path for dividends} \]

**Detailed Analysis of the two policy functions**…
...and the application of the specific control theory model to the utility of dividend function

a). Both components of the policy function which are:

\[ U(D(t)) \text{ and } W(A(t)) \]

Are assumed by literature to display decreasing marginal utility, therefore with regards to dividends, the additional dividend utility will decrease with time as utility of dividends generally increases through time.

With regards to the signaling context for dividends \( D(t) \), the low dividends will give adverse (unfavorable) signals which might result in the managers’ positions being threatened or a threat of a take over can result.

The liquid assets position determines:

1. technical solvency (liquidity state of the organization)
2. level of future dividends


The utility policy functions that are used by the dividend control model mentioned above are from a family of concave utility functions as shown below;

1). The utility of dividends curve: ..................................................(8)

\[ U(D(t)) = \log D(t) \] \hspace{1cm} when \( \varepsilon = 1 \)

And

\[ U(D(t)) = (1-\varepsilon)^{-1} D(t)^{1-\varepsilon} \] \hspace{1cm} when \( 0 < \varepsilon < 1 \)

2). The weighted Assets function.................................(9)

\[ W(A_T) = b (1 - \eta)^{-1} A_T^{1-\eta} \] \hspace{1cm} when \( 0 < \eta < 1 \)

And
\[ W(A_T) = b \log A_T \] ...............................................When \( \eta = 1 \)

Please note that:

\[
\text{Epsilon} = \epsilon = - \left[ \frac{d \log U'(D(t))}{d \log D(t)} \right] = - \left[ \frac{U''(D(t)) \cdot D(t)}{U'(D(t))} \right]
\]

And

\[
\text{Eta} = \eta = - \left[ \frac{d \log W'(A(T))}{d \log A(T)} \right] = - \left[ \frac{W''(A(T)) \cdot A(T)}{W'(A(T))} \right]
\]

Where \((\epsilon, \eta)\) are constant

Proving that:

\[
\text{Epsilon} = \epsilon = - \left[ \frac{d \log U'(D(t))}{d \log D(t)} \right] = - \left[ \frac{U''(D(t)) \cdot D(t)}{U'(D(t))} \right]
\]

Simplifying \(\epsilon\) in two stages……..

Considering epsilon \((\epsilon)\) stated above, is equal to

\[
\epsilon = - \left[ \frac{d \log U'(D(t))}{d \log D(t)} \right] / \left[ \frac{d \log D(t)}{d \log D(t)} \right]
\]

Stage 2  
Stage 1
Simplifying \( \varepsilon \) in two stages

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 'd \log D(t) )</td>
<td>( \varepsilon = -[d \log U' (D(t))] / [d \log D(t)] )</td>
</tr>
<tr>
<td></td>
<td>but from stage 1:</td>
</tr>
<tr>
<td></td>
<td>( y = \log D(t) )</td>
</tr>
<tr>
<td></td>
<td>Therefore:</td>
</tr>
<tr>
<td></td>
<td>( d \log D(t) = dy )</td>
</tr>
<tr>
<td></td>
<td>Therefore:</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon = -[d \log U' (D(t))] / [d \log D(t)] )</td>
</tr>
<tr>
<td></td>
<td>and</td>
</tr>
<tr>
<td></td>
<td>( \varepsilon = -[d \log U' (D(t))] / dy )</td>
</tr>
<tr>
<td></td>
<td>Let:</td>
</tr>
<tr>
<td></td>
<td>(-[d \log U' (D(t))] = www )</td>
</tr>
<tr>
<td></td>
<td>Therefore:</td>
</tr>
<tr>
<td></td>
<td>( 'dwww / dy = dwww / dx \cdot dx / dy )</td>
</tr>
<tr>
<td>stage 1 has already established ( dx / dy ),</td>
<td></td>
</tr>
</tbody>
</table>
therefore stage 2 will establish $\frac{d\text{www}}{dx}$

Let :

$X = D(t)$

And

$Y = \log D(t) = \log x$

\[
\frac{dy}{dx} = \frac{d \log x}{dx} = \frac{1}{x}
\]

therefore:

$\frac{dx}{dy} = x = D(t)$

Let :

\[
\frac{d\text{www}}{dx} = \frac{d \log U'(D(t))}{dx}
\]

NB :-

\[
\frac{d\text{www}}{dx} = \frac{df}{dx} \cdot \frac{dD}{df} = \frac{dD}{dx}
\]

where:

\[
\text{www} = D = \log f
\]

and

\[
f = U'(D(t)) = U'(x)
\]

Therefore considering the above :

\[
\frac{df}{dx} = U''(X)
\]

and

\[
\frac{dD}{df} = 1 / f = 1 / U'(D(t))
\]

Therefore:

\[
\frac{d\text{www}}{dx} = \frac{df}{dx} \cdot \frac{dD}{df} = \frac{dD}{dx} = U'' D(t) / U'D(t)
\]

\[
\varepsilon = - \frac{d \log U'(D(t))}{[d \log D(t)]} / [d \log D(t)] = - \frac{dD}{dy}
\]

Therefore :
\[- \left[ \frac{dD}{dy} \right] = - \left[ \frac{dx}{dy} \cdot \frac{dD}{df} \cdot \frac{df}{dx} \right] = - \left[ \frac{dx}{dy} \cdot \frac{dD}{dx} \right] = \left[ \frac{D(t)}{U'' D(t)} \right] / U'D(t) = \varepsilon \ldots (proof)\]

Where (\(\varepsilon\) is a constant)

And Now Proving:

\[\text{Eta} = \eta = - \left[ \frac{d \log W'(A(T))}{d \log A(T)} \right] / \left[ \frac{d \log A(T)}{d A(T)} \right] = - \left[ \frac{W''(A(T)) \cdot A(T)}{W'(A(T))} \right] / W'(A(T))\]

Simplifying \(\eta\) in two stages……...

Considering eta (\(\eta\)) stated above, is equal to

\[\eta = - \left[ \frac{d \log W'(A(T))}{d \log A(T)} \right] / \left[ \frac{d \log A(T)}{d A(T)} \right] \]

Stage 2 Stage 1
Simplifying η in two stages

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘d log A(T)</td>
<td>η = -[d log W’ (A(T))] / [d log A(T)]</td>
</tr>
<tr>
<td>but from stage 1:</td>
<td></td>
</tr>
<tr>
<td>y = log A(T)</td>
<td></td>
</tr>
<tr>
<td>Therefore:</td>
<td></td>
</tr>
<tr>
<td>d log A(T) = dy</td>
<td></td>
</tr>
<tr>
<td>Therefore:</td>
<td></td>
</tr>
<tr>
<td>η = -[d log W’ (A(T))] / [d log A(T)]</td>
<td></td>
</tr>
<tr>
<td>and</td>
<td></td>
</tr>
<tr>
<td>η = -[d log W’ (A(T))] / dy</td>
<td></td>
</tr>
<tr>
<td>Let:</td>
<td></td>
</tr>
<tr>
<td>-[d log W’ (A(T))] = www</td>
<td></td>
</tr>
<tr>
<td>Therefore:</td>
<td></td>
</tr>
<tr>
<td>‘dwww / dy = dwww / dx ∗ dx / dy</td>
<td></td>
</tr>
<tr>
<td>stage 1 has already established dx / dy, therefore stage 2 will establish dwww / dx</td>
<td></td>
</tr>
</tbody>
</table>
Let:

\[ X = A(T) \]

And

\[ Y = \log A(T) = \log x \]

\[ \frac{dy}{dx} = \frac{d \log x}{dx} = \frac{1}{x} \]

therefore:

\[ \frac{dx}{dy} = x = A(T) \]

Let:

\[ \frac{dwww}{dx} = \frac{d \log W'(A(T))}{dx} \]

NB:

\[ \frac{dwww}{dx} = \frac{df}{dx} \cdot \frac{dD}{df} = \frac{dD}{dx} \]

Where:

\[ www = D = \log f \]

and

\[ f = W'(A(T)) = W'(x) \]

Therefore considering the above;

\[ \frac{df}{dx} = W''(X) \]

and

\[ \frac{dD}{df} = 1 / f = 1 / W'(A(T)) \]

Therefore:

\[ \frac{dwww}{dx} = \frac{df}{dx} \cdot \frac{dD}{df} = \frac{dD}{dx} = \frac{W'' A(T)}{W'A(T)} \]

BUT:

\[ \eta = - \frac{d \log W'(A(T))}{d \log A(T)} = - \left[ \frac{dD}{dy} \right] \]

Therefore:

- \[ \left[ \frac{dD}{dy} \right] = - \left[ \frac{dx}{dy} \cdot \frac{dD}{df} \cdot \frac{df}{dx} \right] = - \left[ \frac{dx}{dy} \cdot \frac{dD}{dx} \right] = - \left[ A(T) \cdot W''(A(T)) \right] / W'A(T) = \eta \quad \text{...}(proof) \]
Where (\(\eta\) is a constant)

Applying (7) to (8)......
Applying the control theory to utility functions with regards to changes in dividends, this establishes a simplified path of dividends......

.....(7)....states that .......

\[
\frac{\text{d}D}{\text{d}(t)} = \left[ (\rho - \delta) Y_0 e^{(\rho-\delta)t} \right] / \left[ \frac{d^2U}{(dD)^2} \right] = \frac{U_0 (\rho - \delta)}{U_{DD}} = D(t)
\]

\[\frac{\text{dD}}{\text{d}(t)}\] is the path for dividends..........................................................(7)

From above:

\[
D(t) = \frac{\text{dD}}{\text{d}(t)} = \frac{U_0 (\rho - \delta)}{U_{DD}} = \frac{(dU/dD)}{(d^2U/dD^2)}. (\rho - \delta)
\]

Applying ....(8)...to the above
From ....(8)...... Let:

\[U(D(t)) = \log D(t)\]

Therefore

\[
\frac{\text{d}U}{\text{d}D} = U' = \frac{1}{D(t)} = D^{-1} \text{.......when } D(t) = D
\]

\[
\frac{\text{d}U'}{\text{d}D} = \frac{\text{d}}{\text{d}D} \left( \frac{\text{d}U}{\text{d}D} \right) = \frac{\text{d}U_0}{\text{d}D} / \text{d}D = (\frac{d^2U}{dD^2}) = U_{DD} = U'' = - \frac{1}{D^2}
\]

but from ............................(7)

\[
D(t) = \frac{\text{dD}}{\text{d}(t)} = \frac{U_0 (\rho - \delta)}{U_{DD}} = \frac{(dU/dD)}{(d^2U/dD^2)}. (\rho - \delta)
\]
Therefore:
\[ \frac{dD}{dt} = \frac{(D^{-1})}{-\left(1/D^2\right)(p-\delta)} = -\frac{D^2}{D (p-\delta)} \]

Therefore:
\[ \frac{dD}{dt} = D (\delta - \rho) \quad \text{where} \quad \epsilon = 1 \]

or
\[ \frac{dD}{dt} = D [(\delta - \rho)/\epsilon] \]

But \( D = D(t) \)

Therefore:
\[ \frac{dD}{dt} = D(t) (\delta - \rho) \quad \text{where} \quad \epsilon = 1 \]

or
\[ \frac{dD}{dt} = D(t) [(\delta - \rho)/\epsilon] \]

Establish \( D(t) \) ...... from ...... \( \frac{dD}{d(t)} \)

From above:
\[ \frac{dD}{dt} = D(t) (\delta - \rho) \quad \text{where} \quad \epsilon = 1 \]

Therefore:
\[ \frac{dD}{Da} = (\delta - \rho) d(t) \]
\[ \int \frac{dD}{D} = \int (\delta - \rho) \, d(t) \]

\[ [e^{\ln D}] = \left[ e^{(\delta - \rho)t + c_1} \right] \text{ Where } \ c_1 \text{ is a constant} \]

Therefore:

\[ D = e^{\log D} = e^{(\delta - \rho)t} \, e^{c_1} = A \, e^{(\delta - \rho)t} \text{ where } A = e^{c_1} \]

But \( D = D(t) \)

Therefore

\[ D(t) = A \, e^{(\delta - \rho)t} \text{ When } (t) = 0 \text{ ....... } D(0) = Ae^0 = A \]

Therefore

When \( (t) = (0) \)

\[ D(t) = D(0) \exp (\delta - \rho)t = D(0) \, e^{(\delta - \rho)t} \text{ where } \varepsilon = 1 \text{.............................(10)} \]
NB The same answer as above for \( D(t) \) and \( dD/dt \) is established when:

\[
U(D(t)) = (1-\varepsilon)^{-1} D(t)^{(1-\varepsilon)} \text{ for } 0 < \varepsilon < 1
\]

To prove this:

Let \( U = (1-\varepsilon)^{-1} D(t)^{(1-\varepsilon)} \) where \( D = D(t) \)

\[
\frac{dU}{dD} = (1-\varepsilon)^{(1-\varepsilon)} D^{-\varepsilon}
\]

\[
\frac{d^2U}{(dD)^2} = -\varepsilon D^{-\varepsilon-1}
\]

Therefore:

\[
\frac{dD}{dt} = (U_D/U_{DD}) \cdot (\rho - \delta) = \left[\frac{(dU/dD)}{(d^2U/dD^2)}\right] \cdot (\rho - \delta)
\]

From (7)

\[
= -(\rho - \delta) / \varepsilon D^{-1} = -D(\rho - \delta) / \varepsilon
\]

Therefore:

\[
\frac{dD}{dt} = D (\delta - \rho) / \varepsilon \text{ where } D = D(t)
\]

and

\[
\frac{dD}{dt} = D (\delta - \rho) / \varepsilon
\]

therefore:
\[ \int \left( \frac{dD}{D} \right) = \int \left[ \frac{e^{-\rho}}{\epsilon} \right] d(t) \]

Therefore:

\[ \ln D = \left[ \frac{e^{-\rho}}{\epsilon} \right] t + c^1 \]

\[ = e^{\ln D} = e^{\left[ \frac{e^{-\rho}}{\epsilon} \right] t} \quad e^{c^1} = D = C_2 \exp \left[ \frac{e^{-\rho}}{\epsilon} \right] t \quad \text{.....where } e^{c^1} = C_2 \]

When

\[ t = 0, \quad D(0) = C_2 e^0 = C_2 \]

therefore:

\[ C_2 = D(0) \]

Therefore:

\[ D(t) = D(0) \exp \left[ \frac{e^{-\rho}}{\epsilon} \right] t \]

b). **Detailed Analysis of the Exogenous Income Y**

but in (a). above a detailed analysis of the two policy functions....

...and the application of the specific control theory model to the utility of dividend function was conducted......

In (a). Both components of the policy function were considered.

**Now in (b). we consider Y income.** Income is assumed to exhibit a geometric growth rate ‘g’.

Therefore:

\[ Y(t) = Y_0 e^{gt} \quad \text{..................................................................................................................}(11) \]
Or

\[ \text{Change in } Y(t) = g \ Y_0 e^{gt} = gY(t) \] ..................................................(11)

Considering the Transversality condition - (Time at the end of the planning period ‘T’)

From……(6)………………U'(D(T)) = W'(A(T))

Therefore at time ‘T’:

\[ \frac{dU}{dD} = \frac{dW}{dA} \] can also be written as \[ U_d = W_A = \lambda(T) \]

where \( \lambda(T) \) is the multiplier \( \lambda(t) \) at time \( T = \text{showdown price} \)

Applying the Transversality condition to equation .....(8) and (9)......our utility functions

We note that:

\[ \frac{dU}{dD} = \frac{dW}{dA} \]

But from above earlier:

\[ \frac{dU}{dD} = (1-\varepsilon) (1-\varepsilon^{-1}) D^{(1-\varepsilon^{-1})} = D^{-\varepsilon} \]

and

The utility function \( W(A_T) = b (1 - \eta)^{-1} A_T^{(1-\eta)} \)

\[ \frac{dW}{dA} = W' = b (1 - \eta)^{-1}(1 - \eta) A_T^{(1-\eta^{-1})} = b A_T^{-\eta} \]

Therefore at ‘T’

\[ \frac{dU}{dD} = \frac{dW}{dA} = D^{-\varepsilon} = b A_T^{-\eta} \]

\[ 1/D^{\varepsilon} = b / A_T^{\eta} \]
\[ A_T^{-\eta} = b D^\varepsilon \]

From above liquid Assets at time ‘T’ are :

\[ A_T = (b D_T^\varepsilon)^{1/\eta} \] \hspace{1cm} \text{(12)}

Incorporating utility functions stated in equation (8) and (9) into the original model
The original model is :

\[
\text{Max} \left[ \int_0^T U(D(t)) e^{-\rho t} \, dt + W(A(T)) e^{-\rho T} \right] \hspace{1cm} \text{(1)}
\]

Managerial time preference rate \( \rho = \text{rho} = \text{cost of capital} = ke \)

Dividend utility function

Weighting function for terminal liquid assets
Now after incorporating utility functions stated in equation (8) and (9) it becomes

\[\text{Max} \left[ \int_0^T \left( (1-\epsilon)^{-1} D^{1-\epsilon} e^{-\rho t} \right. \right. \left. + b(1-\eta)^{-1} AT^{1-\eta} e^{-\rho T} \right] \]

Equation ... (13) above is subject to the following conditions as stated earlier which include:

\[\dot{A} = \delta A + Y - D \ldots (14)\ldots \text{this is similar to } \ldots \text{equation of motion } \ldots \text{which is } \ldots (2)\]
A(0) = A₀ \ldots \ldots \ldots \ldots (15)...this is similar to...initial asset position......which is\ldots \ldots \ldots \ldots (3)

D(T) = \left( A_T^n b^{-1} \right)^{1/\epsilon} \ldots \ldots (16)

NB: In this new model Dₜ \textbf{is NOT} always = 0 (equation ...(4) will not bind.....non-negativity of dividends ...as in ...(4)

Simplifying the equations since both changes in D and A are now known:
Firstly using change in A to establish the equation for A(t)

From equation .......(14) we note that

Change in A = Y + δA - D

Multiplying the function by \( e^{-\delta t} \) gives

\[ \dot{A} \, e^{-\delta t} = Y \, e^{-\delta t} + \delta A \, e^{-\delta t} - D \, e^{-\delta t} \]

Therefore:

\[ \dot{A} \, e^{-\delta t} - \delta \, A \, e^{-\delta t} = Y \, e^{-\delta t} - D \, e^{-\delta t} \]

Rewriting the above equation gives

\[ 'd \left( A \, e^{-\delta t} \right) / dt = Y \, e^{-\delta t} - D \, e^{-\delta t} \ldots \ldots \text{since } de^x / dx = e^x \ldots \text{so } 'd \left( A \, e^{-\delta t} \right) / dt = \]

A \, e^{-\delta t} - \delta \, A \, e^{-\delta} \]

Therefore integrating as an inverse function of a derivative

\[ \left[ A \, e^{-\delta t} \right]_0^l = \int_0^l Y \, e^{-\delta t} \, dt - \int_0^l D \, e^{-\delta t} \, dt \]
\[ A e^{-\delta t} - A_0 e^0 = \int_0^t Y e^{-\delta t} \, dt - \int_0^t D e^{-\delta t} \, dt \]

Multiply by \(e^{\delta t}\)

\[ A e^{\delta t} - A_0 e^{\delta t} = e^{\delta t} \int_0^t Y e^{-\delta t} \, dt - e^{\delta t} \int_0^t D e^{-\delta t} \, dt \]

Therefore:

\[ A = A_0 e^{\delta t} + e^{\delta t} \int_0^t Y e^{-\delta t} \, dt - e^{\delta t} \int_0^t D e^{-\delta t} \, dt = A(t) \]

\\………………… (17)\\

But:

\[ Y = Y_0 e^{\gamma t} \quad \ldots (11) \quad \text{and} \quad D = D_0 e^{(\delta - \rho) t / \varepsilon} \quad \ldots (10) \]

Substituting the above functions into equation ...(17) integrating and solving for \(t=T\), gives:

\[ A = A_0 e^{\delta t} + e^{\delta t} \int_0^T Y_0 e^{(\gamma - \delta) t} \, dt - e^{\delta t} \int_0^T D_0 e^{(\delta - \rho) t / \varepsilon} \, dt = A(t) \]

\[ A = A_0 e^{\delta t} + e^{\delta t} \int_0^T Y_0 e^{(\gamma - \delta) t} \, dt - e^{\delta t} \int_0^T D_0 e^{[(\delta - \rho) / (\varepsilon - \delta)] t} \, dt = A(t) \]

\[ A = A_0 e^{\delta t} + e^{\delta t} \int_0^T Y_0 e^{(\gamma - \delta) t} \, dt - e^{\delta t} \int_0^T D_0 e^{[(\delta - \rho) / (\varepsilon - \delta)] t} \, dt = A(t) \]

\[ A = A_0 e^{\delta t} + e^{\delta t} \int_0^T Y_0 e^{(\gamma - \delta) t} \, dt - e^{\delta t} \int_0^T D_0 e^{[(\delta - \rho) / (\varepsilon - \delta)] t} \, dt = A(t) \]

\[ A(t) = A_0 e^{\delta t} + e^{\delta t} \int_0^T Y_0 e^{(\gamma - \delta) t} \, dt - e^{\delta t} \int_0^T D_0 e^{[(\delta - \rho) / (\varepsilon - \delta)] t} \, dt = A(t) \]

\[ e^{\delta t} \left[ D_0 e^{[(\delta - \rho) / (\varepsilon - \delta)] t} / [(\delta - \rho) / (\varepsilon - \delta)] - D_0 / [(\delta - \rho) / (\varepsilon - \delta)] \right] \]
let:

\[(g - \delta) = j \quad \text{and} \quad i = (\delta - \rho)/\varepsilon\]

Therefore \(A(t)\) then becomes:

\[A(t) = A_0 e^{\delta t} + e^{\delta t} \left[ Y_0 e^{(j)t} / (j) - Y_0 / (j) \right] - e^{\delta t} \left[ D_0 e^{(i-\delta)t} / [(i - \delta)] - D_0 / [(i - \delta)] \right]\]

From the above equation \(A(T)\) is then:

\[A(T) = e^{\delta T} \left[ A_0 + Y_0 / (j) (e^{(j)T} - 1) - D_0 / (i - \delta) (e^{[i-\delta]T} - 1) \right] \] ............(18)

Rearranging the above equation enable us to establish \(D_0\)

Since equation......(18) is:

\[A(T) = e^{\delta T} \left[ A_0 + Y_0 / (j) (e^{(j)T} - 1) - D_0 / (i - \delta) (e^{[i-\delta]T} - 1) \right] \] ............(18)

Therefore:

\[A(T) e^{-\delta T} = e^{-\delta T} e^{\delta T} \left[ A_0 + Y_0 / (j) (e^{(j)T} - 1) - D_0 / (i - \delta) (e^{[i-\delta]T} - 1) \right] \]

\[A(T) e^{-\delta T} = e^{(-\delta T + \delta T)} \left[ A_0 + Y_0 / (j) (e^{(j)T} - 1) - D_0 / (i - \delta) (e^{[i-\delta]T} - 1) \right] \]

Divide all functions by ...........(e^{[i-\delta]T} - 1)
\[
A(T) e^{\delta T} / (e^{[i-\delta]T} - 1) = 1 / (e^{[i-\delta]T} - 1) \left[ A_0 + Y_0 / (j) \left( e^{(j)T} - 1 \right) \right] - D_0 / (i \cdot \delta)
\]

Multiply all functions by \( (i \cdot \delta) \)

\[
(A(T) e^{-\delta T})( (i \cdot \delta) / (e^{[i-\delta]T} - 1) = (i \cdot \delta) / (e^{[i-\delta]T} - 1) \left[ A_0 + Y_0 / (j) \left( e^{(j)T} - 1 \right) \right] - D_0
\]

Therefore:

\[
D_0 = (i \cdot \delta) / (e^{[i-\delta]T} - 1) \left[ [A_0 + Y_0 / (j) \left( e^{(j)T} - 1 \right)] - (A(T) e^{\delta T}) \right]
\]

Therefore:

\[
D_0 = (i \cdot \delta) (e^{[i-\delta]T} - 1)^{-1} \left[ A_0 - A(T) e^{\delta T} + Y_0 / (j) \left( e^{(j)T} - 1 \right) \right] \] 

\ldots (19)

From equation \ldots (18) \ldots which gives us \( A(T) \) we can establish the equation for \( A(t) \):

Since:

\[
A(T) = e^{\delta T} \left[ A_0 + Y_0 / (j) \left( e^{(j)T} - 1 \right) - D_0 / (i \cdot \delta) (e^{[i-\delta]T} - 1) \right] \] 

\ldots (18)

Therefore \( A(t) \) is:

\[
A(t) = e^{\delta t} \left[ A_0 + Y_0 / (j) \left( e^{(j)t} - 1 \right) - D_0 / (i \cdot \delta) (e^{[i-\delta]t} - 1) \right] \] 

\ldots (20)

Where \( D_0 \) is given by (19) and \( A_T \) is given by equation (12) then substituting for \( A_T \) in equation (19) and rearranging gives the following:

\[
A_T = (bD_T) \epsilon \] 

\ldots (12)
\( D_T = D_0 \exp \left[ \frac{(\delta - \rho)}{\epsilon} \right] t \) .................................(10)

Let:

\( \frac{(\delta - \rho)}{\epsilon} = i \)

Therefore:

\[ D_T = D_0 \exp [(ii) t] \quad \text{and} \quad A_T = (bD_T^\epsilon)^{1/\eta} = (b(D_0 \exp [(\delta - \rho)/\epsilon] t)^\epsilon)^{1/\eta} = (b(D_0 e^{it})^\epsilon e^{-\delta T} + Y_0/(j)(e^{(j)T} - 1)] \]

Substituting for \( A_T \) in equation (19) and rearranging gives:

\[ D_0 = (i - \delta)(e^{[i-\delta]T} - 1)^{-1} \left[ A_0 - A(T) e^{-\delta T} + Y_0/(j)(e^{(j)T} - 1)] \right] \] ..........................(19)

Therefore rearranging gives:

\[ D_0 + (b(D_0 e^{it})^\epsilon)^{1/\eta} e^{-\delta T} = (i - \delta)(e^{[i-\delta]T} - 1)^{-1} \left[ A_0 + Y_0/(j)(e^{(j)T} - 1)] \right] \]

Equation (21) is very important enables us to find \( D_0 \) when \( A_0 \) and \( Y_0 \) are known.

**Intertemporal Change of Dividend**

Change of dividend within periods (financial trading periods)

To minimize the risk of calculating a figure with high level of errors, the past research on optimal dividend considers the change in divided within financial periods instead of considering the absolute optimum dividend figures for each year end.

Previous period will be denoted as \( (\alpha - 1) \) and current period as simply \( \alpha \).
Form equation ... (19) substitute $D_0$ for $D_{(\infty - 1)}$

$$D_0 = (i - \delta) (e^{[i - \delta]T} - 1)^{-1} \left[ A_0 - A(T) e^{-\delta T} + Y_0 / (j) (e^{jT} - 1) \right] \quad \text{.........(19)}$$

Therefore equation (19) becomes:

$$D_{(\infty - 1)} = (i - \delta) (e^{[i - \delta]T} - 1)^{-1} \left[ A_0^{(\infty - 1)} - A_T^{(\infty - 1)} e^{-\delta T} + Y_0 / (j) (e^{jT} - 1) \right] \quad \text{.........(22)}$$

Therefore also $D_{\infty} = (i - \delta) (e^{[i - \delta]T} - 1)^{-1} \left[ A_0^{(\infty)} - A_T^{(\infty)} e^{-\delta T} + Y_0 / (j) (e^{jT} - 1) \right]$

Therefore change of dividend between periods is $D_{\infty} - D_{(\infty - 1)}$

Therefore:

$$D_{\infty} - D_{(\infty - 1)} = (i - \delta) (e^{[i - \delta]T} - 1)^{-1} \left[ A_0^{(\infty)} - A_T^{(\infty)} e^{-\delta T} + Y_0 / (j) (e^{jT} - 1) \right] - (i - \delta) (e^{[i - \delta]T} - 1)^{-1} \left[ A_0^{(\infty - 1)} - A_T^{(\infty - 1)} e^{-\delta T} + Y_0 / (j) (e^{jT} - 1) \right]$$

Rearranging the above equation gives:

$$D_{\infty} - D_{(\infty - 1)} = \left[ (A_T^{(\infty)} - A_T^{(\infty - 1)}) e^{-\delta T} - (A_0^{(\infty)} - A_0^{(\infty - 1)}) - (Y_0^{\infty} - Y_0^{(\infty - 1)}) \right] (e^{jT} - 1) \right] /$$
\[(i - \delta)^{-1} (e^{[i-\delta]T} - 1)\]

Assuming the following:
\[Y_0(\omega - 1) = Y_0(\omega)e^{-g}\]

and
\[A_0(\omega) = A_{(1)}(\omega - 1)\]

Therefore form equation (20)
\[A(t) = e^{\delta t} \left[ A_0 + Y_0 / (j) (e^{(j)T} - 1) - D_0 / (i - \delta) (e^{[i-\delta]T} - 1) \right] \text{............(20)}\]

Therefore \(A_{(1)}(\omega - 1)\) becomes:
\[A_{(1)}(\omega - 1) = e^{\delta} \left[ A_0(\omega - 1) + Y_0(\omega)e^{-g} / (j) (e^{(j)} - 1) - D_{(\omega - 1)} / (i - \delta) (e^{[i-\delta]} - 1) \right]\]
\[\text{.....(22 a)}\]

Substituting in (22) and rearranging gives the following:

Where \(D_{(\omega - 1)} = (i - \delta) (e^{[i-\delta]T} - 1)^{-1} \left[ A_0(\omega - 1)A_{(1)}(\omega - 1)e^{-\delta T} + Y_0 / (j) (e^{(j)T} - 1) \right]\]
\[\text{............(22)}\]

And also \(D_{\omega} = (i - \delta) (e^{[i-\delta]T} - 1)^{-1} \left[ A_0(\omega)A_{(1)}(\omega)e^{-\delta T} + Y_0 / (j) (e^{(j)T} - 1) \right]\]

Therefore:
\[D_{\omega} - D_{(\omega - 1)} =\]
\[\left[ (A_{(1)}(\omega) - A_{(1)}(\omega - 1)) e^{-\delta T} + A_0(\omega - 1)(1 - e^{\delta}) \right] + \left[ 1 / j \left[ (e^{j} - 1) + (e^{-g} - 1)(e^{jT} - 1) \right] \right]\]
\[Y_0(\omega) + \]
\[\left[ e^{\delta} \cdot (e^{[i-\delta]} - 1) / (i - \delta) \right] D_{(\omega - 1)} \] / \(\left[ (i - \delta)^{-1} (1 - e^{[i-\delta]T}) \right] \]

The above equation is of the form:
\[ D_{\alpha} - D_{(\alpha - 1)} = a + b_1 Y_{\alpha} + b_2 D_{(\alpha - 1)} \] ..........................(23)

Where :

\[ 'a' = (A^T(\alpha - 1) - A^T(\alpha )) e^{-\delta T} + A_0(\alpha - 1) (e^{\delta} - 1) ] z \] ..........................(24)

\[ 'b_1' = [1/ (\delta - g) ] (e^{(\delta g) - 1}) + (e^{g} - 1) (e^{(g - \delta)T} - 1) ] z \]

\[ 'b_2' = [ \varepsilon e^{\delta} / (\rho + \delta (\varepsilon - 1)) \cdot (e^{((\delta - \rho)/\varepsilon) - \delta}) - 1) ] z \] ..........................(26)

and

\[ z = [(\delta - (\delta - \rho/\varepsilon)) \cdot (1 - e^{[i(\delta - \rho/\varepsilon) - \delta]T} - 1) ] \] ..........................(26b)

The values of \( b_1 \) and \( b_2 \) are found to be quite small, these terms capturing mainly interest and elasticity terms. \( b_1 \) is the coefficient of the exogenous income \( Y(t) \). \( a \) shows liquid assets. \( (A^T(\alpha - 1) - A^T(\alpha )) \) is the present value of \( A(t) \). The above equations predict future dividends for organizations.

1.1 Methodology

This study will be carried out in three parts as follows:

Initially, test the current optimal control theory as described above.

Secondly, use relevant literature on dividend policy to enhance the current optimal dividend model described above. This will devise effectiveness, efficient and useful models for dividend prediction process.

Finally, the new models will be tested.

The sample for the study will be drawn from the top 50 companies in the world and from telecommunication industry and financial services industries from the following regions in the world:

USA
UK
Japan
And
Malaysia.
Approximately data from 300 companies will be analysed and evaluated by this research study.

3.2 Methodology process

Step 1: Data gathering
- Identify 300 companies to be studied. Companies will be selected from three distinctive groups, which include:
  1. Top 50 companies in the world for 2005
  2. Telecommunications industry companies. Firms will be selected from four counties (USA, UK and Japan and Malaysia)
  3. and financial services sector companies. Firms will be selected from four counties (USA, UK and Japan and Malaysia)

150 companies out of the 300 selected at random will be used to test, the current optimal dividend models while the other 150 holdout sample will be used to test the effectiveness of the new models produced.

- Collecting relevant information for analysis will be done through the use of available financial databases primarily DataStream and the internet.

Step 2: Analysis of data collected.

All relevant and appropriate analysis techniques will be utilised in this study to attain the desired new models.

Step 3: Finally, the practical usefulness of the new models produced will be tested.

4.1 Preliminary Results of the PhD

Currently the test for the effectiveness and efficiency of the current optimal control theory as presented and described in literature is being conducted and preliminary results will be presented at the symposium.

4.2 Future Plans for the PhD

- Continue to be highly motivated and to complete the PhD as soon as possible
- To establish the results for the first stage of the PhD study soon and commence all the required work within a reasonable time frame.
- To use all the opportunities that are available which will enable me to complete my PhD successfully.
- To continue to build on the literature review chapter
- Write my PhD as I go along
5. References:


FINANCE FOR SHIPPING: PAST FINANCING METHODS

Jonathan Challacombe

Abstract

Often it is stated that Shipping, (Maritime Transport), is the biggest industry in the world with only the banking and insurance sectors likely to dispute the statement. However, when two of these three sectors coincide, as in Ship Finance, a major business sector emerges which has seen both vast fortunes made and colossal sums lost.

Ships or cargo carrying craft have been in existence since man has travelled on the seas. How they have been financed is the scope of this the research. The early Egyptians and Phoenicians had great fleets which were financed by investors or merchants, (equity and debt). Through the centuries equity was the prime method with occasional borrowings. From oars to sail to steam, little changed in the method of finance, except for the occasional government subsidy. This continued to immediately after the Second World War when Onassis claimed that he first thought of the basic conception of borrowing long-term funds, for the construction of new vessels, against the security of a charter. In fact Onassis copied the tactics of the Norwegians for the financing of new buildings.

This presentation will be a review of academic and industry literature, chiefly within the European sector, to identify and clarify the methods of financing ships in the past to post second world era. This forms part of the total research of past present and future trends in ship finance.

Introduction

Often it is stated that Shipping, (Maritime Transport), is the biggest industry in the world with only the banking and insurance sectors likely to dispute the statement. However, when two of these three sectors coincide, as in Ship Finance, a major business sector emerges which has seen both vast fortunes made and colossal sums lost.

The aim of this paper is to explore the history of shipping to try and discover whether there was any mention of ship finance in history.

Review

Shipping, in this context meaning cargo carrying craft, has been in existence since man first ventured on to the seas. Locally at first within rivers and estuaries and
then further afield. This led to the beginning of trading and soon international trading followed. The craft in which the goods were carried gradually developed from the one man craft such as the ‘coracle’ to larger carrying capacity craft requiring more men to propel and operate them. This development is not the discussion of this paper but rather the simplistic question of ‘who paid for these craft and how?’

Shipping was not the result of a sudden discovery but was a gradual evolution (Kirkaldy 1914 pg 2). It probably began by being an adventure, what lies further down river or up the coast? What is beyond the line where the sky meets the sea?

Some theorists, even academics, still reluctantly allow that Egypt was the mother of ‘shipping’. This is said because until only a couple of decades ago, little was known about the Chinese history and only now are details of early life in China being revealed. That aside there exists no visible record of man’s activities in relation to water transport so definite and so ancient as to the inscriptions, the sculptures and the excavated boats in the Nile valley (Chatterton 1931 pg 1). But as Egypt was initially self supporting the Nile was in fact a large convenient internal highway, and so the first craft were bundles of reeds lashed together with tapered ends to assist the travelling across the river from side to side. Then came the dug outs, chipped out with copper adzes leading to longer trips, one way man propelled with paddles and drifted with the flow and the other wind assisted with a sail. Eventually it was the departure from the Nile Delta that started the world’s progress of international trade. The discovery of new materials to exchange with home produced commodities, the exchange of wealth and the grouping of people on the shores of estuaries and coast. It is believed that by the time that the Dynastic epochs began, Egypt had already skilled boat builders and seaman far in advance of other nations. The first dynasty was from BC3500 and in fact it is possible through the hieroglyphics to trace this early development. At this stage it can only be surmised that the craft were built and the materials paid by individuals that could do so.

There is evidence, written in the excavated tomb-temples of Snefru (3rd Dynasty) and Sahure (5th Dynasty), of fleets of vessels trading in the Eastern Mediterranean and Red Sea bringing large amounts of goods, (Lebanese hard wood and Sinai products). This is still only up to BC3200 (Chatterton 1931, pg 3). It has been deduced, assumed, that the payment for the boats and men would be from the Pharaoh’s own sources, in other words Equity! Although it is known that in this period the great Chinese Dynasties were in existence little is known about water transportation, (yet); but as with the Nile, (4,145miles) the Yangtze (3,500miles) must have had a similar nautical history. Historians await with interest for findings of early writings, such as the in the Egyptian tombs to verify this assumption.

From the early development of trade came Human greed and jealousy which, throughout history, has been the main cause for wars. The development of craft in the Mediterranean from the Nile delta was soon not for commerce but for naval fighting and this bias continued for a long time. It is known that in the period of 1500 to 1000BC there were naval wars, with ships with archers from Egypt against the Phoenicians as a result of territory and trade and this is an element throughout history. The Phoenicians were trading the whole length of the Sea of Tarshish
(Mediterranean) in this period as they were collecting silver from Spanish mines. Phoenicia became a natural staging post for Orient goods brought by land with the goods from the shores of the Mediterranean. It is even accredited that the more adventurous seafarers travelled northwards as far as Cornwall for the tin and south along the bulge of West Africa. Ships were built not just for cargo carrying but primarily for warfare with limited space to bring back goods. But it was the early Greeks who started to make the different construction for their vessels. With the Phoenicians, Egyptians and then the Romans all becoming competitors for sea supremacy, the Greeks learnt from their shipbuilding abilities and commenced to build their own designs. The first was a war vessel that was quickly built, light, mobile swift and relied almost exclusively on oars and secondly a slower, heavier tubbier craft intended for cargo carrying and relied on a sail for propulsion but carried sweeps if required when occasionally becalmed. Length gives speed and beam stability and safety so that is why the early vessels were able to trade in the open seas and for long distances. However, the Greek supremacy only lasted for awhile and the Romans soon ruled the waves as well as conquering nations on land. At this time, circa BC300, documentary evidence mentions the Roman ship-owner as a ‘mercator’. These men formed corporations and the state paid them well for their vessels to carry corn to the capital (Chatterton 1931). It was rare for men to risk their capital in shipping, which has a similar ring today! These freight vessels were said to carry approximately 280 tons of corn. Roman senators were forbidden to be in the corn freight business although they could own vessels to carry their own produce of their estates. A few centuries later Diocletian in his re-structuring of the Roman Empire caused all ship-owners to become State servants (Chatterton 1993).

Little is written about how finance was raised for ships except statements such as the mercator, as above. From various sources and literature it follows that Emperors, Kings, Queens, Rulers and Chieftains levied taxes upon their respective subjects to pay for the war fleets that were required. The great battles of the Eastern Mediterranean were hugely costly in respect of vessels and manpower. But actual ownership is hardly mentioned for the small essential cargo carrying vessels that served the various trading nations.

Mention of English fleets does not occur in Literature until the late 9th Century when King Alfred built a fleet to blockade the Danes in Northumbria which saved his kingdom of Wessex, (Stewart 1993). Because at that time Englishmen were living of the land, ships were not required in any great number. Canute had Scandinavian origins and developed a maritime trade to that area, plus he built 40 ships for his protection. William the Conqueror required ships to repel the Danes and a fleet of ships were ordered from five ports, now known as the Cinque ports which became responsible for providing a permanent contribution to naval forces. These vessels were not war ships but transported men to land battles (Stewart 1993, pg 1). It was not until Richard 1st crusade that the English fleet started to grow. He required over 100 vessels and the Englishmen learned from the Mediterranean of other designs and applied these to the latest buildings. When Henry III defeated a French squadron by an English fleet of 40 ships, this composed of royal and merchant vessels.

The presumption is that in this period it was the merchants, (traders), themselves who owned the merchant vessels. In the 14th and 15th centuries the overseas trade
flourished especially to the near continent and the Baltic. Ship types began to change from the ‘cog’ to the ‘carvel’ and then the ‘carrack’. They became better equipped to undertake voyages further afield. It was an English ship that discovered Madeira in 1344, 74 years before the Portuguese arrived there (Hakluyt 1914)!

Richard II imposed tax on goods coming to England on foreign ships and so protectionism occurred at this time. This was the first of the Navigation Laws of England which was connected with coastal trade especially the export of coal from Newcastle to the continent, the object being to restrict the trade to English ships which at the time could not be fully maintained (Kirkaldy 1914 pg 9). He ordered more ships to be built when it was soon apparent that his tax was a burden as there were not sufficient English hulls to carry the goods. Henry IV issued a law in 1402 which required all shipowners to maintain a certain number of ships at readiness for emergencies. Also ports around the coast were commanded to build ships to police the seas (Burwash 1969 pg 113). In this period it has to be stated that Venice was the great shipping centre of the world. It was the Venetian ships that undertook long voyages, increased in size and developed the modifications in rigging and equipment. These vessels traded as far north to English ports and Antwerp, as well as the whole of the Mediterranean.

At this time there was easy interchange between royal and merchant ships. Henry V enjoyed the use of a good fleet, larger in number and size from his predecessor, for the re opening of the 100 year war with France. He was able to use these for the trade with English controlled sectors of France. Henry VI lost all this French land due to Joan of Arc’s exploits and Charles II could not regain them. Henry VII, though, had a revitalised trade pattern based on the settlement of England after the ‘war of the roses’. Because of the lost land on the continent this brought a revival of the maritime trade. Increased commerce from the mills of Lancaster in woollen and cotton goods required more vessels to take these overseas, Henry VII gave a bounty, (now known as subsidies), to shipbuilders who built vessels over 100 tons (Stewart 1993, pg 4).

Any study of Merchant Shipping is fraught with difficulty because any evidence, whether pictorial or literal or even technical is more plentiful for and more detailed for fighting ships rather than for commercial. Therefore the knowledge of when the introduction of two or three masted ships into England comes from naval accounts, which contain the only continuous and detailed inventories and any inference for merchant ships, is only obtained from brief and scattered references (Burwash 1969 pg 87). The actual construction, although not in design, was used for both types of vessels with little difference.

Although still dependent on the merchant vessels to defend the realm Henry VII and Henry VIII introduced differences between the two fleets and built more royal ships. Henry VIII inherited 7 royal ships from his father and had nearly 60 naval ships in commission at the end of his reign. He was known as the Father of the English Navy as he had a great interest in shipbuilding and overseas trade. In 1513 he founded Trinity House to examine officers and supply seaman for his ships and 50 years later this body became responsible for providing navigation lights and buoys (Stewart 1993
pg 5). His flagship the Mary Rose was built in 1510, refitted in 1536. The cost of this was said to be £1016 (Burwash 1969 pg 115).

When Elizabeth 1st came to the throne in 1558 it could be said that the English Merchant fleet was at its nadir. She also only inherited a fleet of 27 naval ships in 1558. Instead of building up her own fleet Elizabeth encouraged private enterprise, in effect piracy, against Spain’s new Atlantic empire. She used men engaged in these activities, such as Sir John Hawkins and Sir Francis Drake, to command groups of Royal and private ships in attacks on the Spanish. When Spain counter attacked and threatened invasion with its Armada of 1588, the Navy of England both Royal and private was mobilised successfully to defend the realm.

The Dutch had the greatest trading fleet and outnumbered the English trading to the Baltic 20:1. The Dutch were also trading all along the Edges of Europe from the Baltic to Gibraltar. In the Mediterranean the French, Spanish and Italian fleets were large in numbers. Also the Spanish and Portuguese had their great fleets trading across the Atlantic and to the Indian Ocean. England could muster a real force in naval terms but a pitiful trading fleet. Just a trickle of ships trading along the east and south coasts with some venturing to the Baltic. All told the total tonnage was about 50,000 tons of every kind, (fishing included) (Davies 1962 pg 1).

If the Spanish Armada had come at the beginning of Elizabeth’s reign instead of in the latter part of it (and the Spaniards, Portuguese, Italians, and Germans who contributed ships and crews could have mustered a strong force at any time during this period) it would have attacked an England deficient not merely in naval strength but even more in supporting merchant ships and in crews to man either. The basis on which England established itself in 1588 as the first naval power of Europe was only created during the two previous decades (Davies 1962 pg 2).

Merchant ship-owners, as men of wealth, and companies during this period were not documented in a manner that is easy to establish identities.

On 31 December 1600, the East India Company was founded to challenge Dutch and Portuguese dominance in the spice trade. It was the major force behind British Imperial expansion throughout the seventeenth and eighteenth centuries, it was not really until the establishment of voyages made on a joint-stock basis and the granting of factory concessions, in the early seventeenth century, that the company began to boom (www.bbc.co.uk).

However during the early days of the 17th century, Amsterdam was the great centre of world trade, and Dutch shipping was to met with wherever a cargo was to be picked up. Sir Walter Raleigh wrote in the early seventeenth century about the huge trading ability of the Dutch and the number of ships they had. Even though there was not a single tree in Holland the Dutch were still able to build a thousand ships each year (Kirkaldy 1914, pg 15). However the English too were pushing out into new countries and new trades and although at one time they were no better than pirates, the English shipowner and seafarer were beginning to compete for the legitimate trade of the world (Kirkaldy 1914). As the trade increased it was felt by both politicians and shipowners that if possible the Dutch must be prevented from
encroaching on this growing trade. The English Government (Cromwell) passed the first Navigation Act in 1651. This enacted that no one in any of the ports of the Plantations of Virginia, Bermuda, Barbados and other places in America shall allow any goods from these area to be carried other than in English ships to foreign ports. This obviously helped the English and to obtain more restrictions a further act was passed which prohibited any foreign ship from trading with the American Plantations, (the colonies). In 1660, when monarchy was restored, Charles II passed the great Navigation Act which did not just cover the colonies but all goods produced in Asia, Africa or America could only be imported into England on English ships (Stewart 1993 pg.9). This in effect was to last 200 years and was the influence for the four wars with the Dutch and by the later part of the 17\textsuperscript{th} century the Dutch fleet was severely cut and London became the shipping capital of the world (Kirkaldy 1914, pg 156). The English fleet expanded and new chartered companies were formed or were privately owned, either by individual shipowners, or by a small partnership. The East India Company (EIC), was by far the greatest chartered company. The East Indiamen, as the merchant ships were called, were the largest merchantmen of their day. Heavily armed to defend their valuable cargoes, they were sometimes mistaken for warships. Only a few were owned by the EIC; most were chartered by their owners who were consortia each led by an influential merchant or gentleman known as the Husband, who had the ‘right’ to charter ships to the EIC at a most advantageous rate. The shareholders might consist of a couple of merchants, someone from nobility, the Master, the shipbuilder and the sailmaker. Involvement in the EIC was common among the British nobility and many of the officers were gentlemen; some were even titled. Command of the East Indiamen could make a captain very rich (www.lr.org).

Henry Winstanley in 1695 was at the height of his fame as an architect who devised the art of engraving prints of Palaces and noblemen houses. Known to Plymouthians as the builder of the first Eddystone lighthouse (Majdalany 1959, pg.36), he was formally the Clerk of Works to Charles II. His various enterprises had earned him a great deal of money and he had invested some of this in one of the most profitable, but risky, fields then available to a man with capital. He had bought five ships. One had been lost on the Eddystone in August and another was due to reach Plymouth at the end of the year. He was in a tavern in London when he received the news that this ship had been also wrecked on the Eddystone. After arriving in Plymouth he vowed to build a lighthouse on the rocks. After obtaining the permission and lease from trinity House he designed and eventually built the first lighthouse by 1698. This structure lasted for five years when it was then washed away in the ‘Great Storm’ of November 1703, with Winstanley, who had come back to oversee remedial work after continuous autumnal gales(Majdalany 1959, pg 64).

The trading conditions increased and in new countries and the Far East the dominance of the trading was by the English. Each trade was a separate venture and mainly trading was a series of ventures. There might be several owners of a ship and many owners of the cargo but the venture was organised as a rule to make the best profit possible according to the conditions obtaining at the terminal port. Thus on every ship there was an official called the ‘super cargo’. The very name explains his duties. He was a business man versed in commercial procedures and capable of
attending to the business of the ship in port and supervised the handling of the cargo. It is documented that sometimes the Captain of a vessel held both positions, but both the Captain and Supercargo held positions of importance in the commercial world. The success of the venture was entirely dependent on the ability of the Master and/or supercargo. The owner could only wait until the vessel returned to the home port. Traditionally the practice had been for the managing owner of a vessel to mobilize funds and reduce his exposure to risk by inviting investors, commonly family, friends, shipbuilders and suppliers, to take one or more sixty-fourth shares in the venture (Souden 2003 pg 69)

It was not until the cable had made communications possible between countries that the venture system became a past product and trade became regularised. Then there came about a revolutionary change in shipping business not caused by any sudden upheaval, but evolving with the progress of civilisation and commerce. The latter half of the 19th century was when the great change took place (Kirkaldy 1914).

When the great trade routes became open competition, lines of fine sailing ships belonging to private firms began regularly trading to India and the Far East and eventually Australia. These were private wealthy firms trading on their own capital. Ships by law were divided into 64 parts of which as long as there were no sub-divisions the 64ths may be registered in the names of any number of people from 1 to 64. Many companies grew on the reliance of Government contacts to carry mail around the world.

The Navigation acts were repealed in 1849, to allow international free trade but it was interesting to note that no other country followed Britain in this unilateral declaration of free carriage of goods and protectionism continued in other trading nations.

The tendency was for big firms to get bigger and amalgamate. At the beginning of the 19th century shipping consisted of many small companies. There were some wealthy private firms trading on their own capital. The character of shipowning changed decisively in the early part of the nineteenth century. The term ‘shipowner’ came into being as a recognised trade description (BPP,1844). The investment in shipping was now seen to be changing. From the previously traditional ‘venture’, passive investors began to drop out and vessel ownership was more concentrated in one or two shareholders’ hands. The very first specialist ship-owning businesses were emerging. A banker from the banking house Baring Brothers wrote in 1826:

“Within the last 30 years, the shipowners of London have entirely changed character; the increase of the population and extent of the metropolis has given a natural division of labour, and except in some peculiar trades, either (those) that enjoy a monopoly from charter or a monopoly from the mortgage of the West India planters land, the merchant and shipowner are becoming two distinct classes.”

(cited in Souden, 2003 pg 70)
The invention of the compound engine made it clear to the forward thinking owner that the days of sail were numbered. As time went by steam unequivocally asserted its superiority.

Many household names in the shipping industry were formed during these times, (e.g. Peninsular & Oriental Steam Navigation Company, Royal Mail Company, and Cunard). These companies were incorporated to be large important companies by their founders. They enjoyed government support and all received a substantial annual subsidy for carrying mail to the empire. Another factor was the Limited Liability Act of 1862. This was a shock to the traditional business people but now meant that capital available for the development of a business was restricted to the amount belonging to one man or a group of men. Shipping companies grew to a point where it was impossible for one person to finance so took advantage of the new method of trading, the Limited Liability Companies. This enabled them to obtain the amount of capital which they could prove to the investing public, by their balance sheets and dividends they were able to operate to advantage (Kirkaldy 1914, pg. 164).

After 1850 which became the age of the great steamships and railways, London’s financial influence spread across the Atlantic and all Europe. Banking, company formation and shipping all went hand in hand. Over the previous century bankers had become shipowners and vice versa; some bankers had underwritten capital and infrastructure improvements and many numbered shipping interest among their mercantile customers. All those factors continued, but now lending on a major scale to shipping interest became a fact of life in the banking sector (Souden 2003, pg.104).

In 1849 there were 99 joint stock banking companies in England and Wales. In 1892 this had grown to 102, 1898 diminished to 53 and in 1913 reduced to 43 (Kirkaldy 1914, pg.161).

At the turn of the century, (19th/20th), the involvement of finance houses came to the fore. American and German banks combined in a series of ventures and an American bank bought the White Star company. This company together with a German shipping line formed the International Mercantile Marine in 1902 which was to challenge the British company Cunard’s supremacy on the North Atlantic. In 1903 the British government gave Cunard financial assistance to build two passenger ships with steam turbine engines. A twenty year mail contract, an additional annual subsidy of £115,000 and a loan of £2.6million at 2.75% to ensure that the company remained British, build liners to be registered in the UK and available to the Royal Navy in an emergency (Stewart 1993, pg.42).

Further Research

Further research will now be undertaken, by the author, to discover the role of specialist banks in the shipping industry and to verify that debt as well as equity is a source of finance. The period from 1833, (the end of the East India trade monopoly) onwards to the Second World War will be investigated as it is known that it was after the war that charters started to influence the tenor and types of loans available.
References

BPP, British Parliamentary Papers, 1844, VIII, *House of Commons Select Committee on Shipping*.
Chatterton EK (1931) *Sailing the Seas*, Chapman & Hall, London. pg 1,4
Hakluyt Voyages, ii, pg 455 cited in Kirkaldy (1914)
Kirkaldy AW. (1914) *British Shipping, its History, Organisation and Importance*.
David and Charles, Newton Abbot.

http://www.bbc.co.uk/history/timelines/britain/tud_eastindia_co.shtml [doa 12/01/06]
http://www.lr.org/services_overview/shipping_information/is007heic.htm [doa 12/01/06]
KNOWLEDGE TRANSFER BETWEEN EUROPE AND ASIA: IS TURKEY A BRIDGE BETWEEN NETWORKS?

Dababrata N. Chowdhury

Abstract

Turkey’s geographic and cultural position represents a bridge between Europe and Asia. This research investigates the way in which knowledge and resource transfer between these two regions may occur through Turkish business. I am exploring the impact of Turkey’s position on businesses in Turkey, studying exploring the impact Knowledge Transfer between Turkish and foreign firms. and I consider the role Turkey currently plays and will in future play in transcultural Knowledge Transfer between international businesses.

Keywords: Knowledge Transfer, Turkey, Competitive Advantage, Network Bridge.

Introduction

Knowledge transfer is one of the very important key areas in knowledge management within any organisation. Gram explains that Knowledge Transfer is an interactive process in which advanced cutting edged knowledge is passed on to the local community by experts and technologists (Gram, 2003). The word “transfer” is used instead of “diffusion” to emphasize the movement of the knowledge (Gram, 2002). Knowledge transfer had also received increasing interest among the new modern business world and new managers. Increasingly popular knowledge management efforts attempt to establish better utilization of knowledge assets and help new knowledge generation. Knowledge transfer is acknowledged as a major challenge for the multi-unit firms in the current economy. ‘A learning organization is an organization skilled at creating, acquiring, and transferring knowledge, and at modifying its behaviour to reflect new knowledge and insights’ (Garvin, 1993). Madeleine Butschier argues that knowledge represents an additional resource for communication and effective processes. She also explain that a majority of the initial contacts between managers can be time and labour intensive. She argues that people need to develop an understanding of other members’ business and communication requirements and in the beginning stage of creating a knowledge transfer network, small victories are important foundations on which to build (Butschier 2002). In my research, I will explain the concept of Knowledge Transfer between Asia and Europe via Turkey.

This paper addresses the following questions: Which are the most important economic aspects of Turkey for Europe? How important are these for Turkey and the Europe respectively? What knowledge transfers from Asia to Europe via Turkey and vice versa? Finally I close the paper with a discussion of “Turkey is a Bridge” for Asia and Europe.
What is Europe?

Europe is geologically and geographically a peninsula or subcontinent, forming the westernmost part of Eurasia. It is conventionally considered a continent, which, in this case, is more of a cultural distinction than a geographic one. It is bounded to the north by the Arctic Ocean, to the west by the Atlantic Ocean, to the south by the Mediterranean Sea, and to the east its boundary is culturally determined and unclear. However, the Ural mountains are considered by most authorities to be a geographical and tectonic landmark separating Europe and Asia.

Figure 1 World map showing Europe geographically (source: en.wikipedia.org/wiki/Europe)

Europe is the world's second-smallest continent in terms of area, with an area of 10,600,000 km² (4,140,625 square miles), making it larger than Australia only. In terms of population, it is the third-largest continent (Asia and Africa are larger). The population of Europe is roughly 700,000,000: about 11% of the world’s population. Europe has a long history of cultural and economic achievement, starting as far back as the Palaeolithic, although this is true for the rest of the Old World as well. (Source http://www.lib.utexas.edu/maps/europe.html)

Geography

Geographically Europe is a part of the larger landmass known as Eurasia. At times “Europe” is defined with greater regard to political, economic, and other cultural considerations. This has led to there being several different Europes that are not always identical in size, including or excluding countries according to the definition of Europe used. Almost all European countries are members of the Council of Europe, the exceptions being Belarus, and the Holy See (Vatican City). The idea of the European continent is not held across all cultures. Some non-European geographical texts refer to the continent of Eurasia, or to the European peninsula, given that Europe is not surrounded by sea. In the past concepts such as Christendom were deemed more important.

In another usage, Europe is increasingly being used as a short-form for the European Union (EU) and its members, currently consisting of 25 member states. A number of other European countries are negotiating for membership, and several more are expected to begin negotiations in the future (source :http://en.wikipedia.org/wiki/Europe#Etymology).
What Is Asia?

Asia is the central and eastern part of Eurasia and world's largest continent. Defined by subtracting Europe from Eurasia, Asia is either regarded as a landmass of its own or as part of Eurasia. (source: en.wikipedia.org/wiki/Asia)

The demarcation between Asia and Africa is the Isthmus of Suez (though the Sinai Peninsula, being a part of Egypt east of the canal, is often geopolitically considered a part of Africa).

The boundary between Asia and Europe runs via the Dardanelles, the Sea of Marmara, the Bosphorus, to the Black Sea, the Caucasus Mountains, the Caspian Sea, the Ural River to its source, and the Ural Mountains to the Kara Sea at Kara, Russia. About 60 percent of the world's human population lives in Asia. Asia as a political division consists of the part of Eurasia and nearby islands in the Indian and Pacific Oceans, often excluding Russia.

Figure 2 World map showing Asia geographically (source:http://en.wikipedia.org/wiki/Asia)

As already mentioned, Asia is a sub region of Eurasia. Asia itself is often divided in the following sub regions:
• North Asia
• Central Asia
• East Asia (or Far East)
• Southeast Asia
• South Asia (or Indian Subcontinent)
• Southwest Asia (or West Asia)
Turkey & EU

In December 1999, Turkey became the European Union’s (EU) first candidate for full membership. For the first time, Turkey participated as a full member candidate at the EU summit in Nice in December 2000.

Straddling Europe and Asia, Turks have always attached great importance to relations with Europe. Turkey began “westernizing” its economic, political and social structures in the 19th century. Following the First World War and the establishment of the Republic in 1923, Western Europe was the model for its new secular democracy.

As a founding member of the United Nations as well as the OECD, a member of NATO and the Council of Europe, and an associate member of the Western European Union, Turkey’s vital role in the defence of Europe are undeniable.

Turkey applied for full membership of the European Economic Community (EEC) in 1959, less than a year after the organization was established. This application resulted in the 1963 Association Agreement envisaging Turkey’s gradual integration and eventual full membership into the EEC. Turkey then became the longest standing associate member of the EU, the EEC’s successor.

In 1987, Turkey applied for full membership to the EU on a direct track, different from the process outlined in the Association Agreement. The EU did not give a definitive response to this application, citing internal reforms, which it was undertaking at the time.

Meanwhile, the Turkey-EU relations developed and a Customs Union, which had been foreseen as the final step before Turkey’s full membership, was completed in 1995.

The United States strongly endorsed EU membership for Turkey, and lobbied intensely in this effort. Already a member of the EU, Greece dropped its objection, providing the required unanimous approval. This situation, to which Turkey protested strongly, has been remedied with the EU’s decision last December to make Turkey a candidate for membership.

A stable, flourishing Turkey benefits Europe and the U.S. as well as Turkey (Asia). Turkey’s EU candidacy will pave the way for a series of additional changes, economic and political as Turkey strives to align its economy and laws with those of the EU. (Source: [http://www.turkishembassy.org/governmentpolitics/issueseu.htm](http://www.turkishembassy.org/governmentpolitics/issueseu.htm))
European Union member states

Figure 3 EU Member States

(Source http://en.wikipedia.org/wiki/European_Union)

On 3 October 2005, membership negotiations were symbolically opened with Turkey, which has been an associate member of the EU since 1963 and an official candidate since 1999. The historic decision on 17 December 2004 by the European Council was confirmed by the European heads of state and government on 17 June. On 29 June, the Commission presented its negotiating framework to Ankara, and after a full day of intense negotiations the EU-25's foreign ministers finalised the document on 3 October 2005. Within hours, Turkey accepted the terms.

Country overview: Turkey

Geography and administrative aspects

The lands of Turkey are located at a point where the three continents making up the old world, Asia, Africa and Europe are closest to each other, and straddle the point where Europe and Asia meet. It has also been as a centre of commerce because of its land connections to three continents and sea surrounding it on three sides. Geographically, the country is located in the northern half of the hemisphere at a point that is about halfway between the equator and the north pole, at a longitude of 36 degrees N to 42 degrees N and a latitude of 26 degrees E to 45 degrees E. Turkey is roughly rectangular in shape and is 1,660 kilometres long and 550 kilometres wide.

Turkey has a total area of 779,452 square kilometres, of which 14,300 square kilometres is water surfaces. Turkey has influential geo-political status because its location serves as a natural bridge between Europe and Asia. It is surrounded by the Black see in the north, the Mediterranean Sea in the south and the Aegean Sea in the west. The actual surface area of the Turkey inclusive of its lakes is 814,578 square kilometres, of which 790,200 are in Asia and 24,378 are located in Europe.

It shares land boundaries with Greece and Bulgaria in the North West, Georgia, Armenia and Azerbaijan in he northeast, Iran in the east and Iraq and Syria in the southeast. The land borders of Turkey are 2,753 kilometres in total, and coastlines (including islands) are another 8,333 kilometres. Turkey has two European and six Asian countries for neighbours along its land borders. The land border to the northeast with the commonwealth of Independent States is 610 kilometres long; that with Iran, 454 kilometres long, and that with Iraq 331 kilometres long. In the south is the 877 kilometres long border with Syria. Turkey's borders on the European continent consist of a 212 kilometres frontier with Greece and a 269 kilometres border with Bulgaria. Turkey is generally divided into seven regions: the Aegean, the Mediterranean, Central Anatolia, the East and Southeast Anatolia regions. Turkey has 81 provinces and 76,457 villages. Capital of Turkey is Ankara city.
Socio-economic features

Population: On a general basis, the population in Turkey is characterized by youth and dynamism. Based on better health care conditions available after the 1950’s, the average life span in the country has taken an upward trend. Statistics in Turkey indicate that elderly are mostly concentrated in the rural areas.

Culture:

Turkey has a very diverse culture derived from various elements of the Ottoman Empire, European and the Islamic traditions. As Turkey successfully transformed from a religion-driven former Ottoman Empire into a modern nation-State with a very strong separation of state and religion, the increase in the methods of artistic expression followed. During the first years of the republic, the government invested a large amount of resources into the fine arts, such as paintings sculpture architecture amongst other things. This was done as both a process of modernisation and creating a cultural identity.

The Economy: Since the Second World War, the Turkish economy has been transformed by the steady growth of industry and services, and has experienced the consequent decline in the share of agriculture in national income. The economic growth rate, albeit erratic, has been one of the highest in the OECD. Turkey followed classic path import-substitution industrialization up to 1980s. Since then it has embarked on a structural economic reform programme, aimed at liberalizing the previously regulated domestic market, and at re-orienting the Turkish economy by adopting outward-looking trade policies.
Turkey's dynamic economy is a complex mix of modern industry and commerce along with a traditional agriculture sector that in 2004 still accounted for more than 35% of employment. It has a strong and rapidly growing private sector, yet the state still plays a major role in basic industry, banking, transport, and communication. The largest industrial sector is textiles and clothing, which accounts for one-third of industrial employment; it faces stiff competition in international markets with the end of the global quota system. However, other sectors, notably the automotive and electronics industries are rising in importance within Turkey's export mix. In recent years the economic situation has been marked by erratic economic growth and serious imbalances. Real GNP growth has exceeded 6% in many years, but this strong expansion has been interrupted by sharp declines in output in 1994, 1999, and 2001. Inflation, in recent years in the high double-digit range, fell to 9.3% by 2004 - a 30-year low. Despite these strong economic gains in 2002-04, which were largely due to renewed investor interest in emerging markets, IMF backing, and tighter fiscal policy, the economy is still plagued with high debt and deficits. The public sector fiscal deficit exceeds 6% of GDP - due in large part to the huge burden of interest payments, which accounted for more than 40% of central government spending in 2004, and to populist spending. Foreign direct investment (FDI) in Turkey remains low - averaging less than $1 billion annually, but further economic and judicial reforms and prospective EU membership are expected to boost FDI. A major political and economic issue over the next decade is whether or not Turkey will become a member of the EU.

Source: http://www.fao.org/ag/agl/swlwpnr/reports/y_nr/z_tr/tr.htm#submenu1

Table 1 Economic Outlook of Turkey

| BASIC INDICATORS | GNP | Worth $ 301.5 billion (2004)  
(sixteenth biggest economy among 30 OECD countries with respect to GNP) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth rate</td>
<td></td>
<td>9.9% (2004, based on real GNP)</td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td>72 million (2005-estimate)</td>
</tr>
<tr>
<td>GNP per capita</td>
<td></td>
<td>$ 4,172 (2004)</td>
</tr>
<tr>
<td>Inflation rate (CPI)</td>
<td></td>
<td>9.3% (Dec 03-Dec 04)</td>
</tr>
</tbody>
</table>
| Central Bank     |     | Independent (since May 2001)                                              
(Main objective is price stability. ‘No lending to the Government’ is an important policy tool) |
| Foreign exchange rates |   | YTL 1.3584 = $ 1  
(November 9, 2005)                                                             |
|                  |     | YTL 1.5940 = € 1                                                          |
|                  |     | YTL 2.3568 = £ 1                                                          |
| Foreign exchange regime |   | Floating Exchange Rate Regime                                             |
(Exchange rates are determined by demand and supply conditions. The Central Bank may hold auctions and intervene volatily)

<table>
<thead>
<tr>
<th>Inflation targeting</th>
<th>Implicit inflation targeting by end-2005</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Main policy tool is short term interest rates, which are based only on inflation outlook)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Imports</td>
<td>$97.2 billion (2004)</td>
</tr>
<tr>
<td>Exports</td>
<td>$62.8 billion (2004)</td>
</tr>
<tr>
<td>Foreign trade balance</td>
<td>-$34.4 billion (2004)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(approximately -$21.3 billion)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consolidated budget primary surplus</th>
<th>+TL 26.2 quadrillion (2004)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(approximately +$18.4 billion)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Primary Surplus (IMF Definition)</th>
<th>6.9% of GNP (2004)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Public Sector’s Outstanding External Debt</th>
<th>$73.8 billion (end-2004)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(All public debts are medium or long-term)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Consolidated Budget External Debt Stock</th>
<th>$68.5 billion (end-2004)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Outstanding Domestic Debt</th>
<th>YTL 231.3 billion (end-February 2005)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Bonds form 87% of the domestic debt stock)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>International reserves*</th>
<th>$52.4 billion (March 2005)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Central Bank Interest Rates</th>
<th>Borrowing</th>
<th>Lending</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overnight</td>
<td>13.75%</td>
<td>17.75%</td>
</tr>
<tr>
<td>Late Liquidity</td>
<td>5.00%</td>
<td>22.00%</td>
</tr>
<tr>
<td>1 Week</td>
<td>13.75%</td>
<td>-</td>
</tr>
<tr>
<td>(As of November 09, 2005)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in stock-market (IMKB)</th>
<th>+47.8% in $ terms (Jan 1, 2004-Jan 12, 2005)**</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Ratings (FX-denominated Government Bonds)</th>
<th>Fitch Ratings: BB- (Stable) (Jan 13, 2005)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard &amp; Poor’s: BB- (Stable) (Aug 17, 2004)</td>
<td></td>
</tr>
<tr>
<td>JCR: BB- (Positive) (Mar 10, 2005)</td>
<td></td>
</tr>
<tr>
<td>Moody’s: B1 (Positive) (Feb 11, 2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Ratings</strong> (Lira-denominated Government Bonds)</td>
<td><strong>Fitch Ratings:</strong> B+ (Stable) <em>(Feb 02, 2004)</em></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Standard&amp;Poor’s: BB- (Positive) <em>(Mar 8, 2004)</em></td>
<td></td>
</tr>
<tr>
<td><strong>Moody’s:</strong> B1 (Positive) <em>(Feb 11, 2005)</em></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Population aged 15 and over</strong></th>
<th>50 million <em>(2004)</em> <em>(about 70% of total population)</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Civilian Labour Force</strong></td>
<td>24.3 million <em>(2004)</em></td>
</tr>
<tr>
<td><strong>Civilian Employment</strong></td>
<td>21.8 million <em>(2004)</em></td>
</tr>
<tr>
<td><strong>Civilian Unemployment rate</strong></td>
<td>10.5% <em>(2004)</em></td>
</tr>
</tbody>
</table>

(*) International Reserves=Central Bank’s “Gross Reserves + Gold - Overdrafts” + Banks’ “Net Reserves”
(**) The seventh best among 25 emerging-markets’ stock-markets in 2004 (Source: The Economist, Jan 15-21, 2005.)
(***) $: US Dollars (Source: http://www.turkisheconomy.org.uk/economy.html)

**Why Turkey?**

Turkey’s strategic location makes it a natural “bridge” between major business producing areas in the Europe and Asia. Turkey has experienced a strong recovery from a severe economic contraction it experienced in 2001 due to a devastating financial and currency crisis. During 2004, Turkey’s real gross domestic product (GDP) grew by a rapid 8.9 percent, with an inflation rate of 8.6 percent. For 2005, real GDP growth is forecast at 5.6 percent, with inflation of 8.0 percent. Unemployment was running around 11.7 percent in the first quarter of 2005, down from 12.4 percent in the first quarter of 2004. However, (Ingmar Karlsson, 2004) argues that Turkish culture and religious heritage - an asset to the European union. He responds to the security and geopolitical arguments that advocate Turkey’s exclusion on the grounds that Turkey will import the instability of its region to the union, by stressing the point that an excluded turkey cannot act as a firewall against the crisis in the middle east as all the crisis in the middle east so far have directly affected Europe and will affect so in the future. He underlines having Turkey as a full member will have significant advantage for the union as it will increase the EU’s opportunities for pursuing proactive policies in the Arab world and, with its stable democracy, stand as a model for a Muslim world that is in need of such models.

(Source:http://www.eia.doe.gov/emeu/cabs/turkey.html).

Steven Everts (2003) from The Centre of European Reform explains that “An asset but not a model: Turkey, the EU and the wider Middle East”, which looks like two set of questions. The first being the consequences of Turkey’s accession for EU policies in the wider Middle East and the ‘Turkey as a bridge’ or ‘model’ argument. With respect to the knowledge transfer to the first issue, he concludes that Turkey has a lot to contribute to EU policies on the Middle East in terms of credibility, political access, know-how and economic leverage. Regarding the ‘model’ or the ‘bridge’ argument, Everts view is that Turkey cannot be a model for democratising the wider Middle East as it is a unique case in three key respects: its long standing
ties with the West, its secular structure and the transformation of its political elite through the prospect of EU membership. According to Everts, the fact that Turkey has ambivalent relations with most of its neighbours in the region also aggravates the need to tone down the ‘Turkey as a bridge’ argument and avoid the ‘Turkey as a model’.

Everts (2003), States that even EU officials admit that the EU is under performing in the Middle East. Institutional incoherence, poor political discipline, risk aversion and insufficient emphasis on promoting good governance and democracy have all taken their toll. EU lacks credibility and clout. Turkey, as a Muslim country straddling Europe in the Middle East, could be of help here, making EU business policies perhaps more acceptable to countries in the region, especially in the pre-accession phase before it becomes a EU member.

Everts (2003) explains the importance of Turkey. Unlike any other country in the region, Turkey has a long standing relation with the West: institutionally through NATO, and bilaterally with the United States and capitals in Europe. No other country in the region has the same or comparable influence. Nor was any other country in the region born out of an empire, which gives Turks greater political self confidence. Secondly, Everts explains that the Turkey’s strict secular nature of its political system has put Turkey in a distinct international category. Thirdly, deepening Turkey’s democratisation took place largely because of the ‘golden carrot’ of EU membership.

Micheal Emerson and Nathalie Tocci (2004) explain that Turkey’s links to Central Asia would take a bilateral rather than a multi-lateral form. Despite these setbacks Turkey became deeply involved in Central Asia in the early and mid 1990s. It immediately provided emergency assistance (food and medicine) to the republics. It also engaged in medium- and long term projects and concluded a number of bilateral agreements with the Central Asia republics, covering economics and trade, business corporation, public administration, media and education, communication and transport. They also explained Turkish companies are strongly encouraged by the state to invest in Central Asia. In 1992, around 60 Turkish banks and businesses invested in Central Asia and approximately 150 established links with the republics. And Turkey also offered assistance in establishing communication and media in the republics. Turkey established several education institutions in the republics including two major universities in Kazakhstan and Kyrgyzstan. Turkey’s role is very important in gathering and transferring the knowledge between European countries and the rest of the world.

In the view of knowledge gathering and transferring in the business world, Ingar Karlsson argued that Turkey now faces three geostrategic choices: strong decision of its European identity, rapprochement with the Arab and Muslim world, and integration with the Turkish speaking peoples of Central Asia. Therefore knowledge comes to Turkey’s businesses from different background which is very important to transfer it between European countries via Turkey. Because the resources of the knowledge transfer from Asia to Europe is Turkey itself which plays a vital role to revolutionise the entire economy of the European and Asian countries.

Knowledge Transfer

A synopsis of the definitions of knowledge taken from Ören (1990) - which also provides a taxonomy of about 500 types of knowledge and knowledge processing
knowledge—follows: “A limited‐scope definition of knowledge is ‘facts, beliefs, and heuristic rules’ (Hayes‐Roth et al. 1983, p. 401). Russell defines knowledge as ‘a sub‐class of true beliefs’ (Russell, 1948, p. 170). Ayer states that ‘the necessary and sufficient conditions for knowing that something is the case are first that what one is said to know be true, secondly that one be sure of it, and thirdly that one should have the right to be sure’ (Ayer, 1956, p. 35). Minsky summarizes these views by saying: “Some philosophers have argued that ‘knowing’ must mean ‘true’ and ‘justified’ belief” (Minsky, 1988, p. 302)

Some scholars contend that to be considered useful to an organization, the definition of knowledge must include at least three concepts. First, it must point out that knowledge is more than data or information. Second, it must describe the tacit and explicit nature of knowledge. Finally, it must describe the personal nature of knowledge (Nonaka & Takeuchi, 1995). Davenport and Prusak (2000) offer the following definition:

*Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knower* (p. 5).

This definition also addresses key cultural components of organizations. These factors include the varied experiences and values of the organization’s membership and a framework for evaluating and incorporating new experiences and information through embedded routines, processes, practices, and norms.

According to Davenport and Prusak (1998) knowledge may be generated through acquisition. In this process originality is less important then usefulness. The most direct and often most effective way to acquire knowledge is to buy it or to hire it individuals who have it. Nonaka and Takeuchi (1995) and Wiig (1993) define two types of knowledge: Explicit knowledge and tacit knowledge. Explicit knowledge can be articulated in natural and formal language, which makes it “easy” to transmit between people via documents and other types of records. Tacit knowledge has to do with personal knowledge that is embedded in personal experience and is therefore not so easy to formalise and record.

Knowledge Transfer from one set of individuals to another has been a key area of interest for knowledge management researchers (Sarker, 2005). Although knowledge transfer is an important component of knowledge management (Davenport & Prusak, 2000). Alavi (2000) highlights the importance of knowledge transfer by suggesting that for superior performance of a social entity, knowledge generation and its successful transfer needs to take place. Cross, Parker, Prusak, and Borgatti (2004, p. 62) also posit the value of knowledge sharing in today’s economy, “where collaboration and innovation are increasingly central to organizational effectiveness in Asia and Europe.” To successfully build a large and complex networks system in Asia and Europe, team members need to learn continuously from each other regarding different issues, including the capabilities of the new networks system, language, culture and laws. This transfer of knowledge is “often laborious, time consuming, and difficult” (Szulanski, 2000, p. 10), and it can become even more daunting in situations where knowledge is being transferred across time and space, such as in a globally distributed team (Alshawi & Al-Karaghoulı, 2003; Davenport & Prusak, 1998; Sarker & Sahay, 200). Thus it is important to understand factors that impede or facilitate such transfer of knowledge.
The study of knowledge transfer predates the study of knowledge management by several decades (Argote, Ingram, Levine, & Moreland, 2000). Indeed, the notion that knowledge transfer could represent not only a competitive advantage within a firm but also a less expensive alternative to knowledge creation and acquisition is well documented in economics (Alchian & Demsetz, 1972) and organizational behaviour literature (Argote & Ingram, 2000).

Interestingly, as organizations consume material assets, they often decrease in value and quantity. On the other hand, when organizations use knowledge resources, these assets tend to increase in that both the giver and receiver are enriched as a result of the transaction (Davenport & Prusak, 2000). For example, more than one individual can use knowledge at the same time, and shared knowledge stimulates the creation of new knowledge. More importantly, this process appears to reduce costs and significantly contribute to overall organizational success by preventing individuals from repeating the mistakes of other individuals (Baum & Ingram, 1998; Gruenfeld, Martorana, & Fan, 2000).

Knowledge transfer is nominally concerned with the process of moving useful information from one individual to another person. In order for this transferred information to have utility, it must be critical to the success of the organization (Davenport & Prusak, 2000). Extant literature provides several instances of organizations skilful at knowledge transfer (Zairi & Whymark, 2000), but most of these case studies do not fully explore why these organizations were successful at this endeavour. To fully understand how to grow this capability, it is probably necessary to understand what factors tend to affect knowledge transfer. A recent literature review within the knowledge domain provided the following five factors that might influence knowledge transfer (A. Ladd & Mark A. Ward, 2002).

1. Relational channels … frequency and depth of two-way human-to-human contact (Rulke, Zaheer, & Anderson, 2000)

2. Partner similarity … degree of similarity (e.g., interests, background, or education) between individuals (Almeida & Kogut, 1999; Darr & Kurtzberg, 2000)

3. Depreciation … loss of knowledge after transfer (Argote, Beckman, & Epple, 1990; Darr, Argote, & Epple, 1995)


( Source http://www.tlainc.com/articl38.htm)

A critical factor enabling knowledge transfer in organizations is the presence of “smart people” (Davenport & Prusak, 1998, p. 88) who have the expertise necessary to accomplish the work. A source with a greater expertise than his/her remote members has the potential to transfer more knowledge to the recipients (Hinds et al., 2001; Zander & Kogut, 1994). Levin et al. (2004) argue that an individual who is perceived to be more knowledgeable (i.e., has more competence in a given subject area than a recipient) is likely to transfer knowledge to that recipient.
Conclusion

The use of information and communication technology can facilitate the process of knowledge transfer between the various communities. The Republic of Turkey is located mainly in Asia Minor, with 3% of its land located in the Balkans. However its strategic location straddles the Bosporus straits that separated South East Asia from South East Europe. Turkey, surrounded by three seas namely Mediterranean Sea, Black Sea and Marmora Sea, has been a historic crossroads and battleground between great civilizations and a centre of commerce. Turkey borders eight countries which are Greece and Bulgaria to the northwest, Georgia, Armenia and the Azerbaijani to the northeast, Iran to the east and Iraq and Syria to the south.

Both the economic and management tradition of international business attribute great importance to the role of resource in generating competitive advantage. As Fahy (2002) explains, the resource based view of the firm offers a systematic framework for analysing any complex and diverse resource pool using barriers to duplication as the key competitive dimension. Turkey is an economic, geographical and cultural bridge and as such acts as a conduit of ideas and resources from Asia to Europe, providing both sides access to the diverse resources necessary for innovation and competitiveness. In this bridging position Turkish should benefit increasingly from the knowledge transfer across their bridge. This research investigates the impact it is having on Turkish small businesses in two main industrial groups: Textiles and IT.

References

Interpretation”, Organization Studies, vol. 16, no. 6, 1995, pp.1021-1046;
Nonaka, (1994)”A Dynamic Theory of Organizational Knowledge Creation”, 
*Organization Science*, vol. 5, no. 1, pp.14‐37


[http://www.ceps.be](http://www.ceps.be)
[http://www.gutenberg.org/author/European+Union](http://www.gutenberg.org/author/European+Union)
[http://www.oecd.org/eu](http://www.oecd.org/eu)
[http://www.zei.de/zei_startseite_neu/startseite_e.htm](http://www.zei.de/zei_startseite_neu/startseite_e.htm)
[http://www.ena.lu/](http://www.ena.lu/)
[http://europeansociety.tripod.com/](http://europeansociety.tripod.com/)
[http://www.lib.utexas.edu/maps/europe.html](http://www.lib.utexas.edu/maps/europe.html)
[http://www.turkishembassy.org/governmentpolitics/issueseu.htm](http://www.turkishembassy.org/governmentpolitics/issueseu.htm)
THE STRATEGIC MOTIVES FOR INTERNATIONAL JOINT VENTURES AND NETWORKS: A STUDY OF SINO-TAIWANESE JOINT VENTURES

Chen-Yu, James, Liu
Lynne Butel
(Director of Studies)

Abstract

Cultural similarity has been identified as a significant factor in internationalisation of firms, both in the choice of new country and the success of new business. Previous studies of international joint venture formation and the establishment of wholly owned subsidiaries in People’s Republic of China (PRC) have generally focused on cultural factors from the Western perspective. This study examines the establishment and utilisation of social and business networks in Sino-Eastern international business formation. The importance of guanxi in all Chinese social and business relations, is widely acknowledged. There have been no studies exploring the significance of the cultural similarities within the context of a politically hostile environment such as exists between Taiwan and Mainland China. Taiwanese investors just like other countries and regions; seek a suitable investment China to take advantage of the low cost labour, cheap operation costs and available resources.

In addition, Taiwan and China have a very similar cultural identity, the Taiwanese speak the same language as much of mainland China and the use of interpersonal social and business networks has been used extensively in Chinese communities. The preference the Chinese government has for international joint ventures on joint ventures has encouraged the use of business networking to search for suitable partners and guanxi has become a significant part of the networking process. The use of guanxi appears to allow the Taiwanese businesses to become established in China despite the hostile environment.

As a consequence, four key strategic motives for international networking which are less easily identified in other international contexts. We analyse the significance of culture factors on guanxi particularly the extent to which they may influence joint venture performance. The research examines the importance of a guanxi network in providing rich and valuable business information which depends almost entirely on the strength of guanxi ties. Finally we consider the extent to which Guanxi provides safeguard for Taiwanese investors; especially, small and medium size businesses who are seeking entry and expansion in Chinese market. The
research draws valuable insight into the operation of social and business networks in hostile environments which are applicable in other business contexts.

Introduction

Increasingly, large portions of the world economy have already depended on China (Goodman and Segal, 1998; Gatiskell, 2000). Many significant economic incentive regulations after “Open Door Policy” in 1979, which have brought a great opportunity for China to become the world’s largest Foreign Direct Investment (FDI) recipient ahead of the United States (Organisation for Economic Co-operation and Development, 2003). It is believed that China had the most joint venture businesses than any other national in the decade of 80s (U.S. Department of Commerce 1991). According to China Statistic Year Book 2003, the total value of JV in foreign direct investment is growing fast than ever from almost nothing in 1978 to 275 billion U.S. dollars in 2003. At the early age of economic reform, Chinese government preferred and encouraged foreign enterprises to form a joint venture (JV) to access both Chinese domestic market and resources. This allowed the Chinese partners to acquire useful knowledge and management skills from their foreign venture partners (Luo, Shenkar and Nyaw, 2001) and very importantly, the Chinese government could monitor the foreign investors’ progresses.

Given the strong economic and cultural links between Taiwan and People Republic of China (PRC), problems of JV formation from Taiwanese firms can be identified from the phenomenon of China’s economic “hollowing-out” effect. Due to the deteriorating of Taiwan’s domestic investment environment (push factor) which in contrast with China’s (pulling factor) that Taiwanese businessmen became one of the earliest “foreign investors” who were able to establish their presence and quickly react to opportunity in China’s market. Sino-Taiwanese JV cases had increased dramatically since China opened up to the outside world. Lured by benefits, it is estimated that there are more than 1 million private Taiwanese businessmen living in China; especially 500,000 living in Shanghai along and total investments have reached approximately100 billion U.S. dollars\(^\text{13}\) (Business Week, 2004).

It has been argued that cultural similarities between Taiwan and China play an important role in JV formation and success. In addition, geographical proximity, has contributed to Sino-Taiwanese JVs emerging far earlier than other foreign investors, when China’s government adopted an open-up policy in early 1979. For example, one of the earliest Special Economic Zones (SEZs), Xiamen (near Taiwan), received mostly Taiwanese investments since 1989 (McKenney, 1993). In principle, the norms of JV investment generally relate to cases involving legally recognised foreign states, which cannot be applied on Republic of China (Taiwan) and China, due to political conflicts between two sides’ recognition of each other. Additionally, the importance of China “hollowing-out” effect is because that the Chinese government offered Taiwan entrepreneurs with a better investment arrangement than other foreign investors (Lin, 1999).

In this sense, the business links between Taiwan and China were forged not only through official channels but through guanxi networks or grey-market mechanisms (Luo, 2000). Luo (1997) emphasized that “guanxi” (interpersonal relationship) in Taiwan and China are...
critical to business transactions and negotiations process in different level of collaborations. So, in Sino-Taiwanese JV formation, the closeness of ethnic ties will have a significant effect in the partners chosen for the equity or contractual joint ventures in China (Luo, 2002). While Taiwan is one of the most important sources of FDI for PRC. A close look at the development of Taiwan base JV in China will improve our understanding of whether the cultural similarity provides with a stronger position that of Sino-foreign JVs.

Understanding Sino-Taiwanese Joint Venture

Taiwan:
Taiwan was originally known to the West as “Ilha Formosa” (Beautiful Island). It has now become one of the most powerful business nations in the world. Historically, Taiwan had been ruled and occupied by Portugal, Holland, Spain, Koxinga, the Qing Dynasty and Japan, and has struggled for liberty against foreign rulers. Taiwan’s earliest settlers might have come from the mainland, thus its history, culture and value systems are very similar to China. Because Taiwan was a colony for many decades, Taiwan’s culture is a blended of its distinctive Chinese heritage (Confucianism), Japanese’s characters (half a century of Japanese rule), and western influences (U.S. troops assistance for civil war).

When the Nationalist Party, KuoMingTung (KMT), retreated to Taiwan in 1949, Taiwan was governed by "Martial Law" for forty years. Hence opposition political parties were banned, publishing and the media were restricted, and relations with China were forbidden. Taiwan and China were forming their own separate political and economic systems including administrative, financial, taxing and currency. Gradually the two autonomous identities under the policy of “One China, Two Systems” conflicting circumstances.

In 1987, Taiwan successfully lifted the Martial Law and started responding to Chinese demand for cross-Strait trading by announcing “No-interference Principle of Indirect Exports to the Mainland.” Since then Taiwan’s trading activities to China were gradually increased, and Cross-Strait economic integration was seen as an unavoidable and irreversible process.

Characteristics of Taiwan Joint Venture in China:

When China decided to open up to the outside world in 1979, there was a strong needed for foreign investments, modern technologies and management skills for the Chinese government to improve its economic environment. Large number of contemporary studies has been conducted on international joint ventures formation in such emerging markets. Sino-foreign joint ventures are a most popular topic related to the Chinese market. (Pfeffer & Salancik, 1978; Killing, 1983; Harrigan, 1988 Kogut, 1988; Geringer & Hebert, 1989; Hamel, 1991; Yan & Gray, 1994; Caves, 1996; Beamish, Delios & Lecraw, 1997; Kostova & Zaheer, 1999; Dhanaraj & Beamish, 2001; Parke, 2001). However, a combination of an unstable political environment and the rapidly developing economic relation, the Sino-Taiwanese JV represent a more complex process. Thus, seeking an opportunity to enter China’s market to sustain their profitability would become the rational investment behaviour, since the short physical distance, the strong ethnic tie and the previous success of Taiwan’s economic experience may help them to create another Taiwan-style economic miracle in the mainland.
Based on the previous researches, for instance, from Taiwan Ministry of Economic Affairs (MOEA), 38 percent of Taiwanese joint ventures had been established in sample size of 1,627 companies. Approximately the 59 percent of Sino-Taiwanese joint ventures cases were holding less than 50 percent shares, and around 31 percent of these Sino-Taiwanese venture businesses cooperated with local Chinese firms, about 20 percent Taiwanese companies collaborated with China’s local governments, and 26 percent of Taiwanese companies joined with foreign and Chinese partners at the same time. Furthermore, the survey results from Chinese National Federation Industries in 2003 found that equity, contractual and multi-partner joint ventures were commonly put into practice when Taiwanese investors adopted joint venture strategy to enter China. Although wholly-owned subsidiaries has become the most favourite strategy for Taiwanese access to the Chinese market (Taiwan Industrial Association, 2003), many sectors in China, such as telecommunication industry, the exploitation of natural resources, and state owned businesses must establish joint venture with foreign partners to remain in control. Taiwan’s economy seems likely increasingly to depend on China’s resources to create the second economic miracle in order to sustain its economy. This research concerns two key aspects of Taiwanese and Chinese JV formation. First, the use of joint venture in China for Taiwan’s investors needs to be addressed, given the unusual relationship that exists between two sides. Secondly, whether Sino-Taiwanese joint venture can be explained by its geographical affinity, cultural similarity and requirements from Chinese government need to be considered.

Intricate Economic Relationship between Taiwan and China:

Taiwan’s investors are gradually seeing the mainland market as a significant contributor to continued development of Taiwan’s economy. Furthermore, Chinese government has cautiously encouraged Taiwanese investors, for example, establishing Taiwanese High-tech Industrial Park or assembling to export special area where Chinese authorities will reduce the time of paper works for Taiwanese investors. This has greatly promoted the substantial inflow of Taiwanese investment. The economic achievements of China are the result of two main factors (Morrison, 2003): large-scale capital investment and rapid productivity growth. First of all, mainland China took a more active role and began to open to the outside world, thus the reasons for martial law in Taiwan were no longer seen as valid; therefore, martial law was ended in 1987. Taiwanese investors have been responding this key historical breakthrough, like other foreign investors Taiwan investors were driven by the availability of cheap land and labour for low-cost expansion (Child, Chung & Davies, 2000).

Taiwanese investments in China grew dramatically. This drove the Taiwanese government to introduce several polices to cool down this phenomenon\textsuperscript{14}. In Figure 1 highlight the growth of Taiwanese investments in China which reflect the importance sources of Taiwan’s FDI.

\textsuperscript{14} The mid-1990s, Taiwan’s President Lee had announced the policy of “No Rush, Be Patient” and “Go South” caused Taiwan’s FDI slow down to inflow to China’s market.
Figure 1. Taiwan Investments in China from 1991-2002

<table>
<thead>
<tr>
<th>Year</th>
<th>Taiwan Investments data from Taiwan Official</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Investment Cases</td>
</tr>
<tr>
<td>1991</td>
<td>237</td>
</tr>
<tr>
<td>1992</td>
<td>264</td>
</tr>
<tr>
<td>1993</td>
<td>1262 (new applications)</td>
</tr>
<tr>
<td></td>
<td>8067 (addendum)</td>
</tr>
<tr>
<td>1994</td>
<td>934</td>
</tr>
<tr>
<td>1995</td>
<td>490</td>
</tr>
<tr>
<td>1996</td>
<td>383</td>
</tr>
<tr>
<td>1997</td>
<td>728 (new application)</td>
</tr>
<tr>
<td></td>
<td>7997</td>
</tr>
<tr>
<td>1998</td>
<td>641 (new application)</td>
</tr>
<tr>
<td></td>
<td>643 (addendum)</td>
</tr>
<tr>
<td>1999</td>
<td>488</td>
</tr>
<tr>
<td>2000</td>
<td>840</td>
</tr>
<tr>
<td>2001</td>
<td>1186</td>
</tr>
<tr>
<td>2002</td>
<td>1490 (new application)</td>
</tr>
<tr>
<td></td>
<td>3950 (addendum)</td>
</tr>
</tbody>
</table>

Resources: Taiwan Ministry of Economic Affairs and Ministry of Commerce of People Republic of China Department of Taiwan, Hong Kong, and Macao Affairs.

Additionally, data of the export and import activities of the two sides illustrate the growth of China’s productivity which can partly be attributed to Taiwanese economic involvement. Tung (2003) notes that Taiwanese enterprises in China usually involved the importing of intermediate and capital goods from Taiwan and the exporting of finished goods to developed countries, such as the United States and Japan. Such economic activities have not just coincided with the spirit of China’s economic reforms policy, but have led China’s manufacture industry to higher productivity. For example, Chinese exports produced by Taiwan invested firms in China were estimated to contribute between 14 percent and 18 percent of China’s total export. One-fifth of Chinese exports to the United States were produced by Taiwan’s firms in China (Tung, 2002). Figure 2 shows the sharp increase of Taiwan’s export and import volume from or to China in the early 1990s which shows that Taiwanese economy was increasing dependence on the Cross-Straits trade and becoming less dependent on the U.S. market. In terms of the total exports from Taiwan to the mainland China, it increased steadily from U.S$ 384.2 million in 1981 (Taiwan Statistic Bureau, 2004) to U.S$ 38,063.1 million in 2002, reflecting an average of annual growth rate of 5 percent. Meanwhile, the total import from Mainland was increasing from U.S$ 75.2 million in 1981 to

15 In 1958, the U.S. and Republic of China governments issued a joint communiqué reaffirming solidarity between the two countries. Through 1960s and 1970s, Taiwan’s economy was heavily dependent on U.S. military aid for recovering.
$6,585.9 million in 2002. Although the Taiwan’s government is enjoying a large trading surplus by restricted import items from China to protect domestic industries, Chinese government has benefited from the Taiwanese FDI strengthening its export ability resulted a trade surplus with other countries (Lin, 1999).

**Figure 2. Trade between Taiwan and China & Taiwan and U.S.**

<table>
<thead>
<tr>
<th>Year</th>
<th>Taiwan export to China</th>
<th>Taiwan import from China</th>
<th>Taiwan Import from US</th>
<th>Taiwan export to US</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>3,639.0</td>
<td>594.8</td>
<td>13,182.5</td>
<td>23,023.0</td>
</tr>
<tr>
<td>1992</td>
<td>5,881.0</td>
<td>698.0</td>
<td>15,250.4</td>
<td>24,596.0</td>
</tr>
<tr>
<td>1993</td>
<td>12,933.1</td>
<td>1,461.8</td>
<td>16,167.8</td>
<td>25,101.5</td>
</tr>
<tr>
<td>1994</td>
<td>14,084.8</td>
<td>2,242.2</td>
<td>17,108.9</td>
<td>26,705.7</td>
</tr>
<tr>
<td>1995</td>
<td>14,783.9</td>
<td>3,098.1</td>
<td>19,289.6</td>
<td>28,971.9</td>
</tr>
<tr>
<td>1996</td>
<td>16,182.2</td>
<td>2,802.7</td>
<td>18,460.3</td>
<td>29,907.3</td>
</tr>
<tr>
<td>1997</td>
<td>16,441.7</td>
<td>3,396.5</td>
<td>20,365.8</td>
<td>32,628.5</td>
</tr>
<tr>
<td>1998</td>
<td>16,629.6</td>
<td>3,869.6</td>
<td>18,164.4</td>
<td>33,124.9</td>
</tr>
<tr>
<td>1999</td>
<td>19,537.5</td>
<td>3,951.7</td>
<td>19,131.4</td>
<td>35,204.4</td>
</tr>
<tr>
<td>2000</td>
<td>25,497.1</td>
<td>4,994.9</td>
<td>24,405.9</td>
<td>40,502.6</td>
</tr>
<tr>
<td>2001</td>
<td>27,339.4</td>
<td>5,000.2</td>
<td>18,121.7</td>
<td>33,374.3</td>
</tr>
<tr>
<td>2002</td>
<td>38,063.1</td>
<td>6,585.9</td>
<td>18,381.7</td>
<td>32,147.9</td>
</tr>
</tbody>
</table>

Unit: U.S. Million Dollars

Resource: Straits Exchange Foundation Business Information, 2003., U.S. Census Bureau, Foreign Trade Division, Data Dissemination Branch, Washignton, D.C. 20233

*Sino-Taiwanese Joint Venture Strength and Weakness:

The Chinese government has provided Taiwanese investors with a better investment arrangement than other foreign investors (Lin, 1999). Taiwanese businesses in China not only could enjoy tax holidays, duty-free imports, land use rights, transfer and inheritance of properties, permission to purchase bonds, but also the special areas designed exclusively for Taiwanese investments (Cheng, 1995). China’s government tailor-made special investment
laws for Taiwanese investors included “Taiwan’s Investment Law” in 1988, “Law of Protection of Investments by Compatriots from Taiwan” in 1994, and “Rules for the Implementation of the Law of the People’s Republic of China on Protection of Investments by Taiwan Compatriots.” Taiwanese investors are benefited from a grey area of legal terms between China’s Foreign Investment Law and Taiwan’s Compatriot Investment Law and they are currently enjoying both of preferential treatments and legitimate autonomy, which was empowered by People Republic of China (PRC) government.

China’s economic development pattern seems to be one of following and learning process from the Taiwanese economic experience. Many foreign investors have been poured substantial funds into the labour intensive manufacturing industry and created large amount of jobs in China. It appears that Chinese government relies on the foreign partners’ skills and knowledge and did not make much effort on the area of research and development which Taiwanese partners have a dominant advantage particularly in IT industry. Prior to 90s the Chinese government encouraged foreign investors to focus on exporting activity, which resulted in only small proportionate goods allowed to be distributed in the domestic market. Thus, the main operational model for Taiwanese investors was “Receiving orders from Taiwan, manufacturing goods in China, and exporting from Hong Kong” to keep making a profit. Recently, in order to attract more foreign investments, Chinese government opened up more restricted businesses areas to foreign investors, such as retailing, telecommunication, financial institutions, insurance, and access to national protected resources industries via a Sino-foreign joint venture formation.

In sharp contrast, the government of Taiwan is more concerning its international status and position than creating an attractive investment environment between two sides. The Taiwanese government will not ease the restrictions on transportation and business between Taiwan and the Mainland. This has a substantial impact on some Taiwanese businesses. Control over the Sino-Taiwanese joint venture seemed to become a priority at start up. Many Chinese partners have viewed joint venture as equity joined relationship not collaboration. As a result, Taiwanese partner firms are keen on having more shares in the joint venture business, because they have more experience and capabilities of delivering quick and tangible results. This is not always possible for the Taiwanese partner to have dominant shares because the Chinese partners expect quick returns regarding their own contributions.

---

16 Minister of Foreign Trade and Economy Commission (MOFTEC) announced that Taiwan Compatriot Investment Law in 1988 to encourage Taiwan investors set up businesses in China to maintain national economy development.
17 At the Sixth Meeting of the Standing Committee of the Eighth National People’s Congress on March 5, 1994, the President Ching-Chang, Tse-Ming announced the Law of the People’s Republic of China on Protection of Investment by Compatriots from Taiwan.
18 On December 5th, 1999, Premier of the State Council, Zhu Rongji also announced “Rules for the Implementation of the Law of the People’s Republic of China on Protection of Investments by Taiwan Compatriots” to protect Taiwan investors’ businesses and promote Chinese economic development on both sides of the Straits.
19 According to the study of Taiwan’s investment experiences in China (Lin and Liang, 2003), it points out that 45.3% of total 481 sample company adopted wholly-owned subsidiaries while investing in China. Among the others, 31.2% and 23.5% of firms utilized majority and minority equity investment. Major ownership was adopted by steel (75%), telecommunication and cable (66.7%) and security (66.7) industry. As for minority ownership was adopted by transportation (62.5%) and auto (57%) industry.
**Literature Review**

**Defined Joint Venture:**

Joint Venture (JV) is often defined as two parties or more than two parties that have joined together to enter new business fields and it has been used at increasing rate as a means of entry strategy to new markets or new countries. More specifically, JV can be defined as a legally and economically distinct organisational entity, created by two or more parent firms that collectively invest capital and other resources to pursue certain strategic objectives (Preffer and Nowak, 1976). Firms choose JV entry modes to reduce costs and uncertainties when operating in foreign market (Griffith, 2001). Although it contributes many benefits when entering a foreign market, it is considered as the highest risk and most problematic of all entry modes (Franko, 1971; Levine & Byrne, 1986; Bleeke & Ernst, 1995).

![Figure 1. The joint venture basic model](image)

Understandably, JV not only provides particular advantages when firms are facing an uncertain environment (e.g., economic crisis, political turmoil, technological change, and competitive response), but it has long been a favoured mode for entering foreign markets (Beamish and Banks, 1987). Harrigan (1988) points out that JVs are more likely to succeed when partners possess complementary missions, resources capabilities, managerial capabilities, and other attributes that create strategic fit in which the bargaining power of the venture’s sponsors is evenly matched. The effectiveness of JV can be clearly identified as foreign firms contribute the technical expertise, equipment and some of the financing, whereas the Chinese partners provide the facilities, the labour force, contacts with government agencies, and market access.

**Transaction Cost Theory Approaches Joint Venture:**

The notable contributor for transaction cost theory is Williamson (1975) who assumes that for expository convenience, at the beginning, there were markets. The existence of the firm
is explained by market failures. The firm is chosen as a governance structure when the costs of carrying out certain exchange transactions in the open market are greater than organising these transactions within the firm. The basic idea of transaction cost has been widely used to suggest that firms transact in a way that minimizes the sum of production and transaction costs (Williamson, 1975; Kogut, 1988; Hennart, 1991). For example, while firms established venture business in a host country, partners can gain access to complementary downstream resources such as local channels of distribution (Killing, 1983), local market know-how (Blodgett, 1991; Inkpen & Beamish, 1996), and understanding of the local regulatory context (Kobrin, 1982) in order to minimize the transaction costs.

JV can be well addressed by transaction cost implication, because it provides a superior alignment of incentives via the mutual dedication of resources and sharing the residual value of the venture, which can neither be replicated nor acquired through market transactions (Yiu & Makino, 2002). Hennart (1988) expressed similar thought, he notes that a necessary condition for JV to emerge is the presence of inefficiencies in markets for intermediate inputs that include raw materials, components, and knowledge, are complicated in their pricing and transfer. Joint venture is an efficient means of internalizing these failing intermediate markets. Several studies have pointed out that using joint ventures by multinational enterprises tend to have higher performance compared to those wholly-owned subsidiaries (Woodcock et al, 1994; Nitsch et al, 1996; Pan, Li & Tse; 1999) and bring our attention to understand the significance of transaction costs within the firms.

Proposition 1: Using Sino-Taiwanese joint ventures can minimize the transaction and related costs so Taiwanese firms are able to achieve economic of scale, while they are integrating with Chinese partners’ market knowledge. Buckley and Casson (1988) have pointed out from the transaction cost perspective that joint venture would be the chosen over other modes of entry when the conditions are satisfied.

Selection of Joint Venture Entry Mode:

A significant process of this research has emerged on the choice of foreign market entry modes (e.g., Gatignon & Adherson; Agarwal & Ramaswami, 1992). The international business strategy selection usually begins with the consideration of internal and external environmental analysis and the determination of objectives (Anderson & Gatignon, 1988; Beamish & Wang, 1989; Gomes-Casseres, 1989,1990; Child, Boisot, Ireland, Li & Watts, 1990; Pearson, 1991; Yan & Gray, 1994; Buckley, 1996; Pan, 1996).

Joint venture is the oldest form of foreign business operation in China (Beamish, 1993). The characteristics of JV included risk sharing, resource pooling, asset protecting, and the ability to react quickly to market changes (Pan & Tse, 1996), which offer the insight view of how Taiwan’s firms adapted JV strategy to enter China’s huge market. Sino-foreign JVs are popular and government rules requiring them (Davidson, 1987; Eiteman, 1990). By complying Dunning’s framework (1980, 1988), several strands of international business theories on cross-border business activities are generally influenced by three types of factors. The first factor is that scale of operation, length of diplomatic ties, and country of origin fall into the area of ownership advantages. Moreover, China’s experience in attracting foreign investment, location and level of government come from the location specific advantages. The third factor is that cultural dimensions represent the internalisation
Proposition 2: Facing of Chinese military external threat and economic internal hollowing-out risk, Taiwanese investors can register in third counties to form a Sino-Taiwanese joint venture in which business its identity will be protected by the Sino-Foreign bilateral agreement when they are accessing the Chinese market.

Guanxi in Sino-Foreign Joint Venture:

Guanxi is an important concept in China (Wu, 1967) and is generally translated into special relationship and connections (Davies et al., 1995). Guanxi literally means personal relationship. But, the concept of guanxi is embedded within the Confucian philosophy (Fock & Woo, 1998), because the Confucian value system stresses hard work, respect for authority, harmonious relationships and obligation to the family (Simth, 1974). Yang (1994) notes that guanxi involves cultivating personal relationships through the exchange of favors and gifts for the purpose of obtaining goods and services, developing networks for mutual dependence, and creating a sense of obligation and indebtedness. Tang (2003) argued that guanxi is one of the most important factors contributing to JV success in China, because the environment of China is very much a rule of people (as opposed to rule of law) society, and thus its market economy is far from mature. Since partners from similar East Asia cultures can simplify the negotiation process and speed up the establishment of JVs (Cheng, 1997), thus will experience better communication and working relationships than Sino-Western JVs (Li, Lam & Qian, 2001). Tsang (1998) asserts that guanxi can provide individuals and firms with an imperfectly imitable resource that provides a distinct competitive advantage over others in modern day China. Business links between Taiwan and China were forged not through official channels but through guanxi networks or grey-market mechanisms (Luo, 2000). Luo (1997) even suggested that in China and in other Chinese communities in East Asia, such as Taiwan, “guanxi” is said to be critical to business transactions and negotiations.

Proposition 3: Due to the political hostilities, regional economic instability and corruption from Chinese tax or government authorities, and underdeveloped social and economic infrastructures, guanxi contributes significantly to success.

Methodology

The statement made by Liebscher (1998) appears to emphasis that qualitative and quantitative research approaches should support each other, because “a quantitative research methodology is appropriate where quantifiable measures of variables of interest are possible, where hypotheses can be formulated and tested, and inferences drawn from samples to populations. Qualitative method, on the other hand, are appropriate when the phenomena under study are complex, are social in nature, and do not lend themselves to quantification” (p. 669). Thus, the “triangulation method” is particularly relevant in this research, where the international strategies of the Taiwanese firms cannot be separated from the political and economic context in which they develop. In addition a continually changing environment results in an evolving strategy. For the qualitative research approach, following Creswell’s strategies for research designs that we have selected interview technique for the initial data collection through observations, interviewing, documents and arti-facts (Creswell, 1998, p. 122). The research focuses on the extension and combination of two areas of theory: joint venture formation and the role of guanxi in the Chinese Diaspora. The depth of information
the twenty cases of interviews (face to face) were likely to provide was particularly important
in this preliminary stage of the research. While, for the quantitative research approach, we
posted 1,000 questionnaires to selected sample to generate a large number of cases of Sino-
Taiwanese JVs for comparison statistically through aggregated data and information.
Consisted with research topic, we used the semi closed-ended questionnaire is structured in
both Chinese and English versions. Practically, the questionnaire was designed based on
easy to understand method of measurement (Ali K and Eggert, 1999).

Results:

This research was designed to investigate evidence to confirm or refute the three propositions
outlined above: first the choice of JV as an entry mode into the PR China, secondly, the
importance of cultural similarity to Sino-Taiwanese JV formation, and thirdly, the
significance of guanxi to JV formation and success. All interview cases appear to indicate
that the Taiwanese investors are concerning the significance of timing and the need in the
Chinese market. As state earlier, the cultural context in Taiwan has many similarities to that
of PR China. We asked the questions of the importance of culture factor in the JV formation,
most of the interviewed managers claimed that cultural similarity did help them in the process
of negotiation and establish trust for each other. Such result reflects the fact that Taiwanese
businessmen are greater understanding of the “Ways of the East.” (Li, Labig and Chalmer,
2001) than other foreign investors. And the interview findings suggest that guanxi holds a key
to run Sino-Taiwanese venture businesses more flexible and smooth because the indigenous
of collaboration relations have been profoundly embedding in the Chinese characters
(collectivism).

For example, the general manager of STJV 2, he said “China absolutely is a great potential
market and opportunity for Taiwanese businesses,” he predicted “in next couple of years,
there will be huge political changes, which would make Taiwanese firms run more efficiency
in China.” When his firm entered the Chinese market in early 1990s via joint venture
formation, his main object was to use the abundant cheap labour to produce blenders and
electronic magnet stoves. However, after three years operation in China because there was a
big conflict of control issues with Chinese partner the venture was terminated in 2000 and
changed to wholly owned subsidiary. Like other foreign investors that they have attracted by
the vast size and potential of the Chinese market and raw materials; therefore, they jointed
with the Chinese State Owned Enterprise (SOE) in Guangshi to co-explore resources and
distribute to the market in 2000. The procurement manager sharply points out that the biggest
challenge is not about producing the products but to deal with government officials and
authorities. Thus, guanxi comes to play a very significant role in forest business.

On the contrary, the aim of the empirical part of this study was to make an exploratory
investigation of strategic motives of Taiwanese companies for using IJV in China. There is no
previous survey has been made of the issues about Sino-Taiwanese JVs, mail survey
technique is an important procedure for us to investigate such phenomenon. Consequently,
there were 132 feedbacks which were answered by either top executives or senior managers,
e.g., president, chairmen, general managers or deputy managers. The main part of this survey
was 56.8 percent of those Taiwanese firms formed an equity joint venture with Chinese
partners to reach Chinese domestic market. More importantly, there was 78 percent of total
Taiwanese investor agrees that the Mainland market is or will be a very important market for
their businesses. Another important point was that among the total selected sample of Sino-
Taiwanese JVs, 21.2 percent of them believed that to have a good guanxi is the first priority
to access Chinese market, and 70 percent of them shows that culture similarities have gained a great advantage for Taiwanese venture parent firms to cooperate with its Chinese partners.

The findings appear to indicate that Taiwanese businesses were gradually shifting from labour intensive to capital-intensive market; therefore, JV provides a vehicle to reach the Chinese domestic market. As we stated earlier that the characteristics of JV literature which related to the risk sharing, resource pooling, asset protecting, and the ability to react quickly to market changes (Pan & Tse, 1996), it offers Taiwanese venture parent firms the advantages in operating their subsidiaries in China. Child (1994) indicates that foreign investing companies often wish to have access to China’s market and the Chinese may desire joint venture to facilitate the access of their parents to international markets. For Taiwanese investors who largely come from the small and medium businesses, are usually lacking of sufficient investments and are unfamiliar with the Chinese market. In the same manner, the objective of Chinese partners is to find a partner who can offer equipment, skill and technology to develop its economy. Consequently, joint venture formation becomes a desirable strategy for Taiwan companies to enter Chinese domestic market and WTO will help them expand quickly to achieve business success in China.

Conclusion and discussion

We think that this paper can make a significant contribution to the convention literature on international joint venture by analysis the case of Sino-Taiwanese joint venture to examine the differences and similarities with other Sino-foreign joint ventures. Typically, China is still largely depending on the foreign investments in businesses to develop its economy, we argued that Taiwan can be served as an important link to engage the global market by utilising and learning the skills and experience from Taiwanese venture partners. Our current results predict that the Sino-Taiwanese joint ventures may have brought new ways of thinking on how to operate firms with increased efficiency based on the effect of the similar culture context. Second, the joint venture formation obviously is still one of the important entry strategies for Taiwanese firms at least in some specific industrial sectors, hence the relationship between Taiwanese and Chinese partner will become even tighter and crucial in the future. More precisely, because Sino-Taiwanese joint ventures will keep practising in many areas, it will capture the significance of culture-performance linkage. In other words, it related to the strong theoretical guanxi reason, since China lacks of well-developed legal and distribution systems, guanxi seems provide the only efficient alternative for foreign firms (Luo, 1997; Kiong & Kee, 1998; Wang & Chan, 1999). Since Taiwan and China became the member of the World Trade Organization (WTO) in 2002, both governments promised they would gradually relax on the restricted items to fulfil the commitment of market liberalization. Indeed, it shall result in more competitiveness environment for Taiwan and China’s enterprises, but through the collaboration strategy can enhance their competitive advantages in order to face challenges in the global competitions.

Reference:


PARTNERSHIP WORKING - ISSUES OF INFORMATION FLOW

Greta Squire

Abstract

Crime and Disorder Reduction Partnerships (CDRP) have been developed through the Crime and Disorder Act (1998) and amended by the Police Reform Act (2002), which placed a statutory requirement upon local Councils and the Police to perform a 3 yearly audit of crime within their area and from this produce a crime reduction strategy. One of the key areas encouraged by the Home Office to be included within these strategies has been domestic violence.

The CDRP is essentially a strategic body with data being fed into it from partner organisations, which are required to work in co-operation with private, voluntary, plus other public and community groups, including the community itself. This particular process of how information is fed from the local agency groups to the strategic CDRP is the basis for this paper.

We will focus the discussion upon the current structure of the CDRP within Cornwall, and how information is disseminated between the strategic CDRP and the operational agencies and community groups, with focus aimed specifically at the role of domestic violence forum groups within this process. The discussion will examine the structure of partnerships within Cornwall and the domestic violence service sector, and what involvement if any the local forum groups have in the contribution toward policy development.

The basis for this paper and the foundation for considering researching a PhD began through a desire to gain a clearer understanding of domestic violence policy and research through to actual service provision. By investigating further it became apparent that domestic violence had become a key target area for crime reduction with large levels of funding being provided for Home Office research, plus both perpetrator and survivor projects.

Through this study, it became evident that in-order to understand more about domestic violence policy and front-line service provision, a more in-depth review of ‘community safety’ policy would be required.

There has been a huge shift in policy direction through the introduction of a more ‘managerialization’ of crime prevention (Hughes.G: 1998; Hughes.G 2002; Matthews.R: 2001), which has placed a greater emphasis upon promoting an almost business-model philosophy to public sector run agencies. Within the remit of criminal justice, it manifests itself through the introduction in 1998 of Crime and Disorder Reduction Partnerships (CDRP), which emerged through the Crime and Disorder Act (1998). The Act placed a statutory requirement upon local councils and the police to perform crime audits, based very much upon a business model; crime data is collated and presented; and a strategy developed based upon the crime level data gathered. The Home Office places a requirement upon local councils to target particular crime areas such as domestic violence and burglary.
within these strategies, and provide funding based upon projects and initiatives introduced to tackle these crime types.

Through further investigation, it became evident that over the past few years domestic violence policy had become a key target area for the Home Office, with two particular documents, which highlighted the issue of the involvement of forum or community groups within the Community Safety remit. Firstly the ‘National Domestic Violence Delivery Plan’, a report commissioned by an Inter-Ministerial Group, set up in 2003. The report itself was delivered in 2005, with a list of domestic violence service commitments that included a section on ‘more support for the Voluntary Sector in the setting of standards’ (National Plan: 2005:15). This commitment involved the closer working relationship between the statutory and voluntary agencies, with the voluntary agencies providing key frontline service provision for both agency training and education, and providing services for survivors. Alongside this document, was another key text, again governmental in origin, the Home Office Report: Domestic Violence and Crime and Disorder Reduction Partnerships: Findings from a Self Completion Questionnaire’. This report aimed to ‘provide a comprehensive picture of the work of CDRPs in relation to domestic violence. A mapping exercise was conducted to assess the nature and extent of work undertaken by CDRPs on domestic violence’. (Home Office no 56/04: 2004: IV). The report lists several recommendations for further research, one being ‘to identify the ways in which domestic violence fora impact upon the effectiveness of CDRP’s in their responses to domestic violence and to compare the different models. (Ibid: 27). From both these texts, it became clear that although the rhetoric was positive regarding the work of partnerships, and how there needs to be continuing involvement from both statutory and voluntary agencies, it was still unclear as to how partnerships do actually ‘work-together’, and to what extent local forum groups’ impact upon the locally delivered crime reduction strategies. However, we will begin by providing an overview of the current literature is regarding what Crime and Disorder Reduction Partnerships are, together with how and why they were developed. We will then go on to examine what the process is for partnerships for gaining evidence-based data. We will review the literature on what data is collated, and by whom, and how this filters through from each agency through to the Strategic Group. We will then go on to focus our discussion upon the specific Crime and Disorder Reduction Partnership structure within the South West Peninsula, to examine how data is collated there, with an emphasis placed upon the involvement of both statutory and voluntary agencies, and local community groups to ascertain how their involvement affects the overall Strategic Plan.

The term ‘Crime and Disorder Reduction Partnerships’ borne through the Morgan Report or to give it its full title the ‘Safer Communities: The Local Delivery of Crime Prevention through the Partnership Approach. (Home Office: 1991)’ has been at the crux of crime prevention policy from the advent of the New Labour Government of 1997. Whereby, the emphasis has been placed upon ‘the notion that crime prevention cannot be effectively tackled by the police nor indeed by any agency on its own, instead what is required is a holistic approach whereby effort, information, resources and expertise are shared and co-ordinated among key agencies’.

---

Although there were government Circulars as early as 1984, under the Conservative Government, which believed the guiding principle for crime prevention, should be to involve the whole community. (Home Office: 1984). However, it was not until 1998 with the advent of a new political party in Government, which introduced early on in their term; the Crime and Disorder Act - which made it onto the Statute books - which actually placed a legal requirement upon the local authorities, the police, health, and probation committees to work together to tackle the problems of crime and disorder within their local geographical area. Section 6 of the Act stipulated that this was to be done through an auditing system, whereby on a three-yearly cycle, all of the key statutory agencies plus others from the voluntary sector and community groups; would produce and publish a workable strategy laying-out their policies and targets for reducing crime within their geographical location. In 2002 this was amended by the Police Act, which added a clause to include drug misuse alongside crime and disorder.

The coming-together of such groups have been referred to as Crime and Disorder Reduction Partnerships (CDRP) there are currently 376 within England and Wales, their remit being to provide a clear statistical analysis of crime within a particular location, and from there formulate a strategy in which to reduce that particular crime problem.

The justification for providing such data is so that CDRP’s can produce and provide clear data on crime and disorder, and since 2002, drug misuse; so to inform policy makers, but to also provide a document for the public to review and report their views.

The audit process described by Power in 1997 states that:

‘Audit has become a benchmark for securing the legitimacy of organisational action in which auditable standards of performance have been created not merely to provide for substantive internal improvements to the quality of service but to make these improvements externally verifiable via acts of certification. As the state has become increasingly and explicitly committed to an indirect supervisory role, audit and accounting practices have assumed a decisive function. (Power: 1997:10-11).

This statement reinforces the rhetoric provided by the Government regarding the CDRP auditing procedure. The audit process is claimed to be a way in which performance of local police and councils can be measured, providing not only the Home Office, but also the ‘community’ a quantifiable measure of crime, and the reduction methods introduced in response. As Power goes onto state ‘the state cannot play this indirect role without assuming the efficacy of these practices at the foot of a regulatory hierarchy’. (Power: 1997: 10-11).

In response to the Governments requirement for audits and strategies, they produced ‘toolkits’ with which to advise Partnerships on how to develop such strategies. (HomeOffice2000)(www.crimereduction.gov.uk/toolkits/index.htm). The toolkits provide step by step advice on what an audit should look like, what information needs to be included, where that information should be sourced. It provides ‘performance management priority areas’ for which the completed audit should include, such as aims and objectives. A key objective is to provide a SMART strategy. This being; Specific, Measurable, Achievable, Realistic and Time-bound. However, there were still considerable variations in what each of the 376 CDRP’s provided within their audit and strategy. (Home Office 8/00: 2000). As we can see
the research conducted into the work of the partnerships was commissioned very early on within the policy, and therefore we can expect an improvement in how strategies are produced. Within the domestic violence sector there has been an improvement in how strategies are produced with many developing a stand-alone strategy to tackle particular crime groups such as domestic violence. (Diamond et al: 2004).

Within the auditing process it has become a requirement for the inclusion of groups within the ‘community’ to be part of the strategic process, we will therefore consider the issues around what we mean by ‘community’, and how this can prove to be problematic for those involved within the strategic process.

By coining the phrase ‘community participation’ we are placing an emphasis upon ‘community’. Hughes states that ‘the idea of community has long been an integral part of common-sense thinking about a wide diversity of social problems and experiences. Just like the concept of prevention, community maybe seen as a ‘feel-good’ word, carrying the presumption that communities are by their nature ‘a good thing’. (Hughes.G 1998: 105).

Over the last twenty years, there has been an increased emphasis by Governments of both political persuasions, to promote the involvement of community participation. This is not just within the criminal justice area, but covers public-sector policy generally. Through such initiatives as Neighbourhood Watch, which it is now estimated to include around 150,000 schemes around the UK with approximately 5 million members, there has been a belief that through a partnership approach between the community and police, crime can be reduced. The success of Neighbourhood Watch may be limited especially within the geographical locations that it is most needed, but it has however provided useful rhetoric regarding the involvement of communities within the criminal justice process. The Governments requirement for there to be involvement by the community within the CDRP structure should include ‘hard to reach’ groups such as victims of domestic violence, ethnic minorities, travellers and sex workers as an example. However, what has become apparent through the work of Forrest et al is that the individuals most involved, as in the case of Neighbourhood Watch; are white, middle-class and affluent, not representatives from the ‘hard to reach’ community. (Forrest et al: 2004: 4).

In Section 5(2) (c) and 5(3) of the Crime and Disorder Act it stipulated a specific list of community groups that should be included within the partnership process, they include - Victim Support, Neighbourhood Watch, and any groups related to the work of women, young people, the elderly, the disabled, ethnic minorities, gay and lesbian people, religious groups, residents, shopkeepers and any other relevant local agencies.

This diverse grouping covers a large variation in the types of voluntary or community organisation. For example, Victim Support or Neighbourhood Watch has a large number of members nationally, both paid and unpaid, so to compare this with a smaller localised group, that in many cases relies totally upon their volunteer workforce can be quite problematic. Hester in his research on the role of voluntary and agency groups within the partnership process states that ‘it is recognised that specific voluntary and community organisations can best contribute to partnerships in different ways, and that one standardised consultative approach is inappropriate’. (Home Office 10/00: 2000: 2)

It is worth highlighting that from the initial contact the researcher has had with the community-based groups willing to participate within this particular research, the
standardised practices of the statutory agencies and some of the large voluntary agencies such as Victim Support has left the smaller voluntary groups feeling marginalized within the strategic process, and has often left them questioning what their involvement could and should be.

The ‘community’ in this instance has taken the appearance of forum groups set up to provide a platform more often than not for survivor groups, many are women-centred, which introduces another dimension to the work of domestic violence partnership working, as issues of patriarchal-hierarchies within the partnership structure can be accused. Hague describes this as the ‘lowest common denominator effect, with many forums having to struggle to evolve common visions, to develop women-centred analyses of domestic violence based on ideas about gender and power in relationships’. (Hague.G: 1999:18).

We will now go on to consider further the work being carried out within the South-West Peninsula. The County has seven separate CDRP’s, with the new recommendations for change in the structure being implemented in March 2003. Each CDRP has its own local strategic plan, but in addition there is also a countywide partnership, which evolved through a recognised need to work together more closely. This is the ‘policy operational group’ system whereby experts in the key crime areas sit, including representatives from private, public, voluntary and community sectors. These working groups meet on a regular basis throughout the year to formulate policy and action plans for strategies, which are then sent for agreement by the Practitioners Group, and finally given overall approval by the Strategic Group.

Policy Operational Groups or POG’s as they are commonly referred too; are working sub-groups; twelve in total, which have been given the responsibility of creating countywide community safety strategies. The sub-groups are:

Audit and Performance Management
Anti-Social Behaviour
Arson
Burglary Dwelling
Domestic Violence
Fear of Crime
Partnership Development
Road Safety
Drug, Alcohol and Substance Misuse
Vehicle Crime
Violent Crime
Youth Offending

They all signed up to a protocol, which states that they all agree to contribute to priority areas as highlighted as of particular problem crime areas.

The POG’s each report to the County-wide Community Safety Practitioners Groups, which in turn reports to the Community Safety Strategic Group. This three-tier authority structure is believed to not only reduce duplication and improve cross-border working, but also larger initiatives can now be tackled, with a pooling of resources. How effective this is however, can be ascertained through the auditing process and through independent evaluation research.
The Strategic and Practitioner Groups are made up of representatives from the Local District and County Councils, the County Fire Brigade, Primary Care Trust, Drug and Alcohol Action Team, Devon and Cornwall Probation, the local Criminal Justice Board, and the Devon and Cornwall Constabulary. The only representative from a Voluntary Agency is through Victim Support, one of larger of the charitable organisations.

According to the Partnership Constitution and Structure, the Strategic group is responsible for:

- Proposing on an annual basis district and countywide crime reduction work that requires the involvement of more than one CDRP.
- Agreeing individual contributions to the central community safety budget.
- Agreeing and approving budget allocations for the countywide CDRP projects.
- Agreeing and establishing priority taskings for the Community Safety Practitioners Group.
- Reviewing progress and receiving reports on key issues, including finance.

The role of the Community Safety Practitioners Group is as follows:

- To implement the directives of the Strategic group
- To appoint lead officers for the POG’s.
- To oversee the delivery of combined community safety strategies.
- To provide quarterly progress reports and make recommendations to the Strategic Group.

As we can see, the criteria for each level of the system has been documented very much within the managerial business-model, with an emphasis been placed upon the aims, objectives and structure of each tier of the process, which is in keeping with the ‘managerialization process’ discussed earlier in this paper.

Supporters of the system believe that rather than the system being more bureaucratic with the establishment of another level of management, it has been hailed as a great success, with its followers believing that it has improved the sense of direction for Community Safety across the region.

The data for the audits comes through the information hub named Project Amethyst. Which has been described as ‘the web enabled crime mapping and data exchange hub, which sits at the centre of the partnership, providing data and geographical information to CDRP’s within a secure environment’. (www.amethyst.gov.uk). The Policy Operational Groups provide progress reports for the Amethyst site on a quarterly basis that present data on how targets are being met, with one of its main criteria being to reduce the public’s perception of crime.

This overview of the procedure within the South-West Peninsula very much endorses the rhetoric provided by the Home Office for the performance of Partnerships. It has actually been hailed as a success, and a model of Good Practice for other Counties to emulate. By bringing together ‘experts’ within each field it is believed that a greater understanding of the issues will be gained. However, work completed by the University of Plymouth in 2005 for their Crime and Disorder Audits, showed that although the structure was in place for improved partnership
working, the reality was not quite as positive. It has been best described by Tilley and Pawson in their 1997 work ‘Realistic Evaluation’, that the programme and policy can be in place within an organisation, however the mechanism for it to work effectively may not be activated within that particular context, within that particular group. The POG system is still very much in its infancy, as is the case of the statutory CDRP structure. However, it is apparent from initial discussion with representatives form the voluntary and community sector, that their contribution is often seen as tokenism, with much of the day to day meetings and discussions being performed on more of an informal basis, behind closed doors, without the involvement of the voluntary or community groups.

Another issue that has been highlighted through initial discussions relates to the idea that often there are two distinct ‘camps’ within the partnership meetings - the statutory agencies and the voluntary agencies. However, an important point was raised about there being disagreement between the voluntary groups within the partnership process, with some agencies working to a more formal structure, compared to others that work within a more informal format, which has caused conflict within the meetings. This again returns to the issues raised by Tilley and Pawson regarding the effectiveness of the partnership group, within its own context.

In conclusion, it must be noted that the partnerships within the South-West Peninsula are acclimatising to a new structure, which understandably does take time. However, many of the relationships forged between the partnerships members had been made prior to the change in structure, with many of the issues remaining unresolved. Partnerships are only successful if they are able to acknowledge and include all its members, but also find that illusive spark that ignites the mechanism within their particular group, which leads to a successful partnership. Although the basis for the audit procedure is based upon a business-model, such standardisation cannot ignore the fact that a partnership is made up of individuals, harbouring differing agendas and backgrounds, which needs to be acknowledged when attempting to understand what constitutes a successful collaboration. The basis for this paper was to examine the structure of partnerships within the South-West Peninsula, and in particular the work of the domestic violence sector, to gain a clearer understanding of what the involvement of local community or forum groups is within this process. When examining domestic violence research, there is an additional area of interest in the diverse beliefs of those involved. This is no different within the CDRP structure. There is a very fine balance to be made between adhering to the formalised structure for the formulation of the CDRP Strategy, in contrast to how this encompasses the often radical views of forum groups or community groups. It is hoped that through a continuation of this research, we will gain a clearer understanding of how and if this balance has been achieved.

Bibliography.


www.amethyst.gov.uk
MCPherson ‘One step black’
DIVERSITY, DISCRIMINATION AND PREJUDICE
WITHIN THE POLICE SERVICE

Mark Harron

Abstract

There exists a mass of excellent and probing research into the British policing system, spanning a range of disciplines; all topical, all relevant and all incredibly provocative in terms of moving the ‘institutional beast’ of Policing into modernity. Perhaps one of the most published and profound pieces of research (in the widest sense) is that of the McPherson Report (1999), (The Stephen Lawrence Inquiry). The driving force behind the report was the operational processes / failures of the Metropolitan Police Service in investigating the killing of the black teenager, Stephen Lawrence, in 1993. Substantively addressing matters of policy, process and organisational failure within the Metropolitan Police Service, the report goes on to recommend that all Police services should adopt policies and practices that help to engage all sections of the community. Probably the most damaging assertion within the report was a reference to the Metropolitan Police Service as being ‘institutionally racist’ in nature. Institutionalised racism was defined by the inquiry as;

‘The collective failure of an organisation to provide an appropriate and professional service to people because of their colour, ethnic origin or culture. It can be seen or detected in processes, attitudes and behaviour which amount to discrimination through unwitting prejudice, thoughtlessness and racist stereotyping which disadvantage minority ethnic people’

This paper seeks to identify a particular area of exclusion that has been left behind by the rhetoric of ‘McPherson’, or perhaps more specifically, the misapprehension and sterility that the ‘McPherson’ legacy has engaged. It then endeavours to provoke and stimulate a programme of research in addressing some of the more ingrained and endemic facets of institutionalised policing; ignorance, inaccessibility, indifference, remoteness and organisational and collective rhetoric. However, despite this legacy, the shining sword of ‘McPherson’ has created a new political and organisational phenomenon. Never has there been so much valuable currency in the terms ‘ethnicity’, ‘diversity’, ‘race’, ‘gender’ or ‘equality’, than post the ‘McPherson’ enlightenment.

---

21 Paragraph 6.34 The Stephen Lawrence Inquiry 1999
‘McPherson’ forms the start of a journey into a very closed and often misunderstood part of the British policing system, namely the Special Constabulary. Through a labyrinth of material, the paper contends that the management and integration of ethnicity, diversity and cultural identity has more foundation in prejudice, ignorance and visual identifications of colour (or more appropriately the absence of colour), than in the human need to embrace and encourage integration in the widest sense possible. Race, ethnicity and gender may have inadvertently become the only visible and valuable commodities in building our police services and delivering policing to our communities. The paper moves to raise suggestions for both quantitative and qualitative research models to address some of the recruiting dilemmas that face British police services, insofar as they can be related to the policing volunteer.

The British policing system has become enslaved by policy driven schemes of recruitment and development. Arguably all well intentioned and all aiming for an admirable utopia of ethnic, racial and cultural equality, both in terms of operational practices and indeed as important, in the make-up of its' staff. Chasing statistical targets and goals may serve a need in itself but behind every policy implementation, achievement or failure exist real and tangible individuals. These individuals, like all humans, react both positively and negatively to their surroundings and are often the overlooked and undervalued building blocks of tomorrow’s ‘institutional diversity’. It is therefore undoubtedly time to look at things somewhat differently.

Introduction

The concept of the ‘Special Constable’ can be traced back to the Saxon period (410-1066 AD) when every citizen was bound by an oath known as the ‘frank pledge; ‘a duty to assist in upholding the peace and repressing crime’. During the reign of King Charles II (1649 -1685) the concept was again re-affirmed when the Crown extended the common law to compel any man to the role of ‘Temporary peace officer’ in times of unrest. The first statutory enactment allowing for the appointment of Special constables arose in 1831 (Special Constables Act 1831), some 2 years after the establishment of the first full time paid Police service, (Metropolitan Police Act 1829), (Leon 1987).

At the time of writing, the latest recorded Home Office figures show a headcount of 11,918 officers serving within the Special Constabulary. This contrasts with a total of 144,535 regular police officers and 6268 Police community support officers (PCSO) throughout the 43 Police forces in England and Wales. Therefore, the Special Constabulary provides a sizeable addition (7.9% of all available uniformed police presence) to the regular Police service. There has however been an incredible decline in the special constabulary over the past decade. Government statistics show a real time reduction in Special constable numbers since 1992 where the headcount strength stabilised at 17543, rising to a height of 20,573 in

---

22 Jan 2006
23 Police service strength England and Wales Home Office Bulletin 12/05
24 Research and Development Statistics, Home Office 2002
1993 and then falling overall to 12,068 officers in 2001. Indeed, the first annual increase for over 10 years occurred between 2003 /2004 where the numbers increased by nearly 1000 officers to 11,918. This may have been due in part, to a recent funding scheme established in 2003 by the Home Office.

The ‘Capacity Building Scheme’ was established in 2003 in order to build the strength of the Special Constabulary and thereby reverse it’s incredible decline. The scheme provides for up to £70,000 per annum for each of the 43 forces (with a managing police authority) within England and Wales, with the requirement that each force applying for funding, agrees to undertake action to increase it’s strength of Special constables by a real term net increase of 25 officers each year.

Funding has been designated for specific civilian support roles of Co-ordinator, administrative assistant and recruitment personnel along with the provision of some of the available funding for extra training support staff. The funding allocated was intended to be available over a two year pilot period but in late 2005, the Minister of State for Policing, Hazel Blears, indicated that the funding arrangements would continue until 2008 in line with the Home Office’s ‘police reform’ agenda and the national ‘neighbourhood policing’ model.

By August 2005, official Home Office figures indicated that the headcount strength of the Special Constabulary had reached 11,856 officers across 42 of England and Wales Police force areas. These figures omit to show headcount numbers of Special constables for one force area and may account for the numerical anomaly. The highest headcount of Special constables during peacetime was approx. 144,000 officers in 1931. (Morgan 1987)

Gender / Ethnicity Divide

In terms of comparison and in accordance with the measures used in the Home Office data on Police strength 2005, a total figure of 142,795 (FTE) officers will be used and for Police community support officers (PCSO) a figure of 6214 officers. The 11,918 Special Constabulary officer numbers remain the same in line with the statistical analysis.

(Gender) From the base line figures, 30,162 (21%) of the 142,795 Police officers are female, 3813 (32%) of the 11,918 Special constables are female and 2609 (42%) of the 6214 Police community support officers are female. (62% of all civilian police staff are female)

(Ethnicity) From the base line figures, 5017 (3.5%) of the 142,795 Police officers are minority ethnic Police officers (self classified), 715 (6%) of the 11,918 Special constables are minority ethnic officers and 870 (14%) of the 6214 Police community support officers are minority ethnic officers.

At a glance the information clearly points to a significant disparity of both ethnicity and gender between the traditional regular police officer role and that of...

---

25 Home Office Statistics
26 Home Office Capacity building scheme. Financial resources available for recruitment and training along with administrative tasks. Not directly available for resources or development of existing officers although indirectly these may occur.
27 As published on Home Office Special Constables website update
28 August 2005 figure (11856) compared to earlier 2005 Home Office bulletin 12/05 figure (11918) a reduction difference of 62 officers
29 Police service strength England and Wales Home Office Bulletin 12/05
30 Support staff
Special constable and (PCSO). In terms of potential visible uniformed police presence (PCSO and Special constable) together, PCSO and Special constables represent approx. (12.7%) of all available uniformed front line officers and yet have provided some (17.6%) of all available female uniformed patrol officers. Of greater statistical significance is an incredible (24%) availability of all minority ethnic officers across England and Wales.

The Home Office has tasked itself in raising the number of both PCSO’s and Special constables throughout England and Wales. The projected PCSO figure of 24,000 officers by 2008. If gender / ethnic profiles continue along the same lines, then the (PCSO’s) profile will provide an extra 7471 female and 2490 minority ethnic officers available for visible police patrol. Similarly, although no specific target has been set for Special Constabulary strength, it is anticipated that in line with the Capacity building scheme requirements, each force that takes part in the scheme will have a net increase of at least 25 additional officers for each year funding has been allocated. For the purposes of statistical analysis, an increase in the headcount of Special constables to 20,000 by 2008 will be used. Assuming the same ethnic and gender profile is achieved, the Special constabulary profile would provide an extra 2587 female and 485 minority ethnic officers.

Diversification

Probably the most interesting and certainly thought provoking pieces of literature on the concept of diversity has been produced by Kandola & Fullerton (1998). Here the subject matter of diversity has been stretched and academically manipulated to extend to every facet of human interaction or identity. Couple this with the Oxford English Dictionary definition of ‘Diversity’ as meaning; ‘the state of being diverse’ – ‘a diverse range; a variety’, and one can easily see how something as simple as dialect or indeed volunteer police officers, could be considered as a matter of diversity.

Much interest in policing reform, particularly in relation to staffing, centres around the cohabitation of the word ‘diversity’ alongside words like ‘ethnic’, ‘gender’ and ‘race’. However, the politicisation and momentum behind racial and ethnic diversity is nothing new, it just seems to have been given a new lease of life. The current Labour Government has endeavoured to force through diversification by setting employment targets for police services. The aim being to achieve a representative proportion of police officers in each force area relative to the communities they police.

Therefore, there seems to be an inevitable assumption that whether minority ethnic communities like it or not, police services will manage to recruit officers to a level of relative equilibrium. Would this then mean that a force area with no minority ethnic inhabitants escapes from the requirement and could choose not to recruit a minority ethnic officer base whatsoever? As Johns (2004) explains, whilst discussing other problems in the public service field, ‘it may be necessary to establish a target in line with the national demographic profile’. This would certainly assist in navigating the problem, rather than forcing some police forces to adopt meaningless targets that do not properly correlate to their own geographical demography. Whether forcing through diversification (ethnic or otherwise), can ever be a substitute for tangible and meaningful community integration has yet to be tested. Research undertaken by Holdaway (1991), between 1987 /89, suggests that the imposition of any initiative that could be perceived as showing any type of...
favouritism or selected treatment to a particular minority group, could have the reverse intended effect and culminate in resentment and total disregard for the policy or initiative under consideration, both from within and outside of the police service.

The changing model of policing

Plural policing, the wider police family, the extended police family, partnership policing, community policing and the mixed policing economy are all incredibly vibrant and popular terms in describing the vision for policing in England and Wales. As Mawby (2005) describes ‘Policing is being transformed and restructured in the modern world’.

Alongside the traditional icon of constable, the public are likely to see an array of individuals performing various tasks which were historically the domain of the regular police officer. PCSO’s were introduced by virtue of the Police Reform Act 2002 along with the power of a Chief Constable to designate certain members of staff to fulfil duties traditionally performed by sworn police officers, (Crawford et al 2005). These designated officers, where empowered, can fulfil roles such as PCSO officer, investigation officer and escort officer to name a few of the designated roles. Un-sworn police volunteers have also been gradually recruited over the past 2 years in furtherance of the police reform agenda and fulfil a range of non-patrol activities vital to the efficient performance of the policing system. Alongside all of these staff, you will find the Special constable.

At the base of the Special Constabulary model can be found a unique and historically significant relationship between the community as service users of the policing system and members of that community who dawn the office of constable and patrol with the regular police officer. Gill & Mawby (1990) wrote significantly about research undertaken in the late 1980’s centred upon the Special Constabulary. This provided a unique and previously un-researched area of police sub-culture and facilitated a new insight into police volunteerism. Amongst an abundance of revelations, the research indicates that motivational drivers for joining and remaining in the police service as a volunteer officer are directed more towards the notion of ‘giving something back’ and ‘doing something good’ for the community rather than the often perceived motivators of fast cars and / or immense power. The research also identified that candidates applying to join the Special Constabulary were typically white, employed and not necessarily those regular police applicants who had failed the selection tests for the regular paid police force.

It should be noted that there is nothing unique about volunteering in the United Kingdom. 51% of all adults undertake a variety of voluntary activities, contributing incredibly to the economy in terms of unpaid services, some of which would necessarily otherwise be funded by government. There is also nothing particularly unique about volunteering in public bodies. A plethora of volunteers operate throughout the Criminal Justice system (Gill & Mawby 1990). What is unique however about the Special Constabulary is the fact that these members of the public who volunteer their spare time to undertake routine policing patrol, often in their local community, are subject to the same responsibility / accountability and legal recourse as that of their full time paid equivalents. There are no statutory exemptions, rules, policies, practices or codes that will avail these

---

214

Institute of volunteering research 2005
officers of greater protection other than the protection afforded to their full time paid regular colleagues. They will however have the same protection in these areas as that of full time paid officers.

From an internal organisational perspective, there is little additional protection to found. There is no current trade union representation and historically the Police Federation have always taken a rather reserved position to the Special Constabulary. Indeed, it was not until the arrival of the first PCSO officers in 2002 that the Federation started to publicly support the role of the Special Constabulary. There is concern amongst the Police Federation that PCSO appointments are back filling places which would otherwise be filled by additional full time police constables. As Whitaker (1979) indicates, the presence of Special constables was often seen as one which took away the opportunity for paid overtime for the regular police service. Have we really moved that further forward in building the wider police family? Announcements made at the Police Federation annual conference in 2005 indicated that they would now look at ways to provide membership to members of the Special constabulary.

Currently, although sworn officers of the Crown, Special Constables do not hold the position of ‘employee’ within the police force that they operate within. Therefore, the vast majority of employment legislation protection does not apply. As volunteers operating within a public body (as part of the wider police family), these unique individuals have full statutory police powers of arrest, both within their own police force and surrounding force areas. From a public perspective, there is little difference between the Special constable and that of a full time regular police officer. Many force areas have recently withdrawn all insignia that has traditionally visibly identified Special Constables from their regular colleagues. Add to the equation that some Special constables are undertaking more operational hours on active duty than some of their regular paid colleagues, one can imagine the conundrum that thereby arises. For example, place two different people in the same role, with the same responsibility, power and authority, the same uniform and the same operational tasking and yet overtly or inadvertently (the resulting outcome is the same) discriminate against one of them based solely upon the fact that the organisation happens to remunerate / employ one of them and not the other and matters of equality start to germinate. Further, add a body of individuals who, statistically at least, represent a higher proportion of gender and ethnicity and the whole picture somewhat changes.

Discrimination in the Special constabulary arises in a variety of guises. The following are but examples: An absence of specialist courses. A refusal to attach officers to specialised departments. An unjustifiable difference in rank structure. A refusal to equip officers with identical uniform when undertaking identical roles. A failure in acknowledging any perceived authority at a particular rank / level. A difference in the level of expense payments based solely upon a volunteer’s status. Lack of uniform and personal issue equipment. A restriction on the provision of basic courses. Deployment disparity. Inequality in managerial autonomy. Objective, task specific criteria applied differently to different groups of individuals.

Discrimination after all, is simply a difference in treatment based upon anything other than individual merit. Indirect discrimination being something capable of total elusiveness and often cloaked in reasons of economy, efficiency or indifference. The Special Constabulary represent a small but sizeable minority.

34 Police Federation 2005. Detailed on Federation website
cultural group, operating within the British policing system. With no collective voice, no visible identity and little, if any incentive for senior officers, both regular and volunteer, in challenging such discriminatory practices, one can understand why resignation rates almost match those of recruitment. Indeed, wastage rates in the Special Constabulary are almost identical to recruitment gains, Leon (1993). There are many reasons for Special constable resignations as both Gill & Mawby (1990) and Alexander (2001) point out.

However, although family commitments and personal reasons rank highly in most Special constable exit interviews, little is known as to whether these are pre-determined reason codes on exit interview questionnaires or whether they are collated from free-text entries. Obviously balancing the needs of a family and other outside activities is an incredible task for even the most enthusiastic Special. Disorientation and discrimination can often be aggravating factors in facilitating a resignation but may not be the ultimate reason for doing so. The family and other outside activities may just become more important than battling on against the dilemma.

The ‘McPherson’ report identified police organisational failure. How far can it now be said that in the treatment of our Special constables there operates an ‘Institutional prejudice’?

‘The collective failure of an organisation to provide an appropriate and professional service to people because of their colour, ethnic origin or culture. It can be seen or detected in processes, attitudes and behaviour which amount to discrimination through unwitting prejudice, thoughtlessness and cultural stereotyping which disadvantage minority cultural groups’

Discriminating against Special constables may indirectly be discriminating disproportionately against minority ethnic and female officers. If the projected figures of up to 20,000 Special constables are achieved and the ethnic / gender profiles realised, then as against the current Home Office Bulletin 12/05 headcount figures of sworn police officers alone, the special constabulary will represent approx. 17.4% of all uniformed sworn female police officers and approx. 19% of all uniformed sworn minority ethnic police officers.

The Special constabulary has long been an excellent recruiting medium for the regular police service. (Holdaway, 1991) With this in mind, is it now time to raise the ‘McPherson gauntlet’ and realise the cultural significance of the special constabulary or frantically look elsewhere?

The research direction
To date, very little research has been undertaken specifically targeted towards matters of gender, ethnicity and discrimination within the Special constabulary. Substantive pieces undertaken by (Gill & Mawby 1990) and (Leon 1992) address other related aspects of the volunteer police model. Gill and Mawby 1990, chronicle the motivational, sociological and operational aspects of Special constables through a variety of research studies in the 1980’s, whilst Leon 1992, provides an historical review of the Special constabulary within the British policing system.

Other commentators like Crawford & Lister (2004), discuss the special constabulary from the perspective of the widening / extended police family alongside other plural policing providers. Holdaway (1991) provides substantive material on the
recruiting processes and dilemmas facing British police forces in recruiting a multi-racial police service. The material provides some incredibly challenging scenarios that face both minority ethnic applicants and the police services that endeavour to recruit them. Specific cross co-relational aspects could arguably be identified but the boundaries may become somewhat blurred at the edges. Research by Alexander (2001) indicates that there are a variety of reasons why members of the public choose to volunteer for the Special Constabulary, none of which involve direct financial reward. The same cannot arguably be said about paid full time officers. Pearce (1993) discusses organisational behaviour of unpaid workers and specifically points towards the satisfaction that volunteers obtain through volunteering in an organisational setting that respects and values their efforts. But few pieces of research relate specifically to the direct relationship between volunteerism and discrimination from within an organisational setting, as occurs between the managing of paid members of staff and their volunteer counterparts.

Suggestions for both research themes and methodology are as follows:

1. Discrimination, Diversity and Equality. **Document analysis**: Involving recruitment, resignation, developmental and equality material from a selection of Police forces and official statistics.

2. Discrimination, Diversity and Equality. **Self completed questionnaires**: Delivered at the point of recruitment, resignation, promotion application.

3. Discrimination, Diversity and Equality. **Structured Observations**: Tasking groups, management meetings, recruitment focus groups, senior management meetings, operational patrol

4. Discrimination, Diversity and Equality. **Ethnography, overt / covert.** Tasking groups, management meetings, policy group meetings

5. Discrimination, Diversity and Equality. **Interview Groups**: Special constabulary officers, senior and middle level regular police officers

6. Discrimination, Diversity and Equality. **Role exchange research**: Research model where roles of Special constables and regular police officers are reversed in a managed programme (possibly across force areas) in order to provide empirical data for analysis.

**Aim**

The overall research aim is intended to provide contemporary data collection and analysis. In doing so, particularly in the field of policing diversity, one is obviously aware of the constraints placed upon any research model where the participants and main beneficiaries are fully aware of the model’s operating boundaries. To this extent one would like to achieve a research process that is as independent and free from the possibilities of targeted participant manipulation as possible. This is not to suggest that potential research participants will actively manipulate answers and courses of behaviour. However, where a participant is aware of the research process, then the integrity and currency of the collected data may have a different perceived value.

---

35 From an American organisational perspective
Conclusion

Policing diversity in terms of ethnicity, culture and gender has never had such momentum as has been seen since the publication of the ‘McPherson’ report. The mass of recruitment programmes throughout the 43 police forces in England and Wales have become fixated upon the external processes of recruiting into the service rather than examining any endemic reasons why there is such a low application rate from certain minority groups. Change by imposition arguably has less value than that that occurs naturally from within. The face and structure of British policing is certainly changing. Whilst recruiting policies like those recently developed by Avon and Somerset Police36 and positive action processes37 that have been principally developed in the Metropolitan Police Service may go some way to increasing applications from under-represented groups from within our communities, the targeted outcomes from these policies have yet to be fully evaluated.

There is arguable a lot of interest in the variety of recruiting programmes that come close to what can be popularly termed ‘positive discrimination’ Although police forces should be actively encouraged to start looking more widely in terms of innovative recruiting practices and the pushing of boundaries, they should keep in mind that there are strong correlations between ‘successful policies’ and the perceived value of the individuals that benefit from them. Holdaway (1991) clearly indicates that some of the research participants in his study felt uncomfortable with the fact that joining a force on the back of a recruitment policy that was ‘perceived’ to be offering them more favourable treatment than that which was afforded to other applicants.

Statistically at least, the Special Constabulary represents an organisation that attracts a higher proportion of female and minority ethnic officers than the regular paid police service. The reasons as to why are to date, not fully understood. The drive towards community cohesion and true representative diversity is a lot more complex than recruiting figures alone. By recruiting a more diverse base of individuals into an organisation and deploying them in relevant and particular community environments may not necessarily bring about the intended result (Johns, 2004). However, the results may be positively surprising, we just don’t know until we try. Gill & Mawby (1990) were undoubtedly correct in assigning the title to their research. ‘A special constable’. They are indeed, very Special Constables.

Bibliography


Crawford, A., Lister, S. The extended policing family: Joseph Rowntree Foundation

Gill, M.L., Mawby, R.I (1990) A Special Constable, A study of the police reserve, Ashgate Publishing Group

---

36 Avon and Somerset Police 2005. A positive action policy using the rejection of male, white applicants based solely upon their gender / racial profile, in order to redress an historical minority ethnic and gender imbalance.

37 Metropolitan Police Service, positive action team. Established in order to facilitate schemes and policies that increase the number of minority ethnic applications to the police service
Gill, M., Mawby, R. 1990 Volunteers in the criminal justice system: OUP


Leon, C. (1987) the Special Constabulary: A historical view. Special Ed, 6.1


Office for National Statistics (2005),

Pearce, L. 1993 The organisational behaviour of unpaid workers: Routledge
A GENDER AWARE ANALYSIS OF PEACEBUILDING - THEORY, POLICY AND PRACTICE.

Annette Elliott

Abstract

This paper will provide a gender aware analysis of conflict resolution and the building of peace in war-torn societies through an examination of current theory, policy and practice. The intention is to understand the significance of gender differentiated roles, the social constructions of 'masculinity' and 'femininity' and the positionings, power dynamics and strategies that these create within a peacebuilding framework. This will necessitate challenging existing frameworks for peace, both theoretical and practical. It will mean asking ontological and epistemological questions about the foundational values and principles of prevailing international policy designed to incorporate the concepts of ‘gender balance’ and ‘gender mainstreaming’ into current peacebuilding policy. This analysis will question the assumption that these policies are sufficient in themselves to achieve the social transformations necessary to be able to fully integrate a non-discriminatory, emancipatory and human centred approach to building sustainable peace. In order to do this is it necessary to reflect on our understandings of peace and war, and why and how a gender aware perspective is essential to gain a holistic picture of attempts to create a peace that is sustainable.

Over the past two decades, in both academic discourse and in the rhetoric of international policy the understanding of what ‘peace’ is has broadened to incorporate a concept much wider than the traditional notion of the mere absence of war. The notion of peacebuilding in terms of international policy has evolved out of existing concepts of peacemaking and peacekeeping, and now embraces, in theory if not in practice, a wide variety of processes with the aim of rebuilding and restoring war torn societies (Boutros-Ghali, 1995). The notions of re-building and re-structuring however tend to marginalise the concept of transformation that is necessary in most societies to create the environment of equity and inclusion that ‘positive’ peace theorists call for (see Galtung, 1996). This inevitably has significant implications for gender equity and indeed, there is considerable evidence of a trend towards the return to pre-existing patriarchal structures in post-war reconstruction (see: Mazurana et al, 2005, and Meintjes, et al 2001).

It is widely recognised that women are almost completely absent from the negotiating tables and higher institutional levels of decision and policy making and yet it is also acknowledged that women have been deeply involved, largely at community level, in a wide range of activities that promote cooperation and reconciliation within and between communities at war. The pre-eminence of the rhetoric of democracy and human rights in international affairs over the past decade lends much credence, in theory, to the struggle for gender equity and non-discrimination. Yet it appears that within the arena of peace processes and peacebuilding policy relatively small advances have been made in translating gender aware perspectives into gender aware practice.
This work rests on the premise that to assume gender neutrality or, the irrelevance of gender, in the field of conflict resolution and peacebuilding, is to pre-determine a partial, distorted view of international and societal relations which cannot lead to fair conditions for a sustainable peace. Indeed, the conventional notions of peace and conflict may actually help to reproduce instability and the potential for further conflict. The complex multi-faceted nature of violent conflict involves different cultural traditions, values, ideologies, understandings and practices, all of which are 'gendered processes'. By this we mean not only that these processes impact differently on the individual depending on their sex and the role this ascribes to them but also that the differentiated power relations that exist between (and within) the different gender roles impacts heavily on the creation and outcome of these processes. The social constructions of gender roles and the inequalities of power relations that arise from these prescribed roles pervade all human interaction at individual, community, state and international levels of analysis.

There are, of course, multiple factors of identity, including cultural, ethnic, racial, religious, class, age, economic status, and education that impact on ones life experience and thus on identity. Whilst it is accepted that to achieve a gender balance in decision making and implementation processes the focus necessarily falls upon the need to correct the lack of women in these crucial areas, the debate about the relationship of gender to peacebuilding requires a perspective that is equally inclusive of the issues of war and peace as they effect and are effected by both men and women in their ascribed roles.

Understanding War and Peace

In order to understand the concept and processes of building peace we need to understand the ontological and epistemological foundations of accepted understandings of both peace and conflict. As Keen (2001, 2) suggests in his discussion of the differences of (or similarities between) war and peace: ‘in order to think sensibly about peace, we need to think clearly about what war actually is’. Definitions of what constitutes war and what are its causal factors are historically and culturally dependent. Throughout the Westphalian era war has traditionally been depicted as violent conflict usually between two states, for territory, resources or geo-political power. During the Cold War the dominant explanation for armed conflict was linked to ideological differences. However, since the end of the Cold War the ways in which war is analysed and depicted has changed considerably with the rise of ethnopolitics and identity struggles (Kaldor, 1999). Contemporary conflicts are often not events with clear beginnings and ends but are elements of a broader process of social change and have thus been more aptly described as complex political emergencies rather than traditional wars (Goodhand and Hulme, 2000; Duffield, 2001; Pearce, 1999).

Until the multi-dimensional nature of war and, crucially, who is benefiting from it and why is clearly understood, any attempts to create an environment conducive to building sustainable peace is likely to be problematic. The levels of analysis and diverse philosophies on war and peace have spanned individual psychological analyses, religious and spiritual beliefs, philosophies on human nature and community, and analyses of state and international relations. Throughout history the theory and practice of non-violence has been closely related to religious teachings, but religion has also been seen as a causal factor, motivator or supporter of many wars. Many of the fundamental precepts of Christianity are linked to pacifism and yet, along with both Islam and Hinduism, Christianity has been a warrior religion throughout history.
The media image of war frequently portrays it over-simplistically as two opposing sides fighting to win a particular goal. David Keen (2001, 2) points out that there are often much less exposed aims that need to be taken into account: ‘Part of the functions of war may be that it offers a more promising environment for the pursuit of aims that are also prominent in peacetime’.

War is clearly not dysfunctional for all of its actors. Many profit from it either financially or in terms of power, resources and influence. The paradox that war has been and continues to be functional for the arms industries of the developed world, those same countries who are the creators and donors of the UN institutions that promote the principles of peacebuilding, should not be overlooked. Although charged with the primary responsibility for promoting global security, the five permanent members of the UN Security Council are still responsible for 85% of global arms sales (Keen, 2001, 5).

The Liberal Peace Thesis
Peacebuilding in contemporary institutional terms is widely seen to be legitimate as a humanitarian act on the part of the international community without deeper analysis of the underlying assumptions. Paris (2002, 2004) contends that international peacebuilders have promulgated a particular vision of how states should be reconstructed based on the principles of liberal democracy and market-oriented economics, which resembles a more benign version of colonial belief that the Western powers had a duty to ‘civilise’ the less developed parts of the world:

“Peacebuilding missions are not merely exercises in conflict management, but instances of a much larger phenomenon: the globalisation of a particular model of domestic governance - liberal market democracy - from the core to the periphery of the international system.” (Paris, 2002, 638).

Perhaps the notion that this is a more ‘benign’ form of imperialism is also part of the assumptions that need to be scrutinised. Pugh (2000, 2004) has suggested that an analysis of the theoretical and ideological context of peacebuilding would serve to clarify what has traditionally been seen as operations of technical and non-ideological conflict management. Pugh provides a deconstruction of the role of peace support operations which suggests that such operations sustain a particular political world order that continues to privilege the rich and powerful states in their efforts to control ‘unruly’ parts of the world. The underlying purpose of such operations is: “to doctor the dysfunction of the global political economy within a framework of liberal imperialism” (Pugh, 2004, 39). To use the construction ‘unruly’ serves to legitimise the domestic order of the imperial states. The main components of what has been described as the Liberal Peace Thesis (see Paris, 2004 and Richmond, 2005), that is, democratisation, human rights, the rule of law and neo-liberal development of open markets are not only Western derived but are also highly gendered processes. Although neo-liberal ideology rarely explicitly addresses gender and talks in very gender neutral terms such as the individual’, ‘the market’ and ‘human rights’, there is an implicit gender politics embedded in the ideology. The structures of the market economy are usually based on masculine interests, for example, the reduction of the welfare state usually impacts hardest on women and the deregulation of the economy usually places more power in the hands of particular male elites (Connell, 2005). The division of labour that the capitalist economy is based on reinforces gendered power differentials as what has traditionally been women’s work - reproduction, homemaking and many of the caring and welfare roles are unrewarded and under rewarded financially and consequently viewed as inferior. The question that
needs to be asked here is are these issues acknowledged and addressed within the framework of international policy on peacebuilding?

The International Institutional framework for Peacebuilding

Current United Nations peace support operations have evolved from peacekeeping in its traditional sense towards what is known as multi-dimensional peace operations. In *An Agenda for Peace*, (1995), Boutros-Ghali talked of peacebuilding largely in terms of post conflict disarming of previously warring parties, the restoration of law and order, the repatriation of refugees, advisory and training support for security personnel, monitoring elections, advancing the protection of human rights and reforming or strengthening governmental institutions (Boutros-Ghali, 1995, 61). There was a growing awareness during the latter part of the 1990s, evidenced in documents such as the Windhoek Declaration (2000) and the Namibia Plan of Action (2000), of the need to include a gender aware perspective into the policies and processes of institutional peacebuilding. This culminated in October 2000 in the United Nations Security Council Resolution 1325 (UN 2000), which was the first Security Council resolution to specifically focus on issues of gender. Resolution 1325 was intended to improve the protection of women and girls during armed conflicts and to encourage the participation of women in conflict prevention, peacemaking, peacekeeping and peacebuilding. Amongst other provisions, Resolution 1325 called for increased participation of women at decision making levels and consultation with local and international women’s groups. There is clearly recognition in the preceding discussion and policy document of the significance of women’s work in peacebuilding activities at local and community level and by implication the understanding that peacebuilding at this level is crucial to the sustainability of peace. However, another substantial United Nations report on peace and security, the Brahimi Report (UN, 2000) only mentions gender once in the concluding chapter and this was linked to the conduct of UN personnel in the field. The complete disregard for a gender aware approach to peace and security in this document reveals how gender issues have traditionally been marginalised from the mainstream within the UN.

Within the discussions preceding Resolution 1325 ‘Gender balance’ refers to the equitable representation of women and men in all areas. Promoting gender balance requires explicit support of women’s participation, particularly in decision making (UN, 2003, Chapter 9). ‘Gender Mainstreaming’ refers to the process of systematically incorporating gender perspectives into areas of work and assessing the implications for women and men of any planned action, including legislation, policies or programmes (UN, 2003, Chapter 9). It is intended as a strategy for making both women’s and men’s experiences and concerns an integral part of all stages of the policy making and evaluation.

A Security Council Report of the Secretary General entitled ‘*Women and Peace and Security*’ of October 2004 reveals that despite some achievements there remain major gaps in all areas and in particular in relation to women’s participation in conflict prevention and peace processes (UN, 2004, section II, paragraph 4). The conclusions of this document were particularly damning: “in no area of peace and security work are gender perspectives systematically incorporated in planning, implementation, monitoring and reporting” and “in the areas of conflict prevention, peace negotiations and post-conflict reconstruction, women do not participate fully and more needs to be done to ensure that the promotion of gender equality is an explicit goal in the pursuit of sustainable peace” (UN, 2004, Conclusions).
What is a Gender Aware Analysis?

Essentially, to understand the necessity of a gender-aware analysis it is necessary to distinguish between the biological entities of ‘men’ and ‘women’ and the social constructions of ‘masculinity’ and ‘femininity’. Within this context it is possible to examine the deeply entrenched belief in the male ‘warrior’ and the female ‘peacemaker’ constructs. There is a duality in the way that women have been traditionally represented in contexts of war and peace. Firstly they have frequently been represented as the ‘beautiful souls’ creating the nurturing environment of hearth and home (Elshtain, 1995), and secondly, there is an overwhelming representation of women as totally passive victims of conflict and violence (Bop, 2001; Manchanda, 2001). Both representations are problematic.

The construction of differentiated gender roles is not static in either time or place and varies across cultures, races, classes, age groups and eras. The construction of these gender identities is also linked with political, social and economic development and organisation of societies. This process has almost invariably privileged masculine type attributes over feminine type attributes. Although the socially constructed ideal of masculinity does not accurately reflect the personality of most men, it serves to sustain patriarchal authority and legitimises a patriarchal political and social order. For Galtung (1996, 40) patriarchy is a clear example of what he defines as cultural violence. The cultural internalisation and acceptance of patriarchy and androcentrism serves to perpetuate the distortion of our understanding of social organisation. In many fields of knowledge we have come to regard what is masculine with what is human or normal because of the removal of the feminine from public life.

It has been claimed that war is the supreme expression of patriarchy and the warrior the ultimate symbol of masculinity (Elshtain 1995, Cooke and Woollacott, 1993). Whilst challenges have been made to the conventional view of men as war makers and women as peace makers, this notion is so deeply embedded within the structures of society and within the socialisation and internalisation of prevailing societal ‘norms’ that it remains paramount, implicitly and explicitly, in many accepted notions of ‘reality’. This can be seen in the exceptional and sensational media coverage of events in which women have been involved in perpetrating violence, for example women’s involvement in suicide bombing or the behaviour of female guards at Abu Ghrabi prison. The use of language and representation is fundamental to the maintenance of the gender divide.

Much of Foucault’s work is based on the idea that power is embedded in institutions or regimes that produce ‘knowledge’ that is claimed as truth - “the problem is not changing peoples consciousness - or what’s in their heads - but the political, economic, institutional regime of the production of truth” (Foucault, 1980, 133). This phenomenon of control of the production of ‘knowledge’, in this case about gender, is highlighted by Steans:

“Hegemonic domination is not just a question of the social relations of inequality and domination, but also about the production of knowledge and the formation of concepts and ideas that set parameters for how we think” (Steans, 1998, 129)

The historical process of state building has reinforced the public/private domain and has embedded deeply gendered constructions that establish power divisions between men and women. The emphasis on the male as ‘citizen’ and ‘political actor’ can be traced back to the classical theorists (see Machiavelli, 1961 and Hobbes 1968) and the early formation of states. The progressive development of the state worldwide has been dependent on women’s unpaid and largely
unrecognised work - reproductive, sexual, educative, social, domestic, and agricultural. This kind of labour is traditionally seen as secondary to public affairs. However, it is female labour and sexual and reproductive services that have supported and allowed the male/public/political realm to function (see Enloe, 1990, 2000). The impact of modernity and the rise of nationalism meant a relocation of social and individual identity from the local to a much larger scale. Within this move changes in gender roles have often been marginalised by essentialist calls for women and men to resume their ‘rightful’ roles and positions (Smith, 2001). Much state discourse assumes that ‘citizens’ and ‘workers’ are gender neutral whereas in reality state policies frequently affect men and women in different ways because of their different roles and experiences. State building involves telling a particular story about the nation (Anderson, 1991, Steans, 1998, Moghadam, 1994). Steans (1998 , 66) points out the conflicting qualities of nationalist discourse whereby nationalism presents itself as a modern project creating new identities, but at the same time draws on cultural values taken from an imagined past. The burden often falls upon women to be the guardians of cultural authenticity which imposes boundaries and constraints. Women have often been used as symbols for political goals or for cultural identity and authenticity during processes of state building or revolution (Moghadam, 1994). Symbols and imagery around women and their roles have been used to create a sense of national consciousness in many parts of the world, both East and West.

**Gender and Peacebuilding**

Establishing an epistemological basis for an analysis of gender and peacebuilding is problematic in that it is easy to fall into the essentialist argument that women, by their very nature, have something special to bring to the field of conflict resolution and peacebuilding and that this forms the basis of an argument for greater inclusion in peacebuilding processes. As Dan Smith points out, (2001, pp32-46) attempts to base political strategies against gender inequality on essentialist conceptions of particularistic traits of femininity fails to recognise the multiplicity of factors contributing to identity and its ultimate fluidity.

Many have argued that there is a significant connection between women and peace and men and war, and indeed this is the major storyline that is fed to us through much media representation. Considerable evidence supports this understanding as women have a longstanding connection with peace movements (see for example the history of the Women’s International League for Peace and Freedom (WILPF), or the women involvement in the Greenham Common protests). Whether there is an ‘essential nature’ of women that tends towards peacefulness has been hotly debated among feminist scholars (see Reardon, 1985, 1990, 199; Ruddick, 1989). It is easy to fall into the essentialist trap when arguing women’s association with peace; however as Smith (2001) points out political strategies arguing against gender inequality go astray if they rely on essentialist conceptions of femininity. If it is argued that women are by nature inherently different to men in terms of their relationship to violence and peace, then arguments for divisions of labour hold some credibility: ‘Mobilising women as women, however progressive the cause, must run a strong risk of playing directly into the hands of gender traditionalism, especially if traditional gender concepts are the basis of the mobilising appeal’ (Smith, 2001, 44). Such qualities as understanding, cooperation, dialogue and a sense of fairness are often invoked as those traits necessary for the optimum rearing of future generations and these qualities are seen by some as transferable to the international system in terms of a more humanistic approach to peace building and conflict resolution.
Ruddick (1989) argues that gender stereotypes are myths which prevail even in the face of historical and empirical contradictions. Skjelsbaek’s (2001) analysis of the construction of femininity in the wars of former Yugoslavia, El Salvador and Vietnam reveal significant differences in women’s reaction to war. She reveals that with the best interests of their children at heart, the different groups of women reached different conclusions regarding their rejection of war or their support of war and participation in combat. Whereas the Bosnian and Croatian women tried to remove their children from the conflict, the El Salvadorian and Vietnamese women saw support and involvement in the conflict as essential to their children’s future interests.

It is contended here that beliefs and myths about gender play a crucial part in maintaining violent and war-like social structures and thus an understanding of the processes that create and maintain these gender constructions is essential to any study attempting to understand peace. Masculinity is often constructed as something that can be reinforced through violence, and that many of the violent acts perpetrated in armed conflicts are the result of men being convinced that this is the way to show their masculinity - ‘in such a social context, mobilising people for reconciliation may be impossible as long as the dynamics of the male-female division of labour are ignored’ (Skjelsbaek and Smith, 2001, 3). Beliefs about what it means to be a real man or a real woman are instilled and reinforced through a variety of images, symbols, expectations and behaviours in everyday life. In the field of IR the type of language used to describe strong dominant states and practices is closely identified with traits attributed to masculinity, whilst that used in a derogatory way to describe weakness is frequently associated with femininity (See Zalewski, 1996, Cooke and Woolacott, 1993).

Conclusion
The approach of this work is not to problematise women and their inclusion into the field of peacebuilding, but rather to problematise the notions of masculinity and femininity and to deconstruct the traditional hegemonic structures that surround these concepts and that have almost always privileged masculinity in the state and international context. Whereas it is widely, but not unanimously, acknowledged that ‘gender balance’ and ‘gender mainstreaming’ is a positive goal, the policy proposals often fail to be based on an ontological understanding of the socially embedded nature of gender divisions of power. Women’s right to participate at all levels of decision making is a basic human right and a responsibility as a citizen, but the argument for a gendered perspective of peacebuilding is more complex than a call for women’s human rights. Although the human rights argument for equitable participation is extremely valuable in itself, it is made more complex with contemplation of the issue of legitimising the status quo, the spectre of victimhood as a basis for inclusion and the debate about whether gender equity and balance will actually make a qualitative difference to policy principles, policy making and policy implementation. The assertion is that an adjustment to conflict resolution procedures to incorporate more women for the sake of ‘political correctness’ fails to acknowledge that the gendered nature of social reality is fundamentally incorporated into the complex web of conflict causation, and as such must be acknowledged and understood to be an equally crucial part of conflict resolution procedures. Thus, simply adding more women to existing structures that remain ontologically and epistemologically unexamined for biases and power interests is clearly insufficient to transform conflict resolution procedures into the more emancipatory processes called for by those within the positive peacebuilding school of thought (Galtung, 1996, Lederach, 1997). Chenoy and Vanaik (2001, p123) have suggested that “the politics of feminisation is an
integral part of a wider politics of democratisation and empowerment which has now decisively altered domestic and international politics alike”. But what is meant by ‘the politics of feminisation’? It cannot be defined as merely having greater representation for women but must refer to the transforming of political projects by injecting a different set of values, those traditionally associated with femininity (not necessarily with biological womanhood). For a meaningful gender aware perspective of peacebuilding, it is not enough to simply counterbalance the existing power politics with a slightly more representative approach to mitigate its worst excesses but it is necessary to initiate new models and frameworks based on a reflective, critical re-evaluation of mainstream theory and practice.

Bibliography


Reardon, B - (1985) - Sexism and the War System. 1985, New York, Syracuse University Press.


A REALISTIC APPROACH IN EVALUATING URBAN REGENERATION PROGRAMMES

Osbourne Glover

Abstract

This paper presents the incorporation of realistic practice into the formulation of evaluation methodology of activities and projects relating to Devonport’s New Deal for Communities (NDC) Programme. Given that urban policy and neighbourhood residents are closely connected with the affects of social and economic regeneration, and that realistic evaluation (Pawson and Tilley, 1997) is grounded in scientific realist philosophy, provides for new ways in evaluating programmes and projects. Realistic evaluation focuses on testing theories of how NDC programme outcomes of inclusion, participation and provision for black and minority ethnic (B&ME) residents are produced by specific mechanisms and contexts. Such a model provides for a means of discovering and understanding which regeneration program comparisons relating to context-mechanism-outcome configurations are efficacious thus providing policy makers with theoretically based typologies of effective context-mechanism-outcome combinations.

Introduction

The need to integrate knowledge with action is nowhere greater than in the area of urban regeneration particularly given local social and economic conditions, change and innovation in which programmes must be planned, monitored and evaluated in order to achieve ‘best fit’ which meets the needs of the many diverse and different groups within our neighbourhoods, and draws on the achievement of good practice lessons for the benefit of future activities. It is within these linkages and there interrelation with one another and coupled with social and economic policy issues in general greatly add to scientific analyses and policy responses required to address regeneration programme concerns adequately.

In defining urban regeneration programme tasks they should have an overriding aim of improving the well-being, quality of life and physical infrastructure for people in local areas. Specific objectives may vary between different programmes, depending on local conditions, problems affecting each area, and the general priorities set by each regeneration partnership. Thus, a key report by the Social Exclusion Unit (SEU), Bringing Britain Together, argues that regeneration programmes have simply failed to work citing four reasons:

- lack of a holistic and joined-up approach
- Insufficiently focussed and lacking an intensive strategy
- Resources spread too thinly
- Failure to engage local people and communities in a meaningful way (Cabinet Office, 1998:63).
Thus, rather than key decision bodies, local government departments and local people working together on issues relating to inclusion and participation have tended to be given to one interest group or prime agency as opposed to all community interests addressing issues holistically. This new approach by Labour suggests that joined-up problems require joined-up solutions, which in turn requires all partners, agencies and residents working together to achieved a shared vision (Perri 6 et al, 1999).

A further concern relates to community participation, an issue which features prominently in New Labour's criticism of past regeneration policy failure. New Labour suggest that active citizenship is a key in ensuring success in the sustainability of its regeneration programmes. However, Geddes and Benington's (1995) study of a City Challenge Programme in Tyneside, found that 'community involvement in formal decision-making processes remains rather superficial' and that 'partnerships achieved limited success in involving community interests as equal partners’ (1995:52). Like its predecessor, New Labour's approach to community representation has sometimes served simply as a means of legitimising decisions that have been taken by, and worked in the interests of, local authorities, private sector and key players nominated by these partners. Thus, in the same way that past regeneration policies have sustained on-going criticism, Labour's new approach to regeneration was the introduction of a New Deal for Communities flagship model, which stated:

New Deal for communities (NDC) is a key programme in the Government’s strategy to tackle multiple deprivation in the most deprived neighbourhoods in the country, giving some of our poorest communities the resources to tackle their problems in an intensive and co-ordinated way. The aim is to bridge the gap between these neighbourhoods and the rest of England ... tackling five key themes of; poor job prospects; high levels of crime; educational under-achievement; poor health; and problems with housing and the physical environment ... 17 pathfinders were announced in 1998, followed by a second round of 22 partnerships in 1999. Approximately £2bn has been committed to the 39 partnerships (DETR, 1998:1).

One of the main features of NDC is that it must involve and engage with all people in local communities affected by the programme. Failure to engage with local people can lead to an areas funding being rejected, or having funding withheld until participation is achieved. Thus the local focus should be clearly on how the community better identifies its problems and encourages local ownership in order to achieve long term sustainability.

As earlier highlighted New Labour's NDC programme embodies a new approach to local regeneration that redresses some of the many failings of previous Area Based Initiatives (ABI's). As such, New Labour has placed great emphasis on the value of 'community'. Indeed, the notion of community has been positively promoted by the Government, none more so than its leader Tony Blair, whose roots lie in the communitarian philosophy of Macmurray (in Driver and Martell, 1998:27-9) who believed that all individuals are defined through their attachments to other people, within families and within communities. Thus, Blair believes that society is stronger when working together as a team, rather than a collection of individual players.

In this sense, we can clearly identify New Labour’s vision of community in which all its residents would have a sense of belonging, and would come together to work
towards a shared collective vision in order to shape the future and sustainability of its community.

However, the following section of this paper discusses how such an approach has affected and impeded the inclusion, participation and provision for disparate and fragmented black and minority ethnic people who live in low areas of settlement which might include them as residents within an NDC regeneration area. Similarly, in terms of realistic evaluation of a local NDC programme which might be based on the idea that programs deal with real problems, whilst ensuring that its primary intention is to propose realistic developments in urban policy making that benefit all programme stakeholders. Thus a programme such as NDC might not be initially conceived as a ‘working’ one rather it is the actions of local stakeholders that will make it achieve success. Indeed, the casual potential of any initiative must take the form of providing resources and reasons to enable NDC programmes and its stakeholders to make them work. Thus, an event or action may generate a change if only it is in the right condition in the right circumstances for example, the embodiment of knowledge that identifies what works for whom in what circumstances (Pawson and Tilley, 1997).

**Solving the Problem**

Past research on B&ME participation in regeneration programmes (Beazley and Loftman (1998); Cheliah, (1995); Lawless, (2004); Loftman and Beazley, (2001); Oc et al, (1997) and Stewart and Whitting (1983) on behalf of government departments and information units have commented that at best the City Challenge Programme, the Single Regeneration Budget (SRB) Challenge Fund which followed the Urban Programme have not really benefited B&ME groups as most funds went to larger more established partnership organisations. Indeed a Home Office Report which set out a code of good practice in relation to Compact and its relationship with B&ME groups and the wider voluntary and community sector that:

To date, funding for the BME voluntary and community sector, whether from existing or new sources has been significantly below that of similar organisations in the mainstream voluntary and community sector. This is all the more detrimental given that the BME sector does not generally have the windfall legacies, income streams and leverage opportunities that the more established mainstream sector attracts (Home Office, 2001).

Indeed, such issues relating to resourcing B&ME groups is not only relevant to medium to high areas of minority ethnic settlement, as within Devonport NDC programme a particular B&ME group were turned down in an application totalling £1.2 million (out of the £49.3 million allocated to the Devonport area) in which to capacity build B&ME groups within the area over the ten year term of the programme. It was felt that the reason cited for turning down the application were discriminatory doubts about this particular black led organisations capacity to manage such a large sum of money, coupled with assumptions that as a black organisation you are written off before you start with responses such as: ‘someone will run off with the money’, and that ‘They won’t deliver’. Such discrepancies highlight the need for decision-making structures to provide positive redistribution mechanisms within grant allocations in order to redress such imbalances. It is within this framework this paper will briefly explore the Devonport NDC programme, an area described as having a small B&ME population who find it difficult to navigate the invisible barriers placed between them and the significant white majority population.
The area of Devonport a ward within the city of Plymouth has a resident population recorded at just over 4,853 (ONS, 2003) may not appear to be an obvious and immediate example of urban settlement for B&ME people as most research and studies relating to urban regeneration have usually been carried out in larger cities such as London, Birmingham, Manchester and Leeds which have significant conurbations of B&ME communities. However, as this article will show, that although the numbers of B&ME people may be small in comparison to larger cities, the example using Devonport where B&ME people are less visible, offers some useful insights into thinking about the local experiences of how B&ME groups might fit into realistic evaluation of this particular NDC programme.

Due to a lack of research undertaken at citywide and neighbourhood levels relating to the demography of B&ME populations within the city of Plymouth, the B&ME population and its many diverse groups remain unknown. Indeed the neighbourhood Census records for Devonport recorded in 2001 reveals a local B&ME population numbering 144. However, since the Census the National Asylum Seeker Service (NASS) dispersed a number of asylum seekers to the Devonport area, which no doubt increased the number of B&ME residents within the area. Due to changes in government legislation relating to housing asylum seekers and the decanting of some Devonport residents in order to improve local housing stock, a number of people including those from B&ME communities have relocated to other areas of the city. Hence, since the start of this particular piece of research in late 2004 the number and ethnicity of B&ME residents including black mixed heritage children who live within white families has been estimated to be in the region of about 100 people.

Given that the NDC programme insists that all local residents are to be involved in all activities across the board; ‘NDC partnerships and programmes are being driven by their communities, and residents are fully involved in the planning and delivery of NDC programmes’ (DETR, 2001:11). Efforts are being made to ensure that traditionally marginalised groups such as those of B&ME backgrounds are to be fully included in the process through the introduction of the first ever Race Equality Guidance for New Deal for Communities. This guidance has emphasised the need for NDC partnerships to fully include B&ME residents in all aspects of the NDC programme (DETR, 2000). As such, the DETR (now Office of the Deputy Prime Minister - ODPM) makes clear reference for NDC Partnerships to recognise that:

In neighbourhoods where Black and ethnic minority groups make up a relatively small part of the population, the need for race equality measures is possibly even more crucial. In areas with small black and ethnic minority populations, people tend to be more isolated, marginalised and threatened by racial harassment (DETR, 2000:13).

In this respect, and given the small number of B&ME people migrating in and out of the Devonport area over the years, several B&ME residents have commented that because they are so low in numbers increased their likelihood of racism and social isolation. As such, this area of research posits that detachment and social isolation as identified in low areas of minority ethnic settlement suggests that long-term health related psychological problems might arise as residents lack links to key institutional agencies and access to local service providers. Moreover, without support they run the risk of becoming wider estranged from local and citywide social networks. Lacking ‘a sense of place’ in an already deprived community without the support of a wider B&ME population and personal networks very often
leads to affecting their quality of life by reducing levels of confidence, self-esteem, and a sense of belonging.

Indeed, it would also be wrong for this article to suggest that such issues only affect B&ME groups. Clearly, what is evident from current research is that people from the wider community have commented that Devonport NDC has failed disabled people, aged, young people and other minority groups. Thus Devonport NDC’s vision statement which states; ‘Our vision is to create a thriving, vibrant community that raises aspirations, grasps opportunities and which has people queuing to join’, (Devonport NDC Delivery Plan 2001-2011) might well be the litmus paper test in which Devonport will have to show the processes of success or failure in relation to what works for whom and in what circumstances (Pawson and Tilley, 1997).

Improving Opportunities

Since the introduction of New Labour’s New Deal for Communities initiative, a number of evaluation methods have been used in dealing with some of the initial teething problems associated with this new regeneration programme. For example, Roger Tym and Partners (RTP) and the University of the West of England (UWE) were commissioned in 1998 by the then Department of the Environment, Transport and the Regions (DETR now the ODPM) to produce a good practice guide on local regeneration evaluation.

Within the guide RTP and UWE suggest that local evaluation forms an essential part of local area regeneration in its ability to assess results, compare achievements with costs, highlights how programmes can be improved, and draws on good practice lessons in which to benefit future regeneration activities. Coupled with the need to evaluate local regeneration schemes should in effect, also encourage the evaluation of all public policy. However, the DETR (1998) asserts that the need for evaluation is greatest in urban regeneration areas which are dominated by major change and innovation.

The Centre for Regional Economic and Social Research at Sheffield Hallam University (Lawless, 2004) have also been co-ordinating the national consortium responsible for evaluating the NDC programmes (2002-2005) on behalf of the Neighbourhood Renewal Unit (NRU). The evaluation of this major regeneration programme is intended to assist all 39 NDC partnerships deliver on their action plans by undertaking a range of tasks which consisted of household questionnaires designed to establish the ‘base-line’ position within NDC partnerships. Secondly, evaluation involved the collation and analysis related to changes in worklessness, welfare claimants, post 16 education acceptance into higher education, indicators relating to health, mortality and illness and specific crime figures. Interestingly, the themes of crime, health, education, housing and employment fit within the government’s national floor targets which forms the major part of its Neighbourhood Renewal Strategy. Findings from this initial evaluation indicates that very little has been achieved in meeting the needs of B&ME communities within NDC Partnerships.

The third task involved reports produced by 39 NDC partnerships which highlight process issues relating to inter-agency partnership working, barriers to delivery, the involvement of B&ME communities, and wider community engagement (Lawless, 2004).

It is worth noting that this report also highlights how local initiatives might fail if they are not built on genuine partnership arrangements which should include and
involve local residents and others. Moreover, the DETR (2000) and ODPM (2002) suggest that NDC’s have to make a specific effort to engage with B&ME communities who have been a traditionally marginalised sector. Thus although race equality is being mainstreamed across neighbourhood renewal as a whole, further suggests that work undertaken on the national evaluation of NDC programmes seeks to ensure three main issues relating to community engagement. These are:

- Perceptions as to progress made in the area
- Barriers and constraints
- Involvement of B&ME communities in the programme
  (adapted from Lawless, 2004).

Further findings from this NDC evaluation report suggest that in its review of the evidence base for local area regeneration found that much of the existing evidence in relation to numbers of people who might live in deprived areas have cast doubts on the ability of statutory service providers involved in area processes which they see as intrinsically complex. Statutory bodies also cited difficulties with the odd area boundaries, peculiar governance structures and contrasting agendas, which result in their reluctance to invest in interventions which might, or might not work given the wide range of social and economic factors present within deprived neighbourhoods.

As such, sustainability and partnership working remain the major factors in achieving successful local area based regeneration. There are of course many other varying effects, which might impact on breaking the cycles of disadvantage for residents within communities in order to integrate social, economic and physical development which produce beneficial outcomes to all its residents. Area programmes also require long term solutions, which achieve inclusive processes of management, governance and democratic participation, which are all key factors in achieving long term sustainability.

On the other hand, in less stable communities experiencing low levels of educational achievement, employment, high crime and competing agendas such as those experienced by the Devonport NDC programme have triggered anger and an on-going tension on the part of various groups feeling that resident NDC Board members are ‘lining their own pockets’ for the benefit of their own groups rather than looking at the wider picture of ‘who needs what’, and how would ‘it work’ if we looked at issues differently, in other words variance in context.

Such levels of effects and tensions are experienced as earlier highlighted by all disadvantaged groups and not necessarily those from B&ME backgrounds, who might experience different circumstances which trigger different responses, thus producing different outcomes. Effectiveness in such measures is contingent on the context in which it is introduced. Thus according to Pawson and Tilley (1997) what works to produce an effect in one circumstance will not produce it in another.

Tilley, (2004:4) further asserts that mixed findings from his own study on mandatory arrest for domestic violence evaluation studies are typical, and that findings of results are enormously varied. Tilley (2004) cites that this is not helpful for policy makers or practitioners, which is why Ray Pawson and himself developed ideas about realistic evaluation - the ‘what works for whom in what circumstances’.
The counter-proposition of the realist evaluation model developed by Pawson and Tilley is that data construction should be theory-driven, so that theory, expressed as context-mechanism-outcome (CMO) configurations\(^{38}\) is the subject matter of the inquiry, and the subject (stakeholders) is there to find ways of identifying, articulating, testing and refining conjectural CMO configurations.

According to Pawson and Tilley (1997:115) most evaluation studies seem to be one-off affairs, which neither look back and build on previous findings, neither do they look forward to future evaluation. Thus evaluators are often charged with just pecking at the edges of most initiatives rather than feeding into improvements in policy and practice, in essence to make improvements which offer value for money. Thus in deepening our understanding of programme CMO patterns we need to configure and conceptualise their nature and structure in order that they can be explicitly defined.

Thus the researcher would need to indicate the regeneration programme mechanisms (local regeneration of the area to increase educational advancement, reduce crime and so on) together with the context (A medium-sized socially homogenous community, clearly defined deprived neighbourhood) within it works better and identifies possible outcomes, which might be intended or unintended. The focussing on this CMO configuration institutes a process in which, and through, an analytical framework can be designed in which to interpret similarities and differences between activities within the programme.

Furthermore, as realist evaluators seek to understand the ‘for whom and in what circumstances’ a programme should work its way through contextual conditioning. In this case context refers to the social conditions together with the rules, norms, values and interrelationships which would explain the success or failure of a programme in a neighbourhood undergoing regeneration. While on the other hand, a mechanism identifies the process such as ‘empowerment’ or ‘capacity building’ which might theoretically intervene between the differing programmes, initiatives and outcomes. Moreover, as realistic evaluation seeks to understand the ‘why’ programmes work through understanding the action of mechanisms, outcomes specify program goals together with any unintended consequences that they may produce. As such programmes fire multiple mechanisms which cause different effects on different contexts thus producing multiple outcomes. It is through realist evaluation that we are able to examine and analyse multiple outcomes not simply in order to find out if programmes work, but to uncover if the conjectured mechanism/context theories are confirmed (Pawson and Tilley, 1997:54-82).

**Conceptualising the Context, Mechanism, Outcome Cycle**

The versatility of realist research design in evaluating regeneration programmes whether by the use of ethnography to experimental manipulation, or from the use of hard official data, street corner observation or informal interviews with residents and stakeholders may use the three guiding themes within a realist research strategy. These three themes are:

1. They increase specificity of our understanding of the mechanisms through which a programme accomplishes change.

\(^{38}\) A context-mechanism-outcome (CMO) configuration, is a proposition stating what it is about a programme which works for whom in what circumstances. The conjectured CMO configuration is the starting point for an evaluation, and the refined CMO configuration is the finding of an evaluation.
2. They increase specificity of our understanding of the contextual conditions necessary for triggering program mechanisms.

3. They increase specificity of outcome pattern predictions according to context and mechanism triggered. (Pawson and Tilley, 1997:114).

Given that policy makers know a great deal about NDC regeneration programmes which they are responsible for, the policy maker will also have some notions about how the programme will generate positive benefits for neighbourhoods. This according to Pawson and Tilley (1997) is the policy maker’s program theory.

Thus assuming that NDC regeneration programmes such as Devonport NDC involve multiple mechanisms, they may also include multiple contexts. The field of evaluation relating to urban regeneration programmes suggests within the design and analysis we try to identify situations such as residents within the area who through an activity or project might benefit. Thus, implementing an intervention programme clearly involves residents who are exposed or receive the benefits and effects of the project.

The mechanisms within a regeneration programme may determine whether it works, or what it is about this particular programme that makes it work. A mechanism, further demonstrates how outputs might follow from the stakeholders’ choices and how these are put into practice. Thus regeneration programmes work through the actions of mechanisms, through a combination of bringing together reasoning and resources. If mechanisms were not taken into account within the evaluation process, outcome activities would be unexplainable.

As NDC regeneration programmes were created with a purpose of providing sustainable services through problem solving. It is these purposes which are the goals or the intended outcomes usually used by stakeholders to guide their activities and to determine resource allocation. Thus outcomes, whether those that might have been instituted within a community to serve local government political purposes, they may also concern unintended outcomes such as key agencies ignoring the needs of minority groups. It is within this social dimension that the realist may conceive a programmes success on the action of key stakeholders activities in actively engaging with, and providing provision for B&ME groups. Such incentives, reasoning and allocation of resources enable participants in this case B&ME groups to better accept and cooperate in the implementation of further activities within the programme.

**Conclusion**

Realistic evaluation in relation to NDC regeneration programmes attempts to achieve evaluation through a process of programme realisation. The basic task of evaluation research is one of context-mechanism-outcome configurations which focuses on learning in greater detail ‘what works for whom in what circumstances’ (Pawson and Tilley, 1997). Thus in terms of best practice realistic evaluators evaluating regeneration programmes would need to understand why mechanisms activated by the programme generate outcomes in specifiable outcomes. Secondly, According to Pawson and Tilley (1997) realistic evaluation recognises that the effects of government regeneration policy objectives are often achieved through its potential in intervening in right situations. A realistic evaluator’s role is therefore to conduct evaluation which improves programme and policy objectives.
In this way, realistic evaluators would be able to apply appropriate methodologies to construct necessary data, which is a key aspect in creating theoretical structures under investigation. Key stakeholders are then able to make informed contributions to them. The use of interview techniques, and qualitative methods provide a strategy which is highly specific enabling opportunities to elicit or interpret findings of evaluation in realistic terms. Finally, such a procedure enables policy makers an opportunity to better understand the process from the realistic perspective of the evaluator and key stakeholders responsible for planning and delivering local regeneration programmes. Thus it is not just about what evaluation can teach us, but what use we make of it in order to achieve long-term sustainable regeneration (the why it works) which positively values and recognises a social acceptance model of difference within communities which very often plays a significant role in the selection and decision making processes of who gets what and why within current regeneration programmes.

Bibliography


EVIDENCE IN SOCIAL POLICY: CAN WE SEE PAST THE HIERARCHY OF EVIDENCE?

Mark Pearson

Abstract

The use of the ‘hierarchy of evidence’ from Evidence-Based Medicine (EBM) in social policy has met with harsh and robust critique. It has been argued that methods which synthesize (predominantly quantitative) knowledge are inadequate for complex social interventions where context and the nature of implementation itself are of prime importance (Pawson 2002; Eraut 2004; Hammersley 2004; Victora, Habicht et al. 2004). This paper acknowledges the importance of such critiques in challenging the more mechanistic attempts at moving towards an ‘evidence-based society’. However, it also cautions against an unthinking resistance to ‘Evidence-Based Policy & Practice’ (EBP) where the research synthesis that takes place under this banner has started to evolve in ways that endeavour to address these concerns.

The paper begins with a brief history of EBM and an outline of the social and political context in which it arose. The rationale for, and contentions over, the hierarchy of evidence will then be considered. Much of the critique of the ‘EBM approach’ to evidence has focused on the lowly status of qualitative research within this hierarchy. It will be argued that, whether the favoured methods are quantitative or qualitative, methodological arguments focusing on the superiority of particular approaches are counter-productive.

The paper concludes with an analysis of the contention over EBP between two leading participants in the field; Martyn Hammersley and Iain Chalmers (Chalmers 2005; Hammersley 2005). Using these two papers as an example, it will be argued that it is necessary for social policy researchers to be very precise in their critiques of EBP if they are to engage productively with debate in the field.

A Brief History of Evidence-Based Medicine.

EBM, as detailed in the seminal Evidence-Based Medicine Working Group paper (1992), proposed to offer a method for reducing doctors’ reliance upon “…intuition, unsystematic clinical expertise and pathophysiologic rationale…” (p.2420). This paper, largely concerned with reporting developments in the medical curriculum at McMaster University (Ontario, Canada), strongly promoted the randomised, controlled trial (RCT) as the research method with the highest validity. In RCTs, the random allocation of research subjects to intervention or control groups is designed to distribute all possible incidences of bias randomly between the two groups (Oakley 2000; Reynolds 2000). Assuming a minimum sample size and no violation of assumptions, statistical inference can be performed.
that allows a statement to be made regarding the likelihood of the effects of the intervention having occurred by chance. Moreover, these results are then generalisable to the population from which the sample was drawn. It is contended that this approach, being both systematic and transparent, is the strongest way to establish evidence upon which rational policy and practice can be based (Gray 1997; Sackett, Richardson et al. 1997). Thus, generally accepted but potentially harmful practices such as bed rest can be successfully challenged (Allen, Glasziou et al. 1999), and debates about the effectiveness of treatments, such as the administration of human albumin, can be resolved (Wilkes and Navickis 2001).

The Socio-Political Context in which Evidence-Based Policy & Practice Emerged.

Despite the protestations of some of the stronger advocates of EBM and EBP, evidence-bases have not arisen in a socio-political vacuum. The macro-political context in the UK has seen a significant shift in the past quarter century from an interventionist Keynesian economic approach to a monetarist model which pursues deregulation, privatization and reductions in taxes (Gray 1993). The rise to power of New Labour in 1997 arguably intensified this Neoliberal approach, albeit one which was tempered with a concern for social inequalities.

This emphasis upon the Neoliberal nature of policy at the end of the twentieth century is not intended to suggest that certain conclusions will necessarily follow. Arguably, the manner in which neoliberal policies are the result of an ongoing political and historically-contingent process means that the outcome for social policy is by no means already decided. Taking such a position enables an analysis of the use of evidence in social policy that is sensitive both to the phenomenon's ideological nature and the specific national context out of which it has grown (Peck and Tickell 2002). As such, the analysis presented here endeavours to tread a line between both the ‘end of history’ narratives positing neoliberalism as the natural (and best) result of human societal development (inter alia Green 1987; Fukuyama 1992; Minford 1998) and structuralist analyses which view power as an oppressive, rather than productive, force. This position is based upon analyses that view neoliberalism as a phenomenon to be ‘worked’ rather than resisted per se (Larner 2000; Peck and Tickell 2002; Larner 2003).

One of the key manifestations of neoliberal policy was the institutionalisation of a managerialist agenda. In the UK, New Labour continued the work done by the Conservative government on the ‘New Public Management’ (see Hood 1991), although now under the guise of ‘Modernization’. Managerialism was narrativised as providing a means of overcoming immutable and ‘rule-bound’ bureaucracy, paternalistic and opaque professional practice, and ideological interference by politicians (Clarke and Newman 1997). It was thus a thoroughly ‘modern’ institution. It was contended that a ‘modern’ management approach would provide a non-partisan means of assessing options and their costs as they would affect the organisation (Clarke and Newman 1997), provide a transparent means of professional regulation (Osborne, Bovaird et al. 1995), and would enable each defined unit to be held to account through the assessment of outputs (Boyne 1999).

The managerialist approach to evaluation required outputs to be explicitly measurable (Broadbent and Laughlin 2002). In EBP, this takes the form of an audit trail whereby a professional can justify their actions through reference to the ‘best practice’ identified in a systematic review. However, doing so risks belittling aspects of professional practice, in particular those which substantially depend upon tacit knowledge. These areas are typically not amenable to straightforward (usually quantified) measurement (Broadbent 1998). This approach also has a constitutive role; the visibilities created by the audit system substantively shape
public service policy in the future (Power 1997; Broadbent and Laughlin 2002). Areas of practice not visible through the auditing lens cease to exist for the purposes of decision-making under managerialist rationale; the hue of EBP as it is practiced now will play a key role in how professional practice and its regulation takes place in the future.

The Utilisation of EBM’s Hierarchy of Evidence in Social Policy.

The hierarchy of evidence involves a methodological ranking of research. It is founded upon the notion that certain ways of conducting research are more rigorous than others, and hence have a superior claim to producing knowledge that is valid as evidence. Whilst it is not claimed that every research question can be answered by using the methods towards the top of the hierarchy, it is an explicit aim that every effort should be made to utilise these methods whenever possible. Within EBM, the accepted hierarchy of evidence focuses upon differentiating between varying implementations of (systematic reviews of) randomised, controlled trials (RCTs), cohort studies, and case-control studies (see http://www.cebm.net/levels_of_evidence.asp). Research that is not amenable to synthesis in a statistical sense is ranked at the bottom of the hierarchy, along with ‘expert opinion’.

The core tenets of this hierarchy were promoted by some leading figures outside of EBM as the substantive means by which the perennial problems of social research could be addressed. The opacity of much social research, where disentangling the research findings from the values of the researcher are contended to be highly problematic, and where claims of ‘naturalistic’ observation could be argued to be highly biased, are at the root of the desire to utilise methods that were claimed to be more transparent and rigorous (Oakley 2000). There was a belief, which resonated strongly with New Labour policy, that “… whether something works is the primary question… how something works [theories, models]... is secondary.” (Oakley 2000, p.310). The argument for the primacy of the experimental approach in social research was thus advanced, and RCTs were contended to be the most rigorous way of performing such research.

The vigour with which RCTs were promoted to the social research community tended towards presenting its tenets as universally accepted within Medicine. However, EBM was subject to an ongoing debate regarding its utility for informing professional practice. In what way did the synthesis of large-scale datasets provide the detail needed for clinical decision-making at the level of the individual patient (Tannenbaum 1993; Feinstein and Horwitz 1997)? Upon what basis was it assumed that numerical data could meaningfully inform practice when complex moral decisions needed to be made (Polychronis, Miles et al. 1996; Goodman 1999)? Did the techniques of meta-analysis (used to synthesise datasets) adequately address the issues of study heterogeneity and the potential for synergistic interactions within studies (Polychronis, Miles et al. 1996)? It was perhaps no surprise that transferring the hierarchy of evidence to the field of social research resulted in significant contention over its appropriateness.

In some respects, the debates over the hierarchy of evidence in the field of social research echoed those that had occurred in Medicine. It has been contended that the lack of contextual detail in RCTs severely restricts the ability of professionals and policy-makers to judiciously apply research findings (Pawson and Tilley 1997; Greenhalgh, Robert et al. 2004). In addition, the assumption that research findings will in and of themselves inform practitioners and policy-makers of the best course of action has been vigorously contested. The notion that a substantive element of
professional judgement is not required is dismissed as untenable in a number of fields (Tilley 2001; Hammersley 2004; Victora, Habicht et al. 2004, in the fields of criminal justice, education, and public health, respectively). Similarly, Simons (2004) emphasizes the importance of professionals’ ‘integrated knowing’, the idea that decision-making in practice has an essential ‘tacit’ component that is not easily elucidated (see Polanyi 1967; Schon 1983).

There are other areas where contentions in the social sciences over the hierarchy differ from those in Medicine. These reflect the posited nature of the more open systems in social research than those that are (perceived) to exist in medical research. The utility of RCTs in providing meaningful knowledge in the field of Public Health has been questioned by Victora, Habicht et al (2004) on the grounds of the near impossibility of controlling confounding variables. In a similar vein, Tilley (2001) questions how RCTs can rigorously control variables that interact in dynamic social situations; for example, in a crime prevention programme, interventions may develop in response to the activities of offenders, and offenders in their turn adapt to these interventions. The reality of the situations in which professionals and policy-makers have to make decisions, namely where information is imperfect and numerous competing demands are made upon the decision-maker, militates against any straightforward utilisation of ‘the best evidence’ (Pawson and Tilley 1997; Eraut 2004; Hammersley 2004).

Further critiques of the use of RCTs in social research exist. In the field of special education, Gallagher (2004) makes a comprehensive case for there being inherent problems in utilising RCTs. Inter alia, it is contended that it is impossible to ensure that a programme is delivered in exactly the manner in which the researchers intended. Moreover, in Gallagher’s example, the notion that a teacher would not retain some autonomy in their pedagogy (for example, tweaking the programme in response to individual students) is untenable, as it runs counter to the professional culture of which teaching is a part. Gallagher also highlights the problem of ‘implementation fidelity’; how is it known whether it is the programme that is causally related to the outcome, or the particular skills of the teacher concerned?

‘Debates’ About Evidence-Based Policy & Practice.

In the previous section, I have outlined perspectives from a range of professions in which the implementation of EBP, in a form satisfactory to the managerialist agenda, have been contested. I do not want to dwell any longer on the specifics of this debate; instead, I will make an analysis of a recent exchange of papers by Martyn Hammersley (Hammersley 2005) and Iain Chalmers (Chalmers 2005) in the journal Evidence & Policy. As an illustration of the way in which debates over EBP can become polarised and stray quite far from the ideal of academic debate, these two papers are very instructive.

The journal Evidence & Policy is edited by a team based at the UK Centre for Evidence-Based Policy & Practice, an Economic & Social Research Council funded Centre based (at the time) at Queen Mary & Westfield, University of London. The journal aims to provide a forum for substantive discussion regarding EBP and to promote methodological progress in the field. However, these two papers provide a salutary example of the gulf that can exist between those who adopt different positions in the debate regarding EBP. I have chosen these papers for close analysis not because I wish to fuel combative positions within EBP, but rather to illustrate the need for extreme care over terminology and the characterisation of practice lest one’s position be negatively misinterpreted. If those of us involved in the debate regarding EBP (albeit at a somewhat more junior level than Hammersley
and Chalmers) can learn from the analysis of these two papers how carefully we need to tread in the debate, perhaps genuine methodological advancement in EBP can be achieved.

Hammersley’s paper, although written as a critique of how EBP as a whole is currently practiced, largely draws upon a paper published by Chalmers in the Annals of the American Academy of Political and Social Science (Chalmers 2003). Chalmers’ paper argued the case for systematic reviews contributing substantively to a more ‘evidence-based’ society. Chalmers’ conception of rigorous systematic reviews accords with that utilised in EBM (see http://www.cebm.net/levels_of_evidence.asp) where RCTs are the highest form of evidence. Hammersley’s central concern was that Chalmers placed too much emphasis upon research evidence being able to directly inform practice. For example, if a systematic review had shown a certain intervention in the classroom to improve learning outcomes, Hammersley would defend the professional who provided plausible contextual rationale for not utilising it in a particular case, or who suggested how it may work if implemented with certain alterations. However, Chalmers (in Hammersley’s characterisation) would point to the bias and relatively narrow experience of the professional in question, and require him or her to bow to the wide breadth of evidence in the studies synthesized in the systematic review. Hammersley extensively details his reasons for not automatically preferring ‘research evidence’ to the judgement of professionals (Hammersley 2005, p.89-92); space precludes detailing them here, but a very brief summary would highlight the importance that Hammersley accords to context, the limits to generalising research findings, and the dangers inherent in assuming that certain methodologies will always provide superior evidence to others.

I cannot hide my own values in attempting to assess Hammersley’s paper, for his arguments resonate strongly with the position I would take regarding EBP. My position has been informed both by my doctoral study and prior professional experience. Nevertheless, in attempting an objective assessment, I would still argue that Hammersley’s paper is closely argued and cites a rich source of literature. On one occasion (p.86), Hammersley does use a conception of positivism that might be argued to be an unfair characterisation which inadequately represents the scientific practice that takes place under this banner, but taken in the context of the paper as a whole it would seem a minor indiscretion. It might also be possible, for one who was highly acquainted with the very latest developments in EBP (and without the hurdle of publication time-lags) to criticise Hammersley for not being completely up-to-date. This is again a relatively minor point, and one which any author with experience of the protracted nature of publishing in a peer-reviewed journal could sympathise with. In summary, Hammersley’s paper was arguably a rigorous contribution to the debate regarding EBP; moreover, it sought to cultivate dialogue through also detailing (p.87) the areas of EBP in which he was broadly in agreement with Chalmers.

Chalmers’ reply to Hammersley’s paper (Chalmers 2005) is a spirited defence of the importance of policy and practice being based upon the best research evidence. The tone of the paper is set by its title: ‘If evidence-informed policy works in practice, does it matter if it doesn’t work in theory?’ - I would only be exaggerating slightly if I characterise Chalmers’ position as “Hammersley is an academic meddling in matters which he does not understand; he should go back to his ivory tower and leave those of us doing the useful work to get on with it”.

I have already stated my own close identification with Hammersley’s arguments, and it is perhaps because of this that I find Chalmers’ lack of substantive engagement with the issues raised by Hammersley even more alarming. I say this as, unless I am wholly misunderstanding Chalmers’ arguments, I struggle to find anything in the paper that adequately addresses or challenges Hammersley’s concerns. Chalmers does not appear able to engage with Hammersley; he dismisses as “polemical” (p.232) Hammersley’s substantive concerns over how much we can realistically expect research evidence alone to improve policy and practice; two-thirds of the paper is simply a restatement of why EBP is supposed to improve policy and practice (p.231-8); and professionals continue to be characterised as clinging to outdated practices despite new evidence.

The above should not be taken to indicate that Chalmers wholly ignores Hammersley’s arguments. However, where discussion is attempted, Chalmers arguably fails to engage on a level that shows an appreciation of the points raised by Hammersley. Whilst Chalmers is justified in challenging Hammersley’s claim that he (Chalmers) advocates that research evidence should directly inform practice, the exemplar used by Chalmers fails to get to the nub of Hammersley’s concerns. In summary, Chalmers compares the (‘non-evidence-based’) medical advice routinely given in the 1960s to place a sleeping baby on their front with contemporary (‘evidence-based’) advice to the contrary, arguing that the mistaken advice of the 1960s has been responsible for a significant number of avoidable infant deaths. Professionals should thus ‘listen to the evidence’ rather than the pronouncements of senior professional figures. It is not clear where, in this example, Chalmers considers professional judgement to be exercised, unless ‘professional judgement’ is confined to agreeing with the results of (Chalmers’ definition of) a systematic review.

This example of medics’ advice in the 1960s does throw some more light on the matter, but not in the way Chalmers intends. I am assuming that the advice given by medics in the 1960s was plausible given a medical doctor’s knowledge of physiology, just as the contrary advice given today is equally plausible. The important point, however, is that medics, both then and now, exercise professional judgement in passing on this advice. It is this matter of judgement (where knowledge from a diverse range of sources is interpreted in the context of the situation at hand) that Hammersley was driving at in his original paper. The exercise of judgement is inherently problematic, but this does not invalidate decisions made using it; neither can it be improved by simple recourse to methodological hierarchies where certain forms of knowledge are considered superior to others (Hammersley 2005, p.88-93).

I shall end this analysis of Hammersley’s and Chalmers’ papers with a brief reflection upon the common aspects of their papers; these relate more to the publishing game than EBP. I bring these up as an issue in order to demonstrate how, in the field of EBP, it is possible to be drawn into quite personal debates that contribute little to the growth of socially-useful knowledge. In their papers, both Hammersley and Chalmers directly state (or imply) that the other is principally interested in protecting their own ‘empire’, and that what they say is basically polemical. The exact meaning of terms such as ‘evidence-based’ and ‘evidence-informed’ get debated at length, but primarily in order to clarify the authors’ respective positions rather than to facilitate the development of ‘EBP’. Both authors also pointedly refer to the other’s publications as ‘articles’ rather than ‘papers’, arguably signifying that neither accords great substance to what the
other writes. Ultimately, the lesson to be drawn from this is that we should make every effort not to be drawn into such destructive ‘debate’.

Why it is Unhelpful to Think Along the Lines of a ‘Hierarchy of Evidence’.

Whilst it is entirely reasonable to critique the utilisation of RCTs in social research, there is arguably a danger in pursuing these critiques dogmatically. I have already stated why I think that Hammersley’s paper (2005) is a rigorous critique; however, this makes the lack of engagement by Chalmers in his reply (2005) even more troubling. If this is what occurs at a high level in the debate, where words are chosen extremely carefully and extensive research informs what is written, it behoves those of us involved at a more junior level to reflect upon how we may substantively contribute to the debate and not be drawn into unproductive argument. It is for this reason that I would warn against expecting impossibly high standards of validity from RCTs; standards that could never be achieved outside of a closed system within a laboratory. It is only a small step from critiquing RCTs upon the basis that they may not be able to provide the certainty of knowledge that some of their keener proponents may argue they do, to dismissing the entire methodology. Gallagher (2004) treads very close to this line (and quite possibly over it) in the use of ‘us and them’ language regarding “… their research methods...” (p.129). Within the social research community, there is arguably still a sense of needing to vigorously promote qualitative research. The contention is that if this is not done, ‘those’ quantitative researchers, with ‘their’ research methods, will drown out the insights provided by qualitative research (for an extreme example, see Freshwater and Rolfe (2004)). Unfortunately, this tendency has led social researchers in the direction of simply rewriting the hierarchy of evidence, with qualitative research promoted near to, or at the very top.

This revision of the hierarchy of evidence does little to aid the development of a pragmatic approach to the use of evidence by professionals and policy-makers. Firstly, placing the focus upon methodology may lead to a replay of the ‘paradigm wars’; by focusing critiques upon the methodologies themselves, they are inevitably defended as a methodology vis-à-vis alternative methodologies. The example of Hammersley’s and Chalmers’ papers could be argued to be a recent manifestation of this. This is not to claim that methodological debate is unimportant, rather it is to suggest that critiques would be far more productive if they were based in a shared framework.

Secondly, methodological critique tends to blind researchers to the potential of other approaches. Instead of condemning RCTs for their inappropriateness, social researchers could consider what elements of a piece of qualitative research may be amenable to testing in an experimental approach. This informed use of RCTs may actually add weight to its critiques, for it would demonstrate an appreciation of the strengths and weaknesses of the different approaches. Thirdly, reiterating the hierarchy of evidence leads swiftly back to the ‘incommensurability of paradigms’, as it obliges researchers to prefer one form of research over another rather than looking at how they may be complementary, or how they could be meaningfully synthesised.

There is a danger that we will become very good at critiquing RCTs at the expense of gaining any appreciation of the role that they could play in social research. Instead of dwelling upon the hierarchy of evidence, social researchers should arguably endeavour to become literate in both qualitative and quantitative methods. This would allow social researchers to not only have a better grasp of the most appropriate methods for performing research themselves, but also inform
cogent critique of others’ use of a wide range of methods. It may also help foster a little humility with regard to how problematic it can be to attain absolute validity within any research approach, and foster more rigorous debate that works towards a common goal rather than the defence of any particular approach.

Bibliography


MAKING WHOOPEE: WOMEN AND ALCOHOL, FREEDOM AND RESTRAINT

Patsy Staddon

Abstract

This paper looks at society’s expectations of women’s behaviour and the reasons, both overt and covert, for these expectations. It considers to what extent women’s use and misuse of alcohol is perceived as a threat by society as a whole and how we might view its alarm at, for example, young women’s binge drinking.

My research has indicated that some women who are treated for alcohol dependence may be worse off than had they remained at home with a bottle or two for company, despite a few scintillating success stories. Sometimes we are unable to recapture aspects of ourselves which were only accessible when we were drunk, and little sympathy or understanding for our loss is likely to be encountered within conventional medical treatment. Sometimes our lives were only tolerable with alcohol for company.

Is freedom to drink as much as we like, or the responsibility to behave with decorum, a particular dilemma for women? To what extent do we carry society’s moral parcel? When we are described as being in need of “help” or “treatment” what is being implied about the way we live our lives? To what will we be subjected if such socio-medical intervention is accepted? Are the issues different for those of us already out of step with heteronormative, able-bodied, Caucasian society?

If capitalist society makes it essential that some of us have less so that others can have more, perhaps women’s position at the bottom of the heap can only continue to deteriorate, resulting in an ever-growing need for them to be dependent on substances simply to survive. Or not.

1) Introduction

Binge Drinking is on the increase!
Young women at risk!
Government fails to act responsibly!

These may be the headlines of any local or national paper on a Monday morning. Pictures abound of young women sprawled helplessly across park benches, an empty bottle alongside; of scantily clad young girls leaning drunkenly against each other in discos.

The headlines and pictures titillate and attract, despite their avowed intent of expressing moral outrage. The young women photographed may look sexually desirable, their helpless drunkenness feeding into expectations as to women’s
passivity and need of protection; or they may be vomiting hideously, provoking
disgust.

Despite the moral panic over young women’s binge drinking, the Government is
pursuing its plans of increasing the licensing hours, and hence its income, in the
face of evidence that this will increase these very conditions. (Room 2005) Is this
evidence of the primary importance that we place on profit? Or, is it evidence of the belief of a white, masculinist, heterosexist culture that
only a minority of people will be adversely affected, and that none of them will be
people like themselves?--only irresponsible or “sick” people who can be controlled
or discounted?

2) Why the outcry occurs

Society’s expectations of women’s behaviour

Young men’s drunkenness has traditionally been a regular feature of Saturday
nightlife; why the excitement, panic and furore over the drunkenness of young
women?

For a start, expectations are different for the conduct of young women. So-called
equality between the sexes has not extended into equal freedom as regards
behaviour, with young drunken women often being considered as responsible for
sexual assaults carried out upon them. (Amnesty 2005) The vomiting, fighting and
yelling on nights out which is no more than we have seen as acceptable, if unwise,
among young men, provokes disgust and fear when the behaviour is that of young
women. Fear because traditional values are being overturned: what will happen
next? Disgust because these women are soiling an image of purity and decency to
which most of society believes they should at least appear to conform.

Young women are meant to be icons of public morality in a way that young men are
not, to stand in for the cleanliness, beauty and sexual purity (Allen 2003) which are
valued but not much practised by everyone else. There is an assumption that
women’s behaviour will be appropriate, whatever the irregularities of that of men,
and the conspicuous consumption of alcohol puts this at risk. “Being drunk ... it’s a
brazen refusal to be quiet, well-behaved and ladylike ... femininity ... relies on, is
defined by, inhibition.” (Lawson 2000) Our society has particular expectations of
women, such as slimness, attractiveness and sexual exclusivity to one adult male.
They are also expected to behave with decorum and to take care first and foremost
of family responsibilities and domestic continuity, putting the happiness of men
and children first. Most importantly of all, they must never bring shame on their
families by looking ridiculous or even by revealing family secrets such as sexual and
domestic abuse.

In addition, a woman’s body is the home of reproduction and for her to “abuse”
her body is seen as a crime against nature in a way that it is if a man behaves
similarly. (Ettoire, 1992 p.10) The dangers of foetal alcohol syndrome are widely
publicised, and a mother who develops alcohol problems is extremely likely to fear
the loss of residence (“custody”) of her children. This can be an important reason
for women to avoid seeking help with alcohol problems. (Staddon, 2004)

In a post-Industrialist and heteronormative society, a preoccupation with women’s
sexual monogamy and their iconic role may be economic and may even be
religious.
Economic factors
As long as a community sees itself as responsible for all its members, everyone is family, and the unit may be the stronger for it. There may be less need to worry over which male fathered which child, and greater sexual freedom within that community may be the result, at least on certain conventionally acknowledged occasions. Communal drunkenness of males and females together remains a feature of life at festivals in many societies, where it may constitute “a strong statement about one’s identity” and “promote social cohesion”. (Harvey in McDonald, 1994 p.211)

When the term “family” is reduced to the narrower and more exclusive meaning it tends to have in the West today, i.e. the nuclear family, concern is felt as to which adult males will take financial and other responsibility for which child, and attempts are made to enforce women’s sexual exclusivity, to ensure that a man does not become responsible for another’s child. Such attempts have become more focused with the decreasing presence of the extended family, where children “born out of wedlock” were frequently adopted within the family by aunts and even grandparents, often growing up unaware that an older sister was in fact a birth mother. (Twinam 1999)

The focus on a woman’s sexual exclusivity does not seem to have lessened noticeably even with female economic independence, and many women choosing to bring up children on their own. There may be a class issue here, in that it is acceptable for professional women to be single parents but not for young mothers in council flats. However women from most groups risk disapproval at best if their sexual behaviour is perceived as being promiscuous.

In societies with less social cohesion and greater individual responsibility, the extent to which the women may join in alcoholic binges, even at festivals, is likely to be manipulated (“that’s an irresponsible, unattractive way to behave”) and control is attempted. Alcohol, commonly perceived as an aphrodisiac (Wilsnack 1984), is a risk factor in such a situation, since its influence may assist women to defy social expectations and pursue sexual pleasure. (Sjoo and Mor 1987) This could be seen as dangerous to a controlling and patriarchal structure and inimical to one which is reluctant to spend individual wealth on the community as a whole. The possibility of statutory support for single mothers is likely to increase such sentiments, since the financial responsibility is likely to reside with all tax‐paying citizens, whose representatives are less likely to be single mothers than married fathers and grandfathers.

Religious factors
In the agrarian and pre-industrial societies mentioned, these seasonal festivals usually include a ritual and spiritual ingredient. At these times, drunkenness is not only for pleasure but also a means of releasing the soul of the drinker in order to communicate with the spirit world. Children born of sexual unions at such times may be seen as having special attributes, even as being gods themselves---threads of such earlier cultures may be identified in the story of Jesus Christ. Such heavy drinking, to the point of hallucination, may itself be seen as sacred. (Harvey in McDonald, 1994 p.213–214)

In the Christian churches too, the taking in of small amounts of alcohol in a ritual setting is seen as communion with the spiritual world, and with each other, as a gateway to higher understanding. The difference in volume consumed is probably
irrelevant; it is the spiritual connection which is important, and the alcohol is a
token only. It is interesting that the role played by both the woman and by the
alcohol in the Christian ritual is narrowed down. In Christianity, with the alcohol
performing a tokenistic role, the woman is expected to be, if not virginal, at least
sexually modest and certainly monogamous; she depicts the containment of
fecundity in a form acceptable to a paternalistic society. In religions based on a
more holistic view of the world, she is abundant, and free to favour whom she
pleases. In some of these communities this will include the ample imbibing of
mind-altering substances. (Sjoo and Mor 1987)

The woman’s role as icon is present in both sorts of community, but the latter is
more likely to be oppressive and controlling of her sexuality. Fear of a woman out
binge drinking might run very deep indeed.

3) What are the reasons for women to defy expectations in this way?

This iconic role is typically predicated upon the psychology of heterosexual
relationships, which in itself may lead to the social dissonance it seeks to
eliminate.

i--Psychology of personal power
It has been observed by previous writers that women frequently adopt a supportive
role towards men, facilitating their apparent independence partly because they
feel they have no choice and partly because it may be easier to experience the
more exciting and dangerous aspects of life vicariously. (De Beauvoir 1949)
Dinnerstein builds on this theory to demonstrate the need of both sexes, seldom
fulfilled, to outgrow their feeling of needing protection rather than “the ultimate
necessity to take care of ourselves.” (Dinnerstein 1987, p. 189) I would like to
suggest that when women decide to drink alcohol to a point at which their vision of
themselves and their circumstances has altered, they may, among other things, be
attempting to bridge that gap, to free themselves of authority and to become truly
themselves. However, in so doing, they may leave men rudderless and afraid,
suddenly aware of their own inadequacy. Such a situation might well alarm men
and women alike.

ii--Sexuality and stigma
Women’s decorum is a heterosexual concept, not equally valued, if at all, by the
lesbian community. (Staddon 2005 a) Feelings of disgust and withdrawal around
drunkenness in women do not appear; admiration of someone’s daring is more
likely. Some lesbians value being able to drink a lot; others have concerns as to
health and safety----but whichever is the case, drunkenness is not necessarily
stigmatised as it is in some heterosexual circles. This may be said to demonstrate
the situational aspect of stigma, where what is discrediting in one situation will not
be so in another (Goffman, 1963) and also to illustrate the heteronormative
pattern of control described above. Where there is no perceived need to control
sexual freedom, there may be less perceived need to control substance use.

iii--Social dissonance
It may be the case that, given the constraints upon the behaviour of the majority
of women which still exist (see above), the use of substances by women and by
other disadvantaged groups is endemic to modern society and may be perceived as
protest, or as the construction of a more satisfactory reality. Surrounded as we are
by warnings about the dangers of liver damage, the likelihood of the development
of alcoholism, and the degeneration of femininity, it might be good at this point to
remind ourselves of the positive reasons for the consumption of large amounts of alcohol by women.

iv—Benefits of alcoholic excess

1. Escape from social control: it helps women of most ages to feel, at least briefly, independent of public opinion. Sometimes the control exerted by social conditioning has prevented their experiencing their feelings—such as anger, joy, sexuality—in any other way. (Staddon 2005 b)

2. Sexuality: it is easier for many women, in a society that prefers them publicly at least to demonstrate decorum, to express their sexuality when drunk. (Traeen and Kvalem, 1996; Wilsnack, 1984)

3. Rage: many women are angry because they are poorer, have more demands on their time, have been abused, and are more strictly controlled. Alcohol helps release this anger. (Kitzinger and Perkins 1993)

4. Fun: having a good time is enjoyable and both hearsay evidence and advertisements for alcohol make women feel they will have even more fun if they consume particular products. For example there is the advertisement for Beefeater Gin, showing a young woman in a pair of boxing gloves: “A Bold Spirit Always Stands Out”; there is: “Grand Marnier: it changes everything”; and even: “Real friends. Real Bourbon.” (From web site showing “55 advertisements for alcohol”: Dec. 10, 2005)

5. Regaining ownership of their bodies: forever taught that they should be slimmer, browner, paler, taller, better behaved, sweeter, better cooks, good in bed, successful at work, good mothers, “some women use substances to reclaim their bodies”. (Ettorre, 1997 p36)

6. Dealing with shame: Between 50% and 90% of women attending substance misuse services may have experienced abuse, either in childhood or adult life, or both. (Barron 2005)

7. Dealing with fear: the behaviour of others may have led to alcohol appearing to be the only “safe place”, with bingeing an alternative to suicide. (Little and Ervin, 1984)

8. Feeling a social misfit: alcohol may throw a veil over awkwardness and insecurity and is particularly valued by some marginalised groups. (Staddon 2005 a)

4) The consequences for women of contravening these social expectations

Reaction of the vocal majority

Screaming newspaper headlines apart, the Amnesty study mentioned earlier showed attitudes around drunkenness in young women often to reflect misogyny, ignorance, spite and chauvinism. Most people thought that women who got drunk and were raped “asked for it”---as they apparently did by wearing scanty clothing and by flirting. (Amnesty 2005)

Appraisal-tendency theory has been used to demonstrate that when people are afraid of the perceived consequences of events they take greater precautions and provide further safeguards; if on the other hand they feel anger about the perceived consequences, they will do the opposite. (Lerner, Gonzalez, Small and Fischhoff 2003) The current tendency to draft in ever larger numbers of police to deal with unruly behaviour would perhaps demonstrate the former, and the indignant newspaper articles the latter: how dare women behave in this way? However precautions dictated by fear may swiftly lead to unrealistic and
unjustifiable solutions which are fuelled by anger, such as racial violence and homophobic aggression.

Public opinion demands a quick solution to the problem of disorder. To facilitate the contentment of the greatest number, and to avoid loss of revenue by reducing overall consumption of alcohol, it may be most politically acceptable to view alcohol problems as being a concern only for a minority of people. These are perceived to be individuals in need of control, whose deviant behaviour is a challenge to the established order. (Moncrieff, 1997) However the nature of the problem, and hence its solutions, tends to be understood in diverse ways.

**Popular solutions**

Many believe that the “problem” of women’s drinking is caused by lack of information. Young women are apparently ignorant about the ill effects of drinking, and to deal with this, fact sheets are published by Alcohol Concern; special funds are allocated to the provision of local initiatives to inform and to educate; and the Portman Group, that well-known and meritorious arm of the alcohol industry, while earning its living from the sale of alcohol, “works to prevent sales to and by under 18s and also to reduce levels of underage alcohol misuse.” (Portman Group 2005)

Others believe that there are those who carry the seed of the problem within themselves; they are “alcoholics”, for whose behaviour there is a genetic causal factor. Born that way, a speck of mouthwash on the tongue, and they are irretrievably lost to sobriety. Popular opinion has largely accepted the beliefs of Alcoholics Anonymous (AA) --i.e. that alcohol dependence signifies alcoholism, which is incurable, and treatable only by lifelong abstinence and following a strictly controlled pattern of life with a very definitive view of the self. This cult concedes that it is only a problem for a genetically programmed minority group and everyone else can safely drink as much as they want without fear of addiction. Naturally such a stance is popular with the breweries, the government, and those who like to drink alcohol. It is also popular with many treatment commissioners, since AA is both free and, like alcohol itself, almost universally available.

Unfortunately for those eager for swift and simple solutions, this still fails to account for the majority of the women who like to drink a lot, but do not necessarily continue into alcohol dependence. (Hall et.al. 2001 pp.416-417; Ragge 1998; Prochaska et al. 1983) These women may be the women who are causing upset by having a good time and making whoopee. It also fails to account for those who are drinking at home alone for comfort, perhaps due to loneliness, bereavement, or any one of a host of other issues. Of course these women are a less obvious threat to social stability, so only become a matter for concern if they stagger out onto the streets, commit crimes, or consult their doctors.

Many of the visible revellers will arrive before long in an Accident and Emergency unit, where some enterprising young doctor may seek to perform a Brief Intervention, which should help them to understand that their drinking is bad for their health and they must cut back; the majority probably will. Others will keep coming back to such units, and even appear in the Courts, for a variety of misbehaviours, and may well be offered Rehabilitation, at a day centre or in a Residential Unit, as an alternative to fines and/or imprisonment. A number will accept.
It is necessary to examine these options from a sociological perspective—i.e., their function and their effect as opposed to their stated function and their claimed effect.

5) What happens to women who receive alcohol treatment

It has been shown that the reasons for some women’s being classified as alcohol dependent are both varied and confused, relying as much on social prejudice as on scientific evidence. It is perhaps not surprising that this confusion is mirrored in alcohol treatment.

On the face of it, alcohol treatment for women is aimed at restoring lost and unhappy people to a semblance of normal life, and certainly it may be successful in doing so for a proportion of such people. The value of this semblance has been debated elsewhere in respect of the Twelve Step programme originated by AA and still the commonest form of treatment to be found in rehabilitation centres. (Peele, Buff and Brodsky 2000) Such treatment units field criticisms of poor long-term outcomes and considerable costs by describing alcoholism as “a relapsing condition”—i.e. no-one can blame them if the person treated by them, at great expense, is back in treatment two years later, or commits suicide in the interim.

These units, with their vaunted solutions, together with the multiplicity of mutual aid groups run by AA, are useful tools in the linked domains of public order and profit—namely the government’s lucrative pursuit of greater availability of alcohol sales to the population. There is some evidence that such treatment centres may help some people to operate to some extent within society, at least in the short term. Hard evidence of longer term success is harder to obtain. On the contrary, the latest UK Alcohol Treatment Trials (UKATT) have identified equally great success at lesser expense of time and money to belong to the treatment models of Motivational Enhancement Therapy and Social Behaviour and Network Therapy, both of which non-residential treatments “are highly effective and save society five times as much as they cost to run, say two papers in this week’s BMJ.” (BMJ Sep. 10, 2005)

A lack of interest in the causes of alcohol dependence is entirely consonant with a desire to marginalise those who “abuse” alcohol, as opposed to those who can supposedly drink alcohol (and pay taxes on it) with impunity. To look at causes might imply looking at ways in which some people are disadvantaged by others, and imply criticism of the tax-paying “moral majority”. It is not at all surprising that there is so little interest in helping people to build new lives which may be dissimilar to the desires of the majority—-one could consider “wet centres”, one could consider addressing the many social injustices, one could make a serious and determined attack on domestic and childhood sexual abuse, a prime factor in the causes of women’s alcohol dependence.

Instead, women are usually expected to go back and live a more circumscribed version of their old lives; without the support of alcohol, but perhaps with the support of the AA groups, with their agenda of hold, contain, refrain. This is logical enough when we consider the underlying reasons for supplying mental health treatment in the first place. “The treatment of women who are diagnosed as mentally ill or unbalanced is oriented towards their resocialisation into their “correct” social role……. Resocialising women into a stereotyped feminine role which is no longer acceptable to many women is no solution to the stress that such a role manifests. It is after all quite conceivable that it is the untenable nature of
the traditional feminine role...that produces a high incidence of breakdown in women.” (Smart 1976)

6) The consequences

There can be no doubt that some women, choosing or being encouraged to participate in conventional treatment, are worse off than if they had not attended, with their addiction still present and their emotional selves gravely damaged. (Staddon 2004; Staddon 2005 b) For these women---57% of the cross-populational sample interviewed in the recent “Making a Start” study (Staddon 2004)---the alcohol services were less a lifeline than a noose. This was particularly the case for lesbian and bisexual women, women who had suffered and/or were suffering domestic abuse and/or from childhood or adult sexual assault, and women (the entire sample) who mentioned mental health problems additional to alcohol dependence.

7) The real object of alcohol treatment

Alcohol treatment, then, does not exist to return temporarily or chronically “ill” women to their “normal” state of “health”. It exists to encourage conformity. “Wet houses” are particularly unlikely to be adopted both for this reason and because deviance must not be perceived as being encouraged but as being corrected.

Treatment is not, in consequence, expected to succeed in many cases, if by “succeed” we mean turn those treated into substance free, happy, productive individuals. It is designed to appease majority opinion, to reduce criminality over at least a short period of time, and to benefit those considerable numbers employed by the Mental Health services, the breweries, and of course the Exchequer. There are massive vested interests involved in making treatment appear to be an effective way to use the financial resources of the NHS and Voluntary sectors, so that alcohol may continue to be ever more increasingly available at ever higher profits. Most treatment centres still employ a Twelve Step model of treatment with a high drop-out rate and an absence of convincing long-term successes. (Miller et al., 1995) This makes such claims as “treatment works” resonate with bitter mirth. Treatment is in effect designed further to stigmatise a minority, while persuading the majority that they were right all along in seeing these drunks and dropouts as worthless and their own alcohol use as completely acceptable. This is particularly so if the recusants are female, and doubly so if they are young, single mothers, supported by the “generosity” of society.

Making whoopee is not a luxury to be allowed many women for long.

References


Allen, Vicky. “The stereotypes of the male and female drunkard are the same as they’ve ever been. In a man alcohol is the harbinger of violence; in a woman, casual sex”, Sunday Herald Online, 28 Dec. 2003.
Amnesty International UK “Sexual Assault Research Summary Report” 12 October 2005
Hall, Bodenhamer, Bolsted and Hamblett. The Structure of Personality. Crown House, 2001

Sjoo, Monica and Mor, Barbara. *The Great Cosmic Mother* (San Francisco: Harper, 1987).


Staddon, Patsy. (b) “Women’s Alcohol Treatment as Social Control”. *Jn. New Directions in the Study of Alcohol*, Number 30. 2005


UKATT. “Effectiveness of treatment for alcohol problems: findings of the randomised UK alcohol treatment trial (UKATT)” *BMJ*. http://bmj.com/cgi/content/full/331/7516/541

UKATT. “Cost effectiveness of treatment for alcohol problems: findings of the randomised UK alcohol treatment trial (UKATT)” *BMJ* http://bmj.com/cgi/content/full/331/7516/544

EUROPEAN IDENTITY AND THE EUROPEAN DIMENSION IN ENGLISH AND GERMAN SECONDARY SCHOOLS

Sandra Barkhof

Abstract

This comparative study on European identity looks at the links between European integration, popular support and European identity in two EU member states, namely in Germany and the UK. In very simple terms, it is argued that in order to sustain the continued widening and deepening of European integration, some form of popular support is required, which would also address, at least in part, the often cited democratic deficit of the EU. It could be argued that a more widespread positive identification with the EU ("feeling European" in addition to existing political identities) would boast popular support. In the past, one of the ways to foster positive national identification was the mass education system. Therefore, this study examines the European dimension in secondary education in the UK and Germany, through an assessment of its theoretical framework and a study of the practical implementation.

It became clear early on that there would be some significant differences between the two countries. For example, the study has demonstrated that German teacher’s attitudes towards a European dimension in education are much more positive than those of their English colleagues. The European dimension is much more incorporated and integrated into the German curriculum, actual teaching at German schools and even the language used by the German teachers. For the majority of English teachers, the European dimension is, although important in general, not a priority and usually submerged into a wider global dimension. The study seems to indicate that national attitudes have a far reaching impact on the extent and type of European identity in the different member states.

European identity and education

The following paper is part of a study analysing the relationship between European identity and popular support for the European Union. Post war European integration could be described as a complex network of multi-level interactions and shared sovereignty that is driven by elites and not as such by the population. In the case of the EU, it could further be argued that there are hardly any direct lines of accountability between the decision-makers and the citizenry. This so-called and often discussed democratic deficit becomes most apparent when considering that there is for instance no opportunity for the population to elect a European Parliament or government with full decision-making powers. The existing system of indirect representation would probably be sufficient if there were reasons to
believe that the majority of citizens were strongly supportive of the integration process and happy to leave the decision-making to the appropriate elites. However, public opinion polls have suggested that in most member states there are considerable reservations at various times (Nugent 2003: p.107). The EU currently seems to lack the tangibility and intelligibility that would enable it to capture the imagination and therefore gain the voluntary support of its citizens (Graham 1998: p.12). Not surprisingly, the national leaders decided at a meeting in 1992 that “the EC had to become more open and receptive to public opinion” (Urwin p.258).

Graham (1998: p.42) argues that at present discords exist between the political-economic integration and the level of legitimisation accorded to such a construct because it lacks the validation of a cultural consciousness vested in place. He concludes that it could thus be argued that the successful integration of Europe might demand an iconography of identity that would complement national, regional and local identities. Van Oudenaren (2000: p.16) agrees and argues that the extent to which a European identity emerges and whether there are cultural and geographical limits to how far such an identity can be stretched will be key questions in future and the answers will determine in part how successful the EU will be in managing the conflicting demands of deepening and widening European integration. Nugent (2003: p.502) even argued that the absence of a firm common identity would make it very difficult to construct and maintain a strong and fully effective political system. However, although it may be that large parts of the affected population identify themselves as European to some extent, there is so far no real evidence to support the belief that there is a European demos with a strong affective bond. Such strong sentiments remain so far reserved for the nation-state (and to a lesser degree for regional attachments) and are yet to be rivalled by a sense of European solidarity (Anderson 1991: p.46). As Smith (1999: p.243) argues, whilst it is one thing for EU elites to identify with and work for a united Europe, it is quite another to attribute such sentiments and beliefs to the great mass of the middle and working classes, let alone the surviving peasantry of Southern and Eastern Europe. If a European identity in whichever form was to be encouraged, the public would have to learn an attachment to the EU.

If we agree that some form of European identity is necessary to boast popular support and thus sustain the widening and deepening of European integration, the question remains which possible means could be facilitated to create or at least stimulate the development of such a European identity. Traditionally, nations are and have been created and moulded by both bottom-up as well as top-down approaches. Although the EU does not wish to construe a “European nation” per se, it may still employ this “system” of political (national) identity construction for lack of alternatives. It seems unlikely that a European identity would be generated through “bottom-up” movements as the EU does not generate the sorts of passions and loyalty that people feel towards their nations (Dunkerley et al 2002: p.121). Therefore, the EU has to tackle European identity largely through top-down methods, at least for the time being. Traditionally, top-down approaches of the political apparatus included the uniformisation of language, the use of the mass media and the mass education system as well as the establishment of national icons (Wintle 1996: p.17). The EU has modelled its past attempts at European identity creation on these traditional approaches and has focused largely on the establishment of European icons such as a flag, anthem, passport and the single currency. Apart from such icons and the creation of a European citizenship, a uniformisation of language or of cultural concepts does not seem appropriate given the fact that the EU wishes to preserve its multicultural character. From the remaining two, the mass media and mass education (both of which are controlled by the individual EU member states), this paper will look at the mass education system in further detail.

**Identity, citizenship and the role of mass education**

In the nineteenth century, the development of modern national identities was accompanied by the rise of mass education systems. The purpose of the educational expansion was to instil
from an early age what it meant to be a citizen of the respective nation state. Thus, mass education was an important facilitator in the processes of state building as well as political unification in the nineteenth century in countries such as Germany or Italy. For a national identity to be established, people must perceive themselves as belonging to the respective nation. As this cannot considered to be a natural attachment, people have to learn to think of themselves as part of a nation (Dunkerley et al 2002: p.53). Therefore, at least in the traditional context of the nation-state, education, as a transmitter of national identity, plays an important role in linking the citizens with the public polity.

Education has another important role. Similarly to nation-states, the EU requires educated citizens who are able to understand what is at stake in elections, understand the workings of the EU institutions and are able to contribute to the economy of a modern society. In turn, in order to be able to participate fully in the institutions of a modern society and to fulfil their duties as European citizens, the EU citizens require education (Dunkerley et al 2002: p.131). In sum, education enables citizens to contribute to society as well as enjoy its benefits (Dunkerley et al 2002: p.131). Through the mass education system, people can also learn about other EU member states and how the different layers of democracy, sovereignty and governance work together.

It is hardly surprising, therefore, that education and training have been regarded in recent decades as increasingly important tools for the promotion of both the EU as well as European citizenship and European identity. As the European Youth Forum Jeunesse stresses in their “Vision of Education in the Future” (General Assembly Malta, Nov 2002; Education 0817-02: p.2), education is one of the main tools by which societies help young people to come to terms with their European identity. Since the creation of a EU citizenship and a European trans-national community life, the “European project” now needs to “capture people’s imagination to insure citizens’ allegiance and secure the future legitimacy of the European Union“ (European Youth Forum Jeunesse 2002: p.6). The economic-political structure of the EU must now be “anchored in the minds and hearts of its citizens by instilling in them the sentiment of belonging to a common European entity with shared values and ideals” (European Youth Forum Jeunesse 2002: p.6). Shennan (1991: p.41) argues similarly that Europe will be shaping the lives of its citizens for the indefinite future and therefore, children are entitled to be prepared for life within European society. Therefore, the educational policies in the post-Maastricht era need to respond to the political reality with an accelerated programme to promote a sense of European identity in preparation for the future.

The EU acknowledges the link between education, knowledge and feeling/being Europeans in several publications, for example the Green Paper on the European dimension of education (1993) stresses that

“[…] the improvement of linguistic competence, the mutual understanding of the practices and cultures of other Member States, and even the ability to work with those of other nationalities or in another setting, are among the most important factors which help young people to become integrated into society and to accept more readily their responsibilities as European citizens “(COM (93) 457 final: 1993: p.3).

The Green Paper emphasises that Europe is not a dimension that replaces others, but one that enhances them. Therefore “[…] introducing the European dimension into our lives does not entail conflict with other levels of belonging” Feneyrou (1993: p.37). Furthermore, the Green Paper (1993) recommends that an education for European citizenship ought to include the learning of languages, knowledge of other countries, transnational exchanges and a better understanding of Europe today and its future construction.
Despite the apparent importance of education for the future of the EU, few research has been carried out to explore the role of education and schools in forming children's perceptions (Convery et al (1997: p.19). The majority of existing studies agrees that schools and teachers are one of the most important sources of information on the EU (compare Convery et al 1997 and Barett 1992). This clearly stresses the important role that teachers and school education play in informing young people about Europe.

The practical implementation

In practical terms, however, Soysa claimed in 2002 (p.12) that so far the attempts of the European Union Commission and other European Union institutions to "europeanise" education have remained largely ineffective. The main reason for this is undoubtedly the fact that the responsibility for education remains firmly in the hands of the EU member states. Although the EU can issue recommendations, the ultimate decision about a European dimension in education rests with the member states and is often further devolved to regional and local institutions. The member states responses to a 1988 EEC Resolution on the European dimension in education made it clear that “in general terms, the main emphasis in almost all the countries was on the provision of European content within traditional subject areas: learning about Europe rather than learning to be Europeans” (Ryba 1992: p.15).

A study has therefore been carried out to have a closer look at the implementation of a European dimension in education. For the purpose of this study, a survey has been carried out among German and English subject and head teachers to compare opinions and attitudes with regards to:

- what is considered to constitute a European dimension in teaching
- the perceived importance of a European dimension in teaching
- the problems that might be associated with implementing a European dimension into teaching
- if and what aspects of a European dimension the teachers or the school include in their teaching.

It was decided to restrict the study to England as the other parts of the UK have different school systems. Similarly, in Germany, there are variations between the curricula of the different federal Länder. It was therefore decided to limit the study to two Länder in West Germany, namely Lower Saxony and Bavaria.

For the purpose of this study, it was decided to focus on the European dimension in the subjects of history and politics/citizenship within secondary education. As Shennan (1991: p.41) points out history should be a prime facilitator of the European dimension as history helps to explain the present European situation, the action of its people and the values of the contemporary society. Citizenship/politics has been selected for the linkages between citizenship, knowledge and identity as outlined above.

The survey was questionnaire based with two similar questionnaires for subject and head teachers. The questionnaires were dispatched to a sample of two hundred

---

39 Compulsory secondary education (Germany year seven to ten of the Gymnasium, England key stage three and four of secondary comprehensive schools)
Gymnasien in Germany and two hundred secondary comprehensive schools in England, which have been chosen at random from online school listings.

The first problem encountered was the difficulty to define the concept of a European dimension in education, which proved to be by no means straightforward. As Savvides (2002: p.4) points out, the lack of a definition of a European dimension for educational purposes poses a real problem for theorists and practitioners alike. In fact, the findings of the present study show that half of the German and nearly 60% of the English subject teachers complained that the lack of a clear definition made the implementation of a European dimension difficult.

Both researchers and governmental institutions have long recognised this dilemma and have made attempts to define the European dimension for educational purposes. It was suggested for example that a European dimension could include learning about aspects of Europe’s past (historic ties between the member states), the present (current ties) as well as its future (common interests) (Feneyrou 1993: p.32). The syllabus should reflect the distinctive character of Europe’s past, whilst shared experiences should take precedence over divisive factors. It could be summarised that a curriculum with “an effective European dimension” (Savvides 2002: p.4) would incorporate the compulsory teaching of at least one modern European language, the cultural differences and similarities of the European nations, the history of Europe (with emphasis on post-war European integration), European geography, the roles of European citizens (fostering a sense of European identity), common European values as well as the EU institutions and policies etc. Thus, pupils should learn about Europe and its member states as well as the meaning of living and working in Europe and being a European (Savvides 2002: p5).

This overview, albeit brief, makes it clear that in general, any definition of a European dimension tends to be rather vague leaving it prone to numerous interpretations. This is a reflection of the EU’s approach to education, as mentioned above, which places the responsibility for education and the European dimension firmly with the national governments, who in turn often delegate the implementation of educational policies further. Both the EU and many of the national governments are careful to avoid too prescriptive formulae, which often results in general confusion among those, who have to implement a European dimension, the practitioners such as head and subject teachers.

Apart from this lack of definition, the vast majority of both English and German subject teachers still considered the European dimension in teaching to be important or even very important whereby the German subject teachers valued the European dimension significantly more than their English colleagues as this graph illustrates:
The head teachers, with their overall responsibility for teaching at their school, valued a European dimension similarly. It must be noted that nearly one fifth of English head teachers attached to the European dimension no particular importance and thus gave it a comparatively low priority at their school.

It is interesting to note that many English respondents regarded the European dimension as a more or less specific part of a wider global orientation. As one English head teacher put it: “we are more concerned with a global (including European) dimension rather than being Euro-centred”. The emphasis was placed on raising awareness of European and global issues, broadening horizons and understanding other cultures.

The language used by the German respondents was somewhat different. The German responses seem to indicate an embracement of the European dimension as a necessicity to prepare pupils for their role in a more integrated or even united Europe. It seems to be almost taken for granted that states and people in Europe belong together and that they will grow closer in future within the developing framework of Europe. As one German head teacher put it “the political integration must urgently be accompanied and complemented by an integration of minds and hearts”. Pupils should be able to “actively help to create their and their countries position in Europe”. Overall, the majority of German head teachers considered a European dimension at school to be imperative for future European integration in terms of cultural integration (understanding, tolerance and acceptance) and European identity formation.

In practical terms, the majority of English (80%) and German subject teachers (79%) claimed to include a European dimension in their teaching. To be more specific, the teachers stated that they included the following aspects of a European dimension as indicated by the following graph:
The graph also illustrates that all factors (with one exception) are more taught by the German subject teachers than their English colleagues. This is especially apparent in the extent to which the history of European integration and the EU institutions are taught in the subjects and year groups in question. Again, the language used by the German respondents seems to convey that the European dimension is perceived to be a necessary build-in component of many subjects in the curriculum to such an extent that “Europe” has almost become a “natural” part of education and the pedagogical jargon, whilst the English respondents often perceive the European dimension as a separate issue that needs to be taught in addition to the curriculum.

### Barriers to the implementation of a European dimension

One of the main parts of this study was an investigation into the possible problems and barriers to the implementation of a European dimension at school level, some of which will be explained in further detail below. As already discussed above, the lack of a clear definition of the European dimension is one of the most frequently encountered problems and can be expected to hinder the practical implementation considerably. Other problems frequently encountered included a lack of training, lack of time and lack of resources. However, apart from a few exceptions, the English subject teachers were more likely to have encountered problems than the German subject teachers, which had a marked impact on the inclusion and commitment to teaching a European dimension, which is partly expressed in the graph above. For example, whilst nearly every second English subject teacher complained about the lack of time, less than one third of German subject teachers had the same problem. Equally, 60% of English subject teachers (but only 24% of German subject teachers) reported that they did not have the necessary training or knowledge to implement a European dimension in teaching. The impact of knowledge (and thus ultimately training) can hardly be overestimated. The study
shows that subject teachers with a better knowledge of the EU and its institutions tend to be more disposed to include a European dimension in their teaching whilst head teachers are more likely to foster a European dimension at their school. Equally, subject teachers with a better knowledge of the EU and its institutions tend to rate the importance of teaching a European dimension as significantly more important and are more likely to be committed to teaching a European dimension and more disposed to include a European dimension in their teaching.

Another issue are the different channels of implementation. The English approach to the European dimension seems to depend much more on the individual teacher, school or organisation, rather than on the curriculum or a nationally established framework as is the case in Germany. Convery et al (1997:p. 34) for example observed, “what scant evidence there is of educating for European citizenship [in England] is found not at national, organized level but at best at the individual school level, if not at the individual teacher level”. Whilst it appears that in Germany the European dimension is at least theoretically firmly anchored in the national framework - and the detailed Länder curricula - the English approach can be described to a certain degree as decentralised with a broader, less detailed and less prescriptive outline. The result is that the implementation of the European dimension can be a bit haphazard and dependent on individual preferences and informal networks. Although the DfEE decides policy on the European dimension, it does not have a direct involvement in its delivery or implementation in compulsory schooling. Effectively, it has been left to the LEAs to decide whether or not they wish to take the European dimension forward. Accordingly, over half (51%) of the English subject teachers in this study complained about a lack of guidance from educational authorities, a sentiment that was only shared by 21% of their German counterparts.

The future of European education

Finally, the subject teachers were asked if the EU should have a more important role in educational policies in the future. Again the opinions of English and German respondents differed considerably. More than half of German, but only one in four English subject teachers, hoped that the EU would have more impact on educational policies in future. Similarly, only one third of English but over half of German head teachers were in favour of such an increased role of the EU.

Many German respondents suggested a European educational model based on the German federal system where the German Länder are responsible for educational policy, whilst the central umbrella organisation (Standing Conference of the Educational Ministers) co-ordinates curricular content and leaving qualifications. This was generally considered to be a useful solution that takes regional disparities and specialities into account. Therefore some German respondents thought that instead of a general European centralisation, a European federal model would be a better and more workable model for the EU. Whilst educational responsibility should remain with the national (or in the German case the regional) authorities, many German respondents agreed that within Europe the leaving qualifications should be compatible to give pupils equal standards and chances in the European market. In addition, the school systems and teaching targets should be similar to ensure that pupils leave school with similar education levels. A few German head teachers remarked that within Europe, the member states should learn from each other and teach tolerance for other cultures to create a European knowledge base and European identity. None of the German respondents expressed fears of identity loss as some English respondents did. On the contrary, the creation of a European identity to counterbalance national identities was cited by many German respondents as a reason to be in favour of an increased role of the EU in educational policy.
In comparison, the general consensus of the English respondents was that more central interference should be regarded as a negative impetus on the educational system. As one head teachers put it: “[there is] too much central interference as is… leave education to the professionals”. Although co-operation to some degree among schools, both national and international, for example through e-learning and video conferencing, was considered to be useful, it was generally thought, that schools should have more educational autonomy, not less. This attitude is a reflection of the English educational system, where schools enjoy greater freedom concerning curricular interpretation than German schools, which operate in a much more prescriptive and regulated curricular framework.

In conclusion, the study has shown some interesting differences between England and Germany regarding the perception and implementation of a European dimension in secondary education. Further research is recommended to assess for example in how far school education has influenced the attitudes of young people towards the EU and in how far this has fostered a European identity in whichever form.

Bibliography
London: Verso

Oxford: Butterworth-Heinemann

Brussels: Commission of the European Communities COM (93) 457 final

London: Cassell

London: Routledge

General Assembly Malta, Nov 2002; Education 0817-02

Barcelona: Col·legi de Doctors i Llicenciats en Filosofia i Lletres i en Ciencies de Catalunya

Graham, B. (Ed) (1998) Modern Europe: Place, Culture, Identity
London, Sydney, Auckland: Arnold

5th Ed Basingstoke, New York: Palgrave Macmillan

268

Paper presented at IREDU, University of Bourgogne 02.12.2002

Shennan, M. (1991) Teaching about Europe
London: Cassell

Smith, A.D. (1999) Myths and memories of the nation
Oxford: Oxford University Press


London: Longman

Van Oudenaren, J. (2000) Uniting Europe

Aldershot: Avebury
OVERCOMING THE MUDDLE AND MESS: THE NEED FOR PRAGMATISM AND FLEXIBILITY IN AN EMPIRICAL STUDY OF INFORMATION SHARING IN CHILDREN’S SERVICES

Sue Richardson

Abstract

Anyone working in the field of social research is aware that empirical work rarely goes ‘to plan’. The process is usually far more messy than would be imagined from the plethora of books, articles, dissertations and conference presentations reporting on methods of collecting data ‘in the field’ for social research. Although this post-hoc ‘airbrushing’ of reality is not news and others have written entertainingly on the subject, it is worth reminding ourselves, as research students, that pragmatism and flexibility can get us a long way in navigating a route out of the chaos into some semblance of order. This paper catalogues some frustrations experienced in conducting field work for a doctoral thesis and offers hope as the light is starting to appear at the end of the tunnel. Although the focus of the paper will be on the process of conducting the research, the aim is also, because of the examples that are used, to discuss some of the substantial issues that are being tackled in the research. These are the legal, ethical, governance, cultural, technical and human resource challenges that practitioners working with children face when sharing service user data with colleagues in other agencies.

Introduction

This paper reviews the research carried out in the first two years of a Social Science PhD, focusing on the discrepancy between an idealised ‘real world’ research project and the actuality. Although it is generally acknowledged that conducting empirical research in the social sciences can be ‘messy’ and rarely goes ‘to plan’, the vast majority of research outputs, be they books, articles or conference presentations, do not give mention to this aspect of the work. There are good reasons for this. All professional social science researchers are well aware of the phenomenon and they expect their colleagues to understand that, given limited space in any research output, the most important aspects to convey are the findings and implications of the research. The methodology will also be a necessary component to enable the reader to judge the quality of the work in order to apply appropriate qualifications to the research implications and to allow replication if appropriate. To state explicitly that the research was messy and in places muddled is an unnecessary distraction. So, the reason for the exclusion of these ‘data’ is not simply that it would be embarrassing, and would feel uncomfortable, although this may remain a minor element, even at the highest academic levels.
Having provided justification for not drawing attention to the less ordered expects of our research, it feels that, from time to time, it is important to remind ourselves of its existence and to bring it into the light. This is particularly important in forums where the audience will be student researchers. If the only window onto the world of research is standard academic research outputs and texts on research methodology, the first time that students confront the often chaotic nature of social research will be in their own projects. It could be easy to conclude that they are ‘doing it wrong’ if not adequately forewarned.

This paper therefore aims to take the reader on a tour of the initial two years of an empirical social research project. It examines the possible sources of the mess and muddle in such work and goes on to compare the research plan with the research as it unfolded. This is not an entirely pessimistic account. It concludes that it is possible to navigate a way through, by taking a pragmatic and flexible approach and by remembering to use the anchor of the research; its overall purpose.

**The source of ‘the mess’**.

What I am calling ‘the mess’ can result from the content (i.e. the subject matter) of the research or from the process of the research or from both and in the case discussed here, both were sources and will be discussed in turn. (Mess can of course also derive from the researcher themselves or their interaction with the research but this must be left for another article.)

*Research content as a source of ‘mess’*

The topic that I wanted to investigate was the inter-organisational sharing of personal information in health and social care services. It seemed that every few years there is a tragedy from public service failure that results in an inquiry. The conclusion from such inquiries is usually that agencies should work together better and improve their information sharing. A recent example is the report of the inquiry into the death of Victoria Climbié (Lord Laming, 2002) which criticised the way that agencies worked together, specifically the lack of information sharing. Only last month, another report was published following a case of serious child neglect that concluded that agencies must improve the way that they share information (Cantrill, 2005). On the face of it, the recommendations were straightforward enough, but they have been broadly the same for many reports over recent decades. Why, I asked myself, do tragedies continue to happen if the supposed solution is well known? I concluded that this was not simply a case of the recommendations being ignored. With the Every Child Matters initiative and the implementation of this through the Children Act 2004, the recommendations of the Laming report (2002) are clearly being taken seriously. Another possibility is that ‘improving inter-agency information sharing’ is easier said than done and I decided to investigate this issue ‘on the ground’ and ‘at the front line’ to try to understand the difficulties that are being encountered.

The policy background to such a study at the start of the 21st century is ‘partnership working’, through New Labour’s ‘third way’ approach to public service delivery. This approach criticises both the old forms of ‘market’ and ‘bureaucracy’, favouring a ‘network’ approach to governance (Parsons, 1995, Balloch and Taylor, 2001; Newman, 2001; Glendinning et al, 2002). This has resulted in the current situation where agencies function within a complex of matrices. A social worker might work in a multi-disciplinary (or multi-professional) team within a semi-integrated service that reports to a number of bodies of different kinds, including the ‘local community’ and a local strategic partnership.
He or she may report to a health worker employed by an NHS Primary Care Trust or a voluntary organisation as their team manager but will also have a supervision manager employed by the city council. Local strategy will have been devised by a number of bodies and may not be coherent as a unified ‘strategy’. Training and professional development may also be provided by different bodies that may or may not share the professional framework or ethical stance of the social worker. He or she might use different computer systems (both hardware and software) from their team colleagues. They might not even share common notions of a care pathway or use a common assessment framework, although the rationale for having a partnership approach with integrated teams is that a service user does not need to undergo multiple assessments and duplicate their story to different professionals.

Once upon a time, let’s say, more than ten years ago, an organisational analysis of a service would have resulted in something that looks a little like Figure 1.

**Figure 1. Organigram of a hypothetical service, circa 1990**

Most public services in the 1980s and 1990s (as for decades before) had been fairly autonomous and hierarchical. Strategies, including operational and human resources strategies, would have been devised, either at the level of the Department Heads or at the level above that. Each team would probably have been established at a single location and sometimes all the teams in the service would have been co-located. Since the introduction of New Labour’s partnership policies, an analysis of a public service would look more like figure 2.
Figure 2. Organigram of a hypothetical service, circa 2005

Figure 2 is a more complex picture at both operational and strategic levels. Indeed, in this complex image, individuals have not even been represented in the way they have in figure 1 but they can be imagined as inhabiting the teams or partnerships or organisations that are represented by the blobs. As an example of how things have changed, we can look at the lines of accountability. The lines of accountability for a social worker working within a local authority social services department ten or twenty years ago would have been relatively straightforward (see figure 3).

Figure 3. Lines of accountability for a social worker in a hypothetical service, circa 1990

Now the situation is more like that shown in Figure 4.
Returning to the purpose of the research, my main objective is to understand more about inter-organisational information sharing. To discover why it continues to be a problem, I wanted to think about the potential influences on inter-organisational information sharing behaviour. The challenge, to be met by public service professionals, whether they work in health, social services, education or criminal justice, and whether they work in the statutory, voluntary or private sector, is to walk a tight-rope. They need to ensure that information that is vital to the well-being of an individual, a family or a community is passed on to other professionals who need to be informed while at the same time ensuring that information held about a service user is not shared inappropriately or without good cause, maintaining data privacy and confidentiality. The stakes are high. Falling off the tightrope on either side can result in tragedy.

Based on the logical possibilities, I conceptualised a simplistic model of the patterns of information sharing behaviour (described in detail in Richardson and Asthana, 2005). This model suggests that the four basic types of information sharing behaviour are chaotic (neither shares nor protects information appropriately), over-open (protects information appropriately but shares it without good cause), over-cautious (shares information appropriately but protects it without good cause) and ideal (both shares and protects information appropriately).

The literature review allowed the consideration of the possible influencing factors on such information sharing behaviours. As discussed already, different policies on public service delivery can change the way in which agencies work together and it can be hypothesised that policy context will influence information sharing behaviour. Legislation surrounding information sharing has undergone considerable change since 1989 with both the Human Rights Act and the Data Protection Act in their current form being enacted in that year. The Children Act 2004 has shifted
the emphasis in the balance between sharing and protecting information more in favour of sharing. We have already mentioned changes in governance and accountability in services and changes to the ways in which teams work now that collaboration has a higher priority. How these (often multi-professional and multi-agency) teams are managed again might influence how services share information. Management of the administrative and technical systems could be equally influential and so we could suggest that different forms of records management and information/communication systems management might facilitate or inhibit appropriate information sharing.

Another level of influencing factors that was identified in the literature review was the personal level and a potential influencing factor could be hypothesised to be the individual’s confidence to make the right decision with respect to information sharing; their confidence in balancing the risks involved. This confidence would seem to be mediated by a number of factors at other levels, including the quality of training and the levels of trust in other professionals which may, in part, be determined by the different cultures and world views that exist in different professions (Richardson and Asthana, 2006).

I concluded that I would need to investigate at least ten factors as possible influences of information sharing behaviour:

Environment/contextual level
- Policy context
- Legal framework
- Professional culture

Organisational/structural level
- Governance and accountability
- Team management
- Records and case management
- Information technology
- Training and support

Personal/agency level
- Confidence
- Trust

This research, then, was going to be complicated (and messy) just to conceptualise, let alone operationalise.

The research process as a source of ‘mess’
So much for the mess arising from the subject of the research. What about the research process itself as a source? Texts on research methodology often acknowledge that ‘qualitative inquiry’ can be an iterative process with each part dependent on the outcomes of the last, e.g.,

‘The relationship between literature, question formation, data collection and theory building in qualitative research often involves on-going modification, with data collection leading to emergent theories which themselves redirect the data collection process.’

(David and Sutton, 2004:77)

The implication here is that the totality of the research cannot be planned in detail at the outset and there will be much that cannot be predicted at the
research design stage. Some authors have gone so far as to devote an entire book to the usually unspoken aspects of conducting qualitative research (Byrne-Armstrong et al 2001).

It is worth questioning the standard qualitative/quantitative distinction in this respect however. I particularly take exception to this distinction being made at the research planning stage of a research project. I can agree that at the data analysis stage it is helpful to deal with qualitative and quantitative data differently, in fact we have no choice but to do so. It is also true that the way in which data are going to be analysed needs to be considered at the planning stage. I agree that research that is often classified under the heading of ‘qualitative inquiry’ is taking a particular philosophical approach to the work which is distinct from that which is often classified under the heading of ‘quantitative research’. My preference though, would be to accurately label the philosophical approach being taken, e.g. ‘phenomenology’ or ‘natural realism’ because it is possible to include ‘quantitative’ data in a ‘qualitative’ study and vice versa. With regard to a general description of the research design, I like Colin Robson’s distinction between fixed and flexible research designs (Robson, 2002:4, 163-4). Sarantakos (2005) makes use of the same terminology and shows in diagrammatic form, the cyclical nature of flexible research (2005:113). Robson does suggest that flexible designs are likely to be untidy and states that, ‘Doing flexible research design calls for flexible researchers’ (2002:167), a statement that it is worth remembering, frequently.

I would propose though that all empirical social research includes elements of disorder and can be difficult to control and therefore to plan (with any confidence of the plan being adhered to). The reason for this is that empirical social research is interested in the real world. Even in the most rigid of experimental designs, it does not want to isolate variables and control them in a laboratory. Social research is almost always dependent on a great number of other people doing whatever they do in their lives, wanting to participate in the research. Whether a fixed or flexible design is used and whether quantitative or qualitative data are analysed is irrelevant, or rather, it is only makes a difference in terms of the degree to which there is likely to be a lack of order and predictability. As Robson adds after describing the research process, ‘In the real world, of course, it won’t be as neat and tidy as this.’ (2002:83).

To illustrate the way in which an orderly research plan can disintegrate into a fuzzy mess and (hopefully) re-emerge into something meaningful, I will expose the case history of my research to examination here. Bear in mind however that the very process of structuring the account to relate it involves a degree of ordering and censorship, and you can be sure that any completely embarrassing events will be (at least subconsciously) conveniently omitted.

To look at the muddle and mess resulting from the research process in my case, I will compare the initial research plan as it was presented to the ESRC and the reality of the process to date.

The Plan

The title of the planned research in the initial stages was, ‘Understanding the Complexities of Inter-organisational Information Sharing in Public Service Partnerships’. As with most PhD titles, this has changed (and no doubt will again before submission) reflecting, to an extent, some of the difficulties encountered along the way. The current working title is ‘Inter-organisational Information Sharing: Rising to the Challenge in Sure Start Children’s Centres’.
The research proposal stated three objectives, which were to be achieved through the three proposed stages of research. The objectives were:

- Development of a framework for the analysis of inter-organisational information sharing through a primary case study of an integrated service, in Plymouth
- Replication in two comparative but distinct case contexts to establish a case independent framework
- Development of a practical tool-kit to support inter-organisational information sharing.

The proposal included a planned research timetable:

**Phase 1. Months 1-14**

**Phase 2. Months 15-22**
Data collection in secondary cases, analysis of finding across case study partnerships, establish sampling frame of respondents for phase 3 partnerships and negotiation of access to them. Write-up of results on-going.

**Phase 3. Months 23-30**
Draw up checklist of indicators, administer survey of partners and analyse results, develop policy recommendations and toolkit.

Final collation of results and write-up. Months 31-36

The proposal specified that documentary, observational and informant sources of data were to be used and these were specified in some detail.

**The reality**

Having undertaken my Masters project in a subject that involved interviewing NHS staff, I knew that I would need to obtain National Health Service (NHS) ethical approval for my PhD research as NHS staff were involved again. Indeed the first activity listed in the proposed timetable was ‘Application for NHS Local Research Ethics approval’. I was aware that the process could take many months and that it should therefore be started as soon as possible.

Before the proposal was submitted to the Economic and Social Research Council (ESRC), in fact nine months before the PhD started, I had negotiated access with the service that was being nominated as the primary case; a newly-formed integrated drug and alcohol service in Plymouth. I prepared my NHS ethics form in the autumn of 2003 at the beginning of the research when the other main activity was the literature review.

**Major hiccough number 1: Access**

I felt it was important to re-establish contact with the primary case before submitting the research ethics application for the first phase of the research. I wanted the case service to play an active role in the research process and I knew
that the research could not take place without the willingness of the service as a whole (and sufficient individuals within it) to participate.

I went back to the service and found that a number of changes had taken place since my last visit, including the appointment of an operational manager with whom I had not previously had personal contact. The service director I had negotiated with previously was still willing, in principle, for the research to be done but he felt it was important for the new operational manager to make the final decision.

At our meeting, I presented my ideas about the research, the rationale for the research design, the potential costs and benefits for the service of being involved. It was clearly going to be a difficult decision. The benefits of involvement were acknowledged but there were also the costs to consider. These were mainly in terms of staff time and this was a resource with many demands upon it already. I knew that, difficult although it would be to start from scratch at this stage, it would be better than collecting some data, only for the service to find that it was taking too much staff time and deciding to withdraw at that point. With an incomplete case, I would probably still need to start again and would be even further down the line. On understanding this situation, the operational manager came to the decision in October 2004 that it was best to decline the invitation to participate.

By the time this decision was made, I had been able to review a good deal of the literature and policy documentation on information sharing and was struck by two things which helped to direct me towards a new primary case for the study.

One of these was that the majority of cases in the literature discussing information sharing were within a health and social care context. There were exceptions, for example the case of the Soham murders, where the Bichard inquiry identified problems with sharing information between different police forces with the result that an education service was provided with incomplete information (2004). However, I decided to focus on health and social care and in fact, to narrow the focus to children’s services where the literature on information sharing is concentrated. Part of the reason for this is that when information sharing fails in children’s services, the results can be particularly serious, meaning that this is probably the sector where a) there is the greatest need to tackle the problems and b) where there will be high recognition of the issues involved.

The second way in which the literature review was able to re-focus the research concerned the potential importance of the newly established policy of agreeing information sharing protocols. By combining these two new directions, the revived research plan was beginning to take shape. I could study one or more of the information sharing protocols (ISPs) that have been agreed for children’s services in Plymouth in order to study information sharing at strategic level and investigate one or more children’s services which would be covered by the ISP(s) to understand the influences on information sharing behaviour at the operational level. I chose to study Sure Start Children’s Centres.

**Major hiccough number 2: NHS research ethics approval**

As noted in the previous section, not only had the planned research changed with respect to the primary case to be studied but also there was a new emphasis on the agreement of the local information sharing protocol. The first draft of the
research ethics applications (both NHS and faculty) had to be abandoned. The research design and the detail of the methodology needed to be rewritten for the research protocol and for the completion of the fifty-seven page form for the NHS research ethics application. Yes, that is correct; it is not a typographical error. The NHS form was fifty-seven pages long. This did not include the protocol itself (six copies of which were to be supplied), the participants’ information sheets, the participant consent form, the interview topic guides, the applicant’s and supervisor’s CVs, the application checklist or the covering letter.

The NHS application was eventually submitted on 22 March 2005. It was reviewed by the local research ethics committee (REC) on 12 April 2005. I received the verdict in the beginning of May. The Committee had been unable to make an opinion to accept the application because it had three concerns. One was that I would not be able to conduct the number of interviews specified (around 40) in the time (six months). The second was that one of the sentences should be rewritten as it was too technical and was in a section to be in ‘lay language’. The third was that there would need to be confirmation that tapes used for interview would be kept in a secure location and erased following transcription.

I sent a letter to the chair of the research ethics committee responding to these objections and received notification of a ‘favourable ethical opinion’ by the committee on 19 July 2005. This, however, was not the end. The research governance process had to be undertaken with the Trusts involved and this included applying for an honorary contract with the Primary Care Trust. One of the Trusts requested that I change my participant information sheet to make it clearer that the information I collected would be used for PhD research (although this had been explicitly stated). This meant that I had to go back to the REC to get approval for that change of wording. I received my final notification, allowing me to start the research on 27 October meaning that I had begun the third year of my PhD before being able to recruit participants for the research.

The time delays resulting from these two major hiccoughs meant that my research plan needed to be seriously amended and the first hiccup also had a significant effect on the focus of the research. There were other, more subtle consequences. Two resulted from the fact that my period of ‘literature review’ was longer than anticipated and the period of ‘data collection’ was shorter than originally planned. A positive consequence of this was that I was able to write two journal articles based on the literature review while waiting for ethics approval. One has been published. The other has been accepted for publication and will be published later this year. A further consequence was that, having gathered and analysed a large amount of literature, I have decided to structure my dissertation in a way that had not occurred to me when I began. Instead of having one large chapter on the literature as a background and rationale to the study (as I had intended), I have a standard-size chapter that serves this function but have been able to weave aspects of the literature review into other, subsequent, chapters.

A serendipitous consequence of my data collection being delayed was that a useful article on measuring the integration of services (Browne et al, 2004) has been able to inform my empirical work. For this alone, I am grateful to the peculiar meandering route that the research has taken.

**Learning the virtues of pragmatism and flexibility**

Since I started on a research career in 1979, I have conducted many different projects, most of them (but not all) empirical real world studies. However, this is
the one that has gone least to plan and has probably been one of the best for reminding me of the virtues of pragmatism and flexibility. Key to the ability to make appropriate changes has been the constant anchor of the overall purpose of the research. This has remained unchanged. The experience has taught me (lest I had forgotten) that mess and muddle is inevitable in empirical social research and that I should ‘get used to it and get on with it’. I have learned better how to balance good planning with letting go. Knowing you will need to be flexible does not negate the need for planning; it merely requires continual planning revision. Finally, I have learnt to enjoy dealing with the unexpected and am reminded that it is this that makes real world research exciting.

Bibliography


THE APPROPRIATE ROLE OF THE STATE WITHIN THE ETHICAL PARADIGM

Alan Sanderson

Abstract

The traditional role of the British State in determining and protecting the public interest was informed by deontological ethical principles that determined what actions were right or wrong, what constituted the concepts of good or bad. However, in seeking to achieve an innovative modernisation of the public sector, this hierarchical mode of governance has now undergone a transformation into a hybrid organisational form of new public management that seeks to combine elements of the neo-liberal market-driven model of public administration with its antithesis—the entrenched Weberian bureaucratic model.

It is proposed that lower income communities embraced sceptical tenets of morality as a bulwark against the perceived insincerity of neo-liberal conservative market orientated outcomes. Thus, they acceded to a physiological strife founded on the fundamental difference between deontological and consequential ethical outcomes. Moreover, new labour’s political elite has exacerbated this ethical scepticism through a strategy of managerialism, producing an ethically ambiguous scenario causing uncertainty for all stakeholders.

In this ambivalent contemporary public arena, where efficiency and effectiveness are as important as notions of equity in the delivery of policies that are underpinned by a regulatory vision of “the public good”, political leadership seems destined to cope with an indeterminacy that requires perpetual transfiguration and optimal opportunism. This inexorably leads to the risk of both overt and direct political authoritarianism at any level of government.

In the Oscar winning film “Network” Howard Beale, the mad prophet of the airwaves, exhorts his poor, disaffected and disadvantaged audience to throw open the windows of their dilapidated apartment blocks, put out their heads and yell “I’m as mad as hell, and I’m not going to take this anymore!” As Americans respond in their thousands the unmistakable sense of purposelessness in an uncertain, and often malevolent society, is instantaneously apparent. These people have little faith in the integrity and sincerity of the state having witnessed ambiguity, and sometimes, rampant hypocrisy in the attitudes, opinions and behaviour of their elected representatives. They struggle to find any moral codes or sets of principles that offer a means of interpreting the issues and agendas of modern life. Of course, this bleak picture of public discontent might be designated as belonging in the realm of cinematic invention. However, as the scenario resonates with the scepticism and cynicism that pervades those in contemporary British Society who experience “the restricted citizenship of those who are poor” (Beresford et al., 1999: 27), it seems to warrant more than dismissal as dramatic license.
It is probable that, as the role of the state in advanced liberal democracies has become increasingly restrained due to the ineffectiveness in both the formulation and implementation of public service policy provision (Hult and Walcott 1990, Kooiman 1993, Weimer and Vining 1997) and the policy constraints caused by fiscal controls that have arisen from the globalisation of economic supply and demand (Bovens and t’Hart 1990, Bovens et al., 2001, Gray 1998, Sieber 1981) that the traditional precepts of bureaucracy had to give way to the canons of the hollowed-out state that are informed by the principles of neo-liberal doctrine. However, these clear distinctions offer an overly simplistic framework for an adequate appreciation of the particular course of events that have characterised recent policy-making by the British State. Undoubtedly a traditional bureaucracy was challenged by neo-liberal conservative doctrines that upheld the principle of privatisation and recognised the supremacy of the market-based economy.

However, a subsequent new labour managerialist state, that accords supremacy to regulation rather than ideological conviction, has eclipsed this economic and social experiment. Nevertheless, whilst acknowledging the differing ideological configurations, it is argued here that there has been a sustained growth in ethical scepticism caused by the belief that the state has embraced values that are unjustified, unreasonable and uncertain during the various political administrations of Margaret Thatcher, John Major and Tony Blair. In order to contextualise this assertion this paper begins by briefly reviewing the ethical foundations of the traditional bureaucratic state.

Deontological Ethics and the Bureaucratic Model

In this ethical paradigm the appropriate role of the state accords with deontological principles that restrain individual initiative by imposing “the recurrent patterns of social behaviour that determine the nature of human action” (Parker, 2000: 125). Thus, bureaucratic structures that function like sophisticated machines with a clearly defined hierarchy of full-time and salaried personnel, separated from the resources that they direct, (Weber [1904] 1976) employ their knowledge and that of their professional subordinates to exercise control over individual agency. So, structure has assumed a causal capacity as individuals’ decreed duties become objectively knowable.

Therefore, deontological ethics are concerned with what individual duties are, who has rights, and what strategy is best able to produce justice. These fundamental precepts lead to the development of a code that defines what actions are right and permissible, and thus what actions are wrong. As Blackburn observes, “they take us beyond what we admire, or regret, or prefer, or even what we want other people to prefer. They take us to thoughts about what is due. They take us to demands” (2001: 60). As Kant concluded, these demands are derived a priori or from pure reason instead of individual experience. He insisted that for people to accept moral laws their construction must be “freed from everything which may be only empirical” (Kant, [1785a] 1998: 289). Thus, individuals do not construct their morality by considering the consequences of their actions, but, instead, discover their inherent capacity to act morally or dutifully. This process of enlightenment lies at a deeper level than that of affectation, as individual behaviour should fully comply with the intent of a duty, rather than just observe its tenets, if a person is to achieve the particular postulates of Kantian “good will.” From this process of subjective awareness there arises a code of objective ethics, which accords with the thinking of the elitist in that impartial standards of behaviour are created, which can be subject to dispassionate judgement. As Kant maintains, judgement
must be passed on what is right and what is wrong by the use of pure practical reason thus making morality absolute.

When a person acknowledges their moral obligations, they accept “the categorical imperative,” or that moral rule that recognises that human characteristics – such as loyalty and duty – possess a discrete inherent value. This distinction is clarified by Kant in his statement that if an “action is good only as a means to something else, then the imperative is hypothetical; if it is conceived as good in itself and consequently as being necessarily the principle of a will which of itself conforms to reason, then it is categorical “ ([1785b] 2003: 2). Following this assertion he proceeds to confirm the existence of “but one categorical imperative, namely this: Act only on that maxim whereby thou canst at the same time will that it should become a universal law” (Kant, [1785b] 2003: 6). This fundamental principle is often cast into the popular saying “do unto others as you expect them to do to you” although this cliché does not fully accommodate the extent of Kant’s insight.

Pure human reason, which inspires individuals to observe the categorical imperative, was effectively tempered by Edmund Burke’s political vision of the supremacy of good order. In his envisaged sequence of events, politics was to be inspired with a religious vision to explain the unfair distribution of resources (Gaede, 1983: 110). Thus, as religion is the grand prejudice, using the suffering experienced in this life as a portent of rewards in the next, Burke dismisses his detractors with the assertion “you think you are combating prejudice, but you are at war with nature” ([1790] 1993: 49). Moreover, the well ordered state, a product of a slowly evolving and traditionally informed pattern of governance acts as a moral mainstay as custom reconciles us to everything” (Burke, [1756] 1987: 148). In this schema subjects enter a special relationship with their society as “it is a partnership in all science; a partnership in all art; a partnership in every virtue and in all perfection“ (Burke, [1790] 1993: 96). This ethereal bond is held together with a code of deontological ethics where the elite accept their obligations to their subjects in return for the proper fulfilment of duties owed to the state and its dominant hegemonies.

Burke’s prosthetic to re-inspire political dialogue had an influential and lasting effect. Its enduring relevance and effectiveness is reflected in Walter Bagehot, the Victorian constitutional expert, finding the attitude of the English working class towards authority as rudimentarily “deferential” ([1897] 1963: 235). Moreover, this observation could have been labelled quintessentially European as the masses of Germany, France and Britain enthusiastically marched to war in 1914, more than adequately demonstrating the manner that majoritarian democratic societies were able to harness the notion of patriotism through gradual, and sometimes even overdue, concessions to their poorer citizens to ensure their compliance in an endeavour of unparalleled massacre. However, as the example of the Russian Revolution was to emphatically demonstrate, these manoeuvres were for the highest stakes with permanent social change a consequence of their failure (Hobsbawn, 1987: 164).

The inter-war years witnessed pressure on the British State “to take on greater social responsibilities and to intervene to provide direct help to the most vulnerable sections of the community” (Stevenson, 1984: 306). The further concessions made during this period were then incorporated in the institutions of the Welfare State in the 1940s. This balance of rights and obligations were to become “genuinely popular with the mass of the electorate of all classes” (Glennerster, 1995: 12) during the subsequent decades.

However, by the 1970’s the notion of a political social policy settlement was disintegrating in the realities of rampant inflation, industrial strife and substantial increases in the price of oil. The traditional bureaucracy, with its deontological
ethical underpinnings, appeared incapable of addressing the need for economic re¬structuring thereby ushering in a decade of what was to become known as “Thatcherism”. This body of thought had found some of its fundamental ethical tenets in the principles of consequentialism where free beings can only be motivated by material reward. It is in this imperative, which now informed the formulation of public policy, that the seeds of scepticism might have been sown.

Consequential Ethics and Neo-liberal Conservatism

In this ethical paradigm the appropriate role of the state accords with consequentialist principles that recognises that an individual has both the causal capacity to act and the discernment to objectively assess whether the likely results from their actions will have good or bad consequences. Thus, “the moral value of any action always lies within its consequences and it is by reference to these consequences that actions, and indeed such things as institutions, law and practices are to be justified if they can be justified as all” (Smart and Williams, 1973: 79).

In making the value judgement that an action that results in the greatest happiness for the greatest number of people is morally right the proposition expresses a value-predicate — happiness — that is applied to the subject — the greatest number of people — so, neo-conservatives could accept that ethical statements should be articulated in the terms of social aggregation and expect the value-predicate of happiness to be analysed in objective denominations that measure the extent of material well-being (Taylor, 1975: 176). The need for such an instrument of measurement becomes clearer with the practical application of the act-utilitarian doctrine. This states that “the only reason for performing an action A rather than an alternative action B is that doing A will make mankind (or, perhaps all sentient beings) happier than will doing B” (Smart and Williams, 1973: 30). Thus, the neo-liberal conservative perspective, in embracing act-utilitarian ethics, require an objective means of assessing the anticipated consequences of actions, so as to be able to determine what would constitute the greatest aggregate or accumulative happiness. Therefore, as neo-liberal conservatives would choose to negotiate the preferred constituents of their own well-being with others, they rely on the mechanisms of the free market to act as an instrument of evaluation of social activity that facilitates judgements of ethical consequences. Nevertheless, Plant (1999: 20-1) can identify three propositions that demonstrate the unprincipled nature of the market. When these are combined, they offer a convincing case for the rejection of the market’s capacity to convert the abstract notions of “right” or “happiness” into synthetic statements that can be measured in terms of each individual’s transactions.

The first proposition, made by Hayek (1960) and Acton (1971), maintains that a just market transaction is one devoid of coercion. As individuals enter into free exchanges, where inequalities of power are redressed through the freedom to negotiate and enter into binding contracts in the full awareness of their personal rights and responsibilities and of the outcomes arising from their actions, such transactions cannot be deemed to be unjust.

The second proposition is that premeditation is a necessary pre-requisite for an action to be deemed unjust, which means that outcomes from self-interested market transactions cannot be unjust. Instead, the myriad number of daily transactions, which together constitute market activity, produce a spontaneous order amongst market participants that is not directed by pre-determined measures of income re-distribution (Hayek, 1978: 183).

The third proposition is, as Nozick notes, that while players in the market can serve moral imperatives “the market mechanism does not especially reward us for satisfying those desires, rather than other desires that are neutral towards or even
retard those people’s development” (1981: 514). Thus, as no generally agreed principles for the distribution of goods exists, there can be no moral case for the free market to answer.

However, as Plant concludes, neo-liberal conservatives, by embracing these three propositions and endorsing the act-utilitarian ethical principle, are conceptualising their primary unit of social transaction — the market transaction — “as happenings outside one’s moral self” (Smart and Williams, 1973: 104). By implication, then, neo-liberal conservatives “should be willing to agree that...[act-utilitarianism’s]...general aim of maximising happiness does not imply that what everyone is doing is just pursuing happiness” (Smart and Williams, 1973: 113). Instead, ostensibly rational action to maximise probable benefit can sometimes be irrational. In this case, it can perpetuate a maleficent outcome, or one that, whilst not intended, could or should have been anticipated, on a particular social group without offering any justification that such a situation is inevitable in bringing the best results for the majority.

Therefore, if neo-liberal conservatives wish to address the moral dilemma of foreseeable, adverse unintended outcomes arising from their actions they could consider the ethical consequentialism developed in Rawls theory of justice (1971), with its aim of ensuring the stability of the state. Rawls recognises that if citizens are to obey the state then a basic scheme for ordering society should include an agreement between those citizens and the state as to how that society would be conducted. These aims require a political consensus over the application of the concept of justice that extends to the details of how the principle can be morally justified. So, to achieve such an understanding, Rawls proposes a hypothetical situation. In this scenario self-interested and rational citizens who are ignorant of the position they would occupy in a future society must choose the highest possible level of income and equality of opportunity for the poorest that is acceptable to all in that society. It is assumed that all participating citizens wish to pursue the greater good and would be prudent enough to realise the need for future social stability. Thus, the outcome would establish not just fairness but the following principles of justice, stated in their order of priority, that underpin the structures of a just society (Rawls, 1971: 320):

“First principle: Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all. Second principle: Social and economic inequalities are to be arranged so that they are both:

(a) To the greatest benefit of the least advantaged, consistent with the just savings principle, and

(b) Attached to offices and positions open to all under conditions of fair equality of opportunity.”

These Rawlsian principles are designed to govern the manner the basic political, economic and social institutions, mould and voluntarily constrain the agent. Therefore, Rawls has constructed an artificial situation where citizens have cooperated with the objective of advancing their self-interest by ensuring the fairness of social outcomes. Thus, this agreement is envisaged as the product of a wide reflective equilibrium that has successfully challenged citizens to confront their values and re-assess their priorities to ensure an effective and equitable meritocracy in their own self-interest.

Therefore, if neo-liberal conservatives dismiss Rawl’s theorising then they must face the apparent failure of market mechanisms to eradicate poverty. Although the trickle-down effect from wealth creation may lift some citizens from absolute deprivation they would continue to suffer relative imbalances in their property rights that leaves them dis-empowered relative to the affluent. This outcome reflects a desire in the marketplace to separate economic reality from social reality giving rise to the mechanical economic machine metaphor. So, instead of
realising some individual’s internalised desires for freedom, the marketplace can present a series of constraints that impel acquiescence to economic rationality (Bourdieu, 1998: 96). It follows that “adaptation becomes the highest goal of character formation” (Beck, 1998: 13) in the free market environment leaving those who adhere to the tenets of consequentialism to ponder whether they should “rejoice in the market economy, but reject the market society” (Plant, 1999: 24).

There is no doubt that the decade of Thatcherism benefited some lower income communities. For instance, those who had paid rent for most of their lives where allowed to buy their council houses and experience the middle class benefits of re-mortgaging to buy a new car or in taking the holiday of a life-time. Similarly, lower income families could now own part of the equity of a privatised utility as previously nationalised industries were floated on the stock exchange and neo-liberal conservative acolytes promulgated dreams of a share owning democracy. Thus, many lower income families were motivated to take advantage of wealth creating opportunities as a means of embracing the virtues of self-responsibility and self-achievement in a paradigm that could lead to the realisation of their own self-worth. Nevertheless, the theoretical drawbacks of consequentialism, in the spectre of market inequalities, haunted this ideological experiment as the trickle-down effect failed to materialise and Rawl’s philosophical vision remained strictly theoretical against a backdrop of a UK market driven economy suffering from high levels of structural unemployment. Therefore, the application of a dose of ethical consequentialism had rendered lower income communities bereft of the certainties that were inherent in the ethereal bond of deontological ethics. This state of affairs could have contributed to the nullification of the traditional working class virtues of patriotism, obedience and compliance and the creation of a code of uninspiring ethical scepticism where apathy and disillusion are paramount.

The Ethical Dilemma of New Public Management

The election of the new political elite, branded as new labour, in 1997 did not herald a return to deontological ethical premises but instead signalled the arrival of a hybrid organisational form of new public management that sought to combine elements of neo-liberal conservatism with bureaucratic structures. This synergism results in the state employing a diverse combination of organisations to deliver public services. These encompass traditional centralised provision to devolved provision at both local and regional level and employ a variety of organisational forms ranging from corporatised and commercialised quasi-public to private-for-profit and private-non-profit (Dixon and Dogan, 2002). This complexity reflects the desire of those formulating policy to impose managerialist values and practices on service providers resulting in a counter-productive paradoxical environment (Dixon et al., 2005). In this scenario the most problematic paradox is:

that public managers are expected to manage “efficiently” and “effectively”, and so be accountable for the efficient and effective management of “inputs” used to produce “outputs” (which may be difficult to quantify, or even adequately conceptualise), which generate “outcomes” (which may be difficult to measure, or even adequately conceptualise), which relate to “programme objectives” (which may be difficult to articulate in mutually compatible and quantifiable terms), which must be compatible not only with “policy objectives” (which government may be unwilling or unable to articulate in quantifiable terms, and which may, themselves, be mutually incompatible, particularly in a multi-level political structure) but also with “customer objectives” (which may, also, be mutually incompatible).

This ambiguous public arena is offered by new labour’s proselytisers as a site where civil renewal and active citizenship can take place, facilitated by a plethora of performance management targets, which are to provide a recipe for efficient, effective and economic outcomes that can be contrasted to the failed neo-liberal conservative project. However, this vision is overly optimistic as this scenario seems to offer citizens, who are dependent on equitable public policies, an ethical framework that is too weak to sustain the re-discovery of the virtuous notion of responsibility towards the state. Instead, the question arises as to
whether the exhortations of policy makers for lower income communities to participate in altruistic involvement for the delivery of public services and the governance of community affairs is fundamentally misplaced in an overly confident managerialist rhetoric (Active Communities Directorate, 2004; Blunkett, 2003, 2004; Chanan, 2003; Civil Renewal Unit, 2003, 2004, 2005; Home Office Research, 2003; 2004a, b; ODPM 2005a, b, c; Rodgers and Robinson 2005).

Furthermore, governance by performance objectives, geared to efficiency and effectiveness, throws up the challenge of how desired “outputs” and “outcomes” are achievable without overt government intervention. So, the ethically sceptical citizen, having been deprived of the deontological bond of trust with the state after suffering what could be regarded as the unprincipled nature of market transactions might have their sceptical assumptions re-enforced as the regulatory state exerts its political authority to achieve its aims and objectives. This authority perceives “society as comprised of a web of obligations, which may override individual freedom: obligations amongst individuals in communities and between the citizen and the state” (Driver and Martell, 1998: 169). Thus, the state can demand that the individual fulfil their civil duties without offering reciprocal obligations in return.

Conclusion

In its traditional role, the British State was informed by deontological ethical principles that decreed if the state exercised extensive power to control the lives of its citizens this power should be exercised in conjunction with benevolent paternalism. Thus, an ethereal bond existed that fostered the virtues of paternalism, obedience and compliance in lower income communities as an elite accepted responsibility for the state’s decision–making and action–taking processes. However the neo-liberal conservative project, underpinned by consequentialist ethical principles, was committed to the state refraining from interfering in the lives of citizens as they pursued their own legal pleasures. Implicit within this paradigm is the precept that the cost of government administration should be minimised by use of the mechanisms of the free market, which are more effective and efficient than centralised bureaucracies. However some citizens, in particular those on a low income, experienced the outcome of market transactions that were adverse or difficult to comprehend. Thus, it is contended that ethical scepticism grew alongside distrust of market solutions for the delivery of essential public services.

The second contention is that new labour’s political elite, with their belief in the managerialised, regulatory and ethically hybrid state, have been unable to reverse the growth of ethical scepticism despite their concern to promote active citizenship and civil renewal. This may reflect new labour’s constitutional radicalism, which places an emphasis on individual responsibilities rather than individual rights.

Therefore, it is appropriate that a debate should take place over what constitutes the public good. This should address the importance that society wishes to accord to the values of equity, distributional justice, community solidarity and social stability in the regulatory provision of public services. John Stuart Mill captures the risks inherent in neglecting this dialogue:

a State which dwarfs its men, in order that they may be more docile instruments in its hands even for beneficial purposes – would find that with small men no great thing can really be accomplished (Mill, [1859] 1989: 15).

Bibliography


Office of the Deputy Prime Minister (2005c) *Vibrant local leadership*, London: HMSO.


