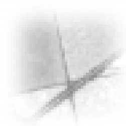


Southern African Business Review
Volume 8 Number 3
December 2004

Contents

Research articles

The contribution of 'familiness' to the performance of family businesses
W.P. Venter & S. Kruger 1

Predicting turnover behaviour of direct salespeople
P. Msweli-Mbanga 14

Consumer spending patterns in the informal retail trade sector of South Africa
D.H. Tustin 26

Evaluation of the need to introduce a system of group taxation in South Africa
J.S. Wilcocks & S.N. Middelmann 38

Quality assurance in internal auditing: an essential tool in ensuring that the expectations of the users of internal auditing services are met
M. Marais 54

The tax deductibility of donations, with specific reference to donations of property made in kind to public benefit organisations
R. Oberholzer 68

Guidelines for contributors to the *Southern African Business Review* 76

The *Southern African Business Review* is a publication of the College of Economic and Management Sciences, University of South Africa.

The contribution of 'familiness' to the performance of family businesses

W.P. Venter & S. Kruger[*]

While family businesses are known to consistently outperform non-family businesses in financial terms over the long run, family businesses have received comparatively little attention from researchers. In this article, an explanation is offered for this superior performance in the form of the concept of 'familiness' – the unique contribution that family involvement brings to any business (which is divided into founder capital and family capital). It is explained that family businesses possess no general competitive advantages over non-family businesses. The unique strength of successful family businesses does not lie in their espoused advantages, but in their ability to sustain and adapt, through family capital, the culture created by the founder. An evolutionary conceptual model of the creation and transmission of familiness is provided to explain how this unique strength influences family business performance over the long run.

Introduction

Even though the family business is the most prevalent form of business throughout the world (Upton & Petty 2000: 28; Kim, Kandemir & Cavusgil 2003: 1–2; Heck, Upton, Bellet, Dunn & Parady 1994: 2), it has received comparatively little attention from researchers (Litz 1997: 55). Family businesses are even more important in emerging market economies (Gibson 2002: 68; Kimet et al. 2003: 5). In South Africa it is estimated that 84% of all businesses in the formal sector are family-owned, in other words, a total of close to 1.2 million businesses. Of this total, approximately 330 000 are companies or close corporations, and 870 000 are sole proprietors (Balshaw 2003: 5).

The conventional wisdom about family business refers to the perceived lack of endurance of family businesses. The statistics that in only one out of three cases does the successive generation continue with the family business, and that the average lifespan of a family business is 24 years, have been repeated in articles on family businesses since they were first published by Beckhard & Dyer (1983: 5). Studies that contradict the conventional wisdom are less known and less frequently quoted. It has been shown in various countries that family businesses last longer (Westhead & Cowling 1998: 44; Wall 1998: 24; Bhattacharya & Ravikumar 2001: 187), perform better (Ryan 1995: 12; Anderson & Reeb 2003: 13–21; McConaughy, Matthews & Fialko 1997: 7–9; Neubauer & Lank 1998: 11–12) and are more resilient (Church 1993: 17; Jones & Rose 1993: 1) when compared with non-family businesses.

Early research into the family business field (for example, Donnelley 1964: 93–105; Levinson1971: 90–98; Davis 1983: 47–56) stressed how to overcome the negative side of family business and so set the research agenda for future family business research. As a result, the perception that family businesses lack endurance has dominated research. Given this negative outlook, the main question family business research addressed was: What is wrong with family businesses, and what can be done to fix it? An alternative area of interest was: How can family businesses avoid failure and survive longer? As Kepner (1983: 69) pointed out, the family business field is occupied with the wrong questions; questions that perpetuate a mindset focused on eliminating the negative rather than developing the positive.

As a result of the negative perspective of most family business research, it comes as no surprise that the field has been preoccupied with issues surrounding preservation. Hoy & Verser (1994: 10) found that research is predominantly focused on succession, governance and survival issues. A review by Dyer & Sanchez (1998: 287–295) of

* Dr W.P. Venter is Executive Chairman of the Altron Group. E-mail: wpventer@altron.co.za. Professor S. Kruger is Professor of Business Management at the Rand Afrikaans University.

articles appearing in the *Family Business Review* since its inception in 1988 came to similar conclusions.

With the emphasis on preservation dominating academic research, few studies have addressed the development of the positive perspective of family business. In many cases, studies that do emphasise the positive aspects of family business rarely go beyond a list of advantages (Habbershon & Williams 1999: 37–39). The purpose of this article is therefore to again draw attention to the superior performance of family businesses relative to non-family businesses, and to find the underlying reasons for this superior performance.

This research presents a conceptual model of the contribution of family involvement to family business performance.

Defining family business

A significant problem in the family business field is the lack of definitional clarity on the central concept of family business. It may not be much of an exaggeration to state that there have been as many different definitions of family business as there have been authors on the subject. From Chua, Chrisman & Sharma's (1999: 25) review of the important definitions of 'family business' offered since 1964, it is clear that most definitions have tended to focus mainly on the ownership and management dimensions.

Most authors (for example, Barnes & Hershon 1976: 106; Donckels & Frohlich 1991: 152; Gallo & Sveen 1991: 181) believe that family businesses are defined as those where a single family has controlling ownership (50% of the voting shares or more) as well as management control. Since this definition is easy to operationalise, it is also most often used in empirical studies of family business. This definition, however, seems to include businesses owned by a single person and businesses where the family has control but has little interest in the affairs of the business or in continuing its involvement in the long run, while it excludes businesses where a family might have a significant influence over the direction of the firm but little formal control.

To clarify, some authors (for example, Donnelley 1964: 94; Churchill & Hatten 1987: 52; Ward 1987: 252) have added further dimensions to the definition, most commonly the involvement of more than one generation in the business and the perception of its being a family business. This has led to increasingly complex definitions, as exemplified by the use by Westhead & Cowling (1999: 40) of a four-level hierarchy to define family businesses and focus empirical research. Attempts to capture the essence of family businesses came to resemble a degenerating research programme, with add-ons regularly being suggested to cover some previously neglected aspect of the family business.

The family business typology of Litz (1997: 57) suggests that the essence of family businesses is better captured by the behavioural dimension, in other words, businesses are family businesses because they behave like family businesses. Some empirical studies pointed out the need for a behavioural perspective, since it is possible for firms to be professionally managed and still exhibit the characteristics of family businesses (Daily & Dollinger 1993: 88). The empirical study of Chua et al. (1999: 28–35) confirmed that the behavioural dimension explains family business behaviour significantly better than any of the definitions that include easily operationalised variables such as ownership and management control.

The definition offered by Chua et al. (1999: 25) appears to come closest to incorporating the behavioural dimension and will therefore be used: A family business is a business governed and managed with the intention of shaping and pursuing the vision of the business held by a dominant coalition controlled by members of the same family in a manner that is potentially sustainable across generations.

Family business performance

In this section, it is shown that family businesses consistently outperform non-family businesses in financial terms. The concept of 'familiness' is offered as an explanation for this superior performance.

Comparative studies

Various empirical studies (for example, Anderson & Reeb 2002: 1–37; Anderson & Reeb 2003: 1–39; McConaughy et al. 1997: 1–17; McConaughy, Walker, Henderson & Mishra 1998: 1–19; Vilaseca 1995: 1–13) have shown that family businesses tend to approach the pursuit of financial performance differently from non-family businesses and consistently outperform non-family businesses according to financial measures.

McConaughy et al. (1997: 8–11) found that family businesses on the Standard & Poor's 500 index

(S&P 500) tend to outperform non-family business according to the measures of market to book equity ratio, stock market return, sales growth, gross and net margins on sales, cash flow per employee, and working capital per unit of sales. McConaughy et al. (1998: 1) conclude that family businesses tend to be more efficient than non-family businesses. Anderson & Reeb (2003: 17–18) found that family businesses on the S&P 500 tend to do better than non-family owned businesses in terms of Tobin's Q-ratio, return on assets, and earnings before interest, taxes, depreciation and amortisation (EBITDA), especially if founders or founder descendants occupy the position of CEO. Anderson & Reeb (2003: 37) confirm that family businesses tend to be better at value maximisation than non-family businesses on the S&P 500. While family businesses show no significant difference in risk-taking, they are more valuable and have a higher economic value added (EVA) than non-family businesses.

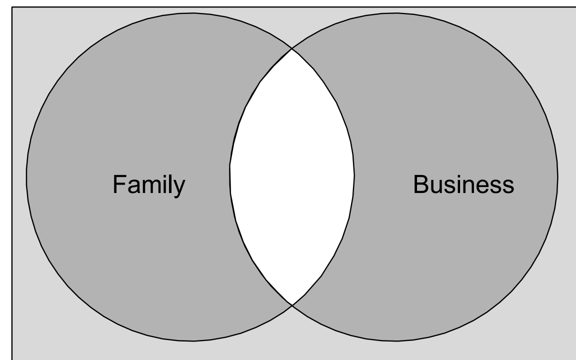
The idea that family businesses, particularly family business groups, are dysfunctional or in crisis, is soundly and consistently debunked by empirical studies. While no broad comparative studies have been done in South Africa, the situation is likely to be similar, according to Ryan (1995: 12) – family businesses listed on the JSE Securities Exchange delivered a 36% return on equity from 1987 to 1992, as opposed to an average 27% return in the non-family industrial sector over the same period.

Orthodox explanations

Orthodox family business research sees family businesses as comprising separable dimensions (namely family and business). One of the earliest depictions of this so-called overlap model of family business was given by Lansberg (1983: 44), as illustrated in Figure 1, and further refined by the developmental models of Gersick, Davis, Hampton & Lansberg (1997: 17) and Carlock & Ward (2001: 27).

The explanation that such overlap models suggest for family business performance is that family businesses were able to separate the family dimension from the business dimension, thereby avoiding the destructive conflicts to which family businesses are prone.

Overlap models create a perception that family and business can and should be separated. This view can be traced back to the good – bad approach, in terms of which family influence is regarded as something that is to be tolerated in a business, and



Source: Lansberg (1983: 44)

Figure 1: The two-circle overlap model

not necessarily crucial to the performance of the family business. It is known, however, that family businesses tend to perform better than non-family businesses. If family businesses outperform non-family businesses, it is because they have some kind of sustained and unique advantage that non-family businesses cannot easily imitate. The better performance can only be due to the single differentiator between family businesses and non-family businesses – a family exercising strong influence as a dominant coalition. Keeping family separate from business is therefore harmful, as it attempts to extract the one thing that gives a family business its advantage over its non-family business rivals. Stafford, Duncan, Dane & Winter (1999: 206) express this succinctly: ''... it is not the business that makes a family business unique from other business arrangements, rather, it is the family''. In other words, a family business that is able to extract and separate the family element from the business will lose the one element that makes family businesses unique and allows them to outperform non-family businesses.

Family influence is the one thing that is unique to family businesses, and could be regarded as a resource to a business. Family influence as a resource is referred to as 'familiness'. Familiness is the unique bundle of resources a firm has as the result of the interaction of the family, the firm and individual family members with one another (Habbershon & Williams 2001: 18). Familiness is regarded as a capability, in the sense that it is firm-specific, is embedded in the firm and its processes, and is not transferable to other firms.

Familiness originates with a founder, so the founder as a resource to the business (founder capital) needs to be considered before investigating the

family as a resource (family capital). In a successful family business, founder capital and family capital interact to create superior business performance.

Founder capital

The combination of the fact that family businesses outperform non-family businesses, with the finding of Chandler & Hanks (1994: 84) that founder competence positively influences firm performance, confirms that the founder can be regarded as a significant resource to the firm. For this reason, this article introduces a new concept to the family business literature – founder capital. Founder capital can be seen as an important part of a firm's human capital.

The concept of founder capital is inspired by Erikson's (2002: 275–278) concept of entrepreneurial capital. The concepts of entrepreneurial capital and founder capital are related, since one of the founder's main roles is to act as innovator and entrepreneur (Chandler & Hanks 1994: 78), and therefore the founder's entrepreneurial competence directly enhances a firm's performance (Chandler & Hanks 1994: 84; Liang & Jek 2000: 15).

According to Erikson (2002: 275), entrepreneurial capital is a concept akin to social capital. He defines entrepreneurial capital as the present value of future entrepreneurial behaviour, or the present value of the infinite series of shadow options that exist as a result of the entrepreneur's involvement with a firm. Entrepreneurial capital is used over time as the shadow options are converted into real options. The more the firm draws on its entrepreneurial capital, the more entrepreneurial capital becomes available, and the firm's performance improves (Erikson 2002: 279).

Erikson (2002: 278) empirically demonstrates that entrepreneurial capital is a multiplicative function of entrepreneurial competence and entrepreneurial commitment. The founder of a successful family business is likely to have both in large doses, especially commitment. Erikson (2002: 282) states that commitment is only meaningful if a vision, and subsequently a goal, exists and that the committed entrepreneur has a strong intention to achieve this goal and invests large amounts of physical, emotional and intellectual energy in pursuit of this goal. In the case of the founder of a family business, this goal is a vision of success "that is potentially sustainable across generations" (Chua et al. 1999: 25). This vision is likely to go beyond financial wealth (although it may often include family wealth creation) and extend to a vision that will satisfy the founder's need for generativity.

Entrepreneurial capital is a more narrow concept than founder capital, because it does not include the founder's other main role (according to Chandler & Hanks 1994: 78–79), namely, to act as a professional manager. By broadening the concept of entrepreneurial capital, founder capital can be defined as the present value of the founder's future entrepreneurial and managerial behaviour.

There can therefore be no doubt that founders are one of the most important resources of a family business in its early years (Costa & Gubitta 2002: 3). This can, however, easily work in the opposite direction. A founder who remains virtually omnipotent and refuses to make crucial changes, or to recognise great opportunities, can become a firm's greatest obstacle and in many cases cause a firm's failure.

However, this resistance to change is less likely to happen while the founder is still with the family business. Boeker (1989: 510) found that strategic inertia and lock-in are less likely to take place while the founder is still around. His study delivered a surprising result: the longer the tenure of the founder, the more likely it is that strategic change will occur. The reason for this seems to be that, firstly, founders have the authority to change the practices and principles they impose, and secondly, founders understand the spirit of the principles and practices and are less likely to follow them to the letter.

Founder legacy in a family business refers to the extent to which succeeding generations refer to and are imbued with the founder's vision and management principles when making strategic decisions (Kelly, Athanassiou & Crittenden 2000: 39). When the founder leaves the firm, founder descendants will invariably continue within the culture the founder created. If the founder was not able to communicate the spirit behind the principles and practices he or she imposed, founder descendants in top management positions will most likely regard them as prescribed law and follow them to the letter. In such a case, the founder's legacy inevitably constrains the business, causing it to become less responsive to change and harming business performance. Ogbonna & Harris (2001: 25) refer to this state of affairs as founder legacy 'hangover'. However, if a founder's descendants focus more on the process and less on the content behind the founder's legacy, then the founder's legacy will

become an 'inheritance' and will enable family members to introduce the cultural adaptations necessary to continue the superior performance of the business (Ogbonna & Harris 2001: 25).

Family capital

Founder capital continues in the family business through the founder's descendants. If the founder left a legacy in the form of an 'inheritance', then the founder's principles, beliefs, practices and general business philosophy are left open to interpretation, and will most likely change as the next generation exerts its influence (Ogbonna & Harris 2001: 25–26). The founder's descendants will absorb the founder's legacy and filter it through their own minds, personalities and view of the business environment as they see it at the time.

A new kind of resource will have to be introduced to refer to the unique resource that the founder's descendants bring to the business, namely, family capital. It is also quite possible that every generation will bring a different brand of family capital to the family business. This article therefore introduces a further refinement to the literature: the concept of generational family capital.

Family capital is a relatively young concept in family business literature, having been introduced only recently by Karakoulaki (2002: 2–3) and Hoelscher (2002: 6–9). Both these authors see family capital and success in business as closely related. Karakoulaki (2002: 7) regards family capital as a necessary condition for social capital to exist, while Hoelscher (2002: 3–4) regards family capital as a more intense and immediately available form of social capital.

Karakoulaki (2002: 2–3, 5) defines social capital as the relationships, networks, contacts, alliances, partnerships, culture, social practices, conventions and institutions within society that generate trust, establish expectations, create and enforce norms, and facilitate collective action. Karakoulaki (2002: 2) similarly defines family capital as the relationships, networks, alliances, partnerships and other non-economic factors within the family that affect the entrepreneurial activities of the family and define a framework for its economic actions.

Hoelscher (2002: 13) sees family capital as a distinctive core competency of the family business that is not available to non-family businesses. While social capital is available to non-family businesses and family businesses alike, family businesses have access to the more intense and more immediate form of social capital, namely family capital (Hoelscher 2002: 34).

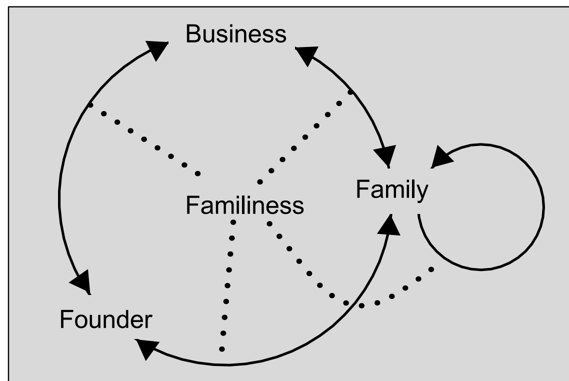
Like other resources, family capital can be eroded. Family capital is usually eroded by anything that makes members of the family business network less interdependent. Interdependence is enhanced by trust, respect, reputation, closure (the completeness of the network) and reciprocity (Hoelscher 2002: 28–29). Moreover, like any other resource, family capital can depreciate (but at an unpredictable rate) and can complement or substitute other resources in the business (Hoelscher 2002: 32).

Hoelscher (2002: 72) found a positive empirical relationship between the amount of family capital and the performance of a family business, as one would expect. Hoelscher (2002: 47–48) specifically tested for the effect that task conflict and relational conflict have on family business performance. He found that task conflict has a potentially positive effect and relational conflict a potentially negative effect on family business performance, but only when there is little family capital present. However, Hoelscher (2002: 77, 81–82) found that if high levels of family capital are present in a family business, then the degree of task or relational conflict has little effect on business performance. The lower the level of family capital, the more sensitive to conflict the performance of a family business becomes.

One can draw several conclusions from Hoelscher's (2002: 47–90) results. It now becomes even more obvious that attempting to separate family and business is likely to make things worse rather than better. A family business with high levels of family capital is resilient in the face of the inevitable relational conflict that occurs in a family business. Family capital is a resource that needs to be nurtured. Instead of separating it from the business so that it falls into disuse and loses its function, it should therefore be strengthened and further integrated into the family business. While family business literature since Donnelley (1964: 93–105) has focused on how to manage and constrain the conflicts that occur in a family business, Hoelscher's (2002: 90) results suggest that this is a misdirection of effort. As family capital compensates for low levels of task conflict and high levels of relational conflict, family businesses will make much better use of their limited resources if they apply them to nurturing family capital than to constantly attempting to contain and manage conflict.

Familiness

The understanding of familiness presented in this article is illustrated in Figure 2, which shows that there is reciprocal influence between the founder, the family and the business. Family members of various generations reciprocally influence one another and the business. Familiness is created by the interactions between the founder (or founder legacy), family members, generations of the family, and the business.



Source: Adapted from Habbershon & Wiliams (1999: 17)

Figure 2: The creation of familiness

It would be wrong to think that familiness is always a positive influence in a family business. Familiness, if not maintained and nurtured, can rapidly become a destructive force in a business (as the popular press is quick to point out, usually under the guise of sibling rivalry). For this reason, Habbershon & Williams (1999: 20) distinguish between distinctive and constrictive familiness. Constrictive familiness develops when founder and family capital are eroded and family involvement becomes an encumbrance to the family business. Distinctive familiness exists when family involvement in a family business provides a firm with a sustainable competitive advantage.

The model

In this section, the effect of familiness is explained further and then expanded into an evolutionary model of family business. The evolutionary model is then extended through a model of the transmission of familiness in every stage of the evolutionary process.

The real effect of familiness

The familiness of a business, which results from the interactions between and within the family and the business, cannot be separated from its corporate culture. Corporate culture can be defined as the values, beliefs and attitudes that influence individual and group behaviour within a business organisation (Miller 2000: 22). Barney's (1986: 657) definition also adds assumptions and symbols as elements of corporate culture. Familiness overlaps with corporate culture of a family business, as the values, beliefs, assumptions and attitudes of the founder and the founder's descendants are absorbed in the corporate culture and influence the way things are done in the business. When culture is transmitted, familiness will then automatically be transmitted as well.

Culture (and the familiness absorbed in it) affects business performance, but for culture to have a positive influence, it needs to be valuable, rare and inimitable (Barney 1986: 658). If cultures that enable success are rare and inimitable, then successful family businesses cannot be said to possess some kind of general 'family business culture'. The same is true of any variable, or combination of variables, that has the potential to create sustained competitive advantage, since it has to be unique to provide an advantage. If family businesses possessed some kind of general family business competitive advantage, they would not have something unique and would thus not exhibit superior performance relative to other firms. Trying to find the general competitive advantage of family business is inappropriate and doomed to failure (Habbershon & Williams 1999: 10; Jones & Rose 1993: 4).

All this suggests that efforts to find the secret of family business success can easily be misguided. One will not find the secret by trying to establish commonalities in corporate culture. Furthermore, since culture decisively influences business performance (Barney 1986: 659), it would be more insightful to investigate how the rare, inimitable and valuable culture that enables superior business performance is created. It is a difficult task to capture the nature of a rare and inimitable culture or explain how such a culture came into being. Barney (1986: 660) attempts an explanation when he states that unique cultures almost invariably have their origin in the unique personality of the founder or the unique circumstances of an organisation's founding or growth.

A business cannot choose the personality of its founder or the founding conditions that create its corporate culture. Both successful family businesses and successful non-family businesses will have unique founders and founding conditions. If

founders and founding conditions are unique to each successful business, this approach is also unlikely to provide any common factor that distinguishes family businesses and explains their superior performance. There are therefore no clear commonalities in family business in the way culture is created, just as there are no commonalities in their corporate cultures as such.

If it is futile to attempt to find the secret of the success of family businesses in their actual cultures or the factors that create their cultures, the question still remains: What allows successful family businesses to consistently, and over the long-term, outperform successful non-family businesses? The central argument of this article is that the crucial difference lies not in commonalities of culture or in how a unique and valuable culture is created, but rather in how a proven culture is sustained, nurtured and adapted to changing circumstances.

As stated before, successful businesses achieve success in large part as a result of the initial and sustained strategic leadership of their founders. In many businesses, this founder capital continues to influence the business even long after the founder has left. The businesses where this sustained influence is most likely to be found are family businesses, where founder capital is transmitted to, and absorbed by, successive generations of the founder's family.

Non-family businesses may still be influenced by the founder's legacy after the founder leaves, but this legacy may become so rigid that it harms the business (for example, in the case of IBM) or dissolves over time (as new employees enter the business and parts of the business are sold, for example, Premier Milling in South Africa). In successful family businesses, the family will have absorbed not the letter, but rather the spirit, of the founder's legacy. They adapt the legacy to circumstances by sustaining the positive and discontinuing the negative aspects of the founder's legacy. This is what makes successful family businesses unique – their ability to sustain and adapt the culture that created the success of the business in such a way that the success is maintained. The combinations of advantages and culture differ from family business to family business, but this almost perfect transmission of a successful founder's spirit is the one thing successful family businesses have in common.

Non-family members can also sustain the founder's legacy, but because of a less than total identification with the business and difficulties in absorbing the tacit knowledge and skills of the founder, this happens with much greater difficulty. Non-family members are more likely to leave in times of severe adversity or when better opportunities arise and are less able to grasp the unspoken and articulated understanding of the founder, and thus an imperfect channel for the transmission of founder capital and family capital exists. While family businesses tend to outperform non-family businesses on average, there is nothing automatic about this superiority. Non-family businesses can outperform even the best family businesses. If there were a more powerful way than the family to transmit and adapt the spirit of a visionary and capable founder, the tables would probably have been turned, with non-family businesses performing better and exhibiting greater resilience in the face of adversity.

There are an infinite number of unique cultures that can provide competitive advantage to a business, and hence an infinite number of ways to turn unique cultures into superior business performance. Investigating how this happens will not yield the insights needed to determine the reasons why successful family businesses deliver superior financial performance. It is more useful to focus on that which enabled the business to create a unique culture, namely its founder and history, and that which causes the uniqueness to be sustained and adapted. As a result, a new proposition to provide a different focus to the study of successful family businesses is derived. The unique strength of successful family businesses does not lie in their espoused advantages, but in their ability to sustain and adapt, through family capital, the culture created by the founder.

Stages of development

In order to understand how familiness influences a family business over the long run, an evolutionary model of the family business system will be presented. This evolutionary model will then form the basis for a more elaborate explanation of how familiness is transmitted, sustained and adapted in a family business.

The family business will evolve over time and start as a much simpler version of Figure 2. A possible evolution of the family business system over time is shown in Figure 3.

Stage one in Figure 3 represents the start-up phase of a family business. At this stage, the founder's descendants are young, and the founder is more concerned with the firm's survival than generational continuity. The founder influences the rest of the

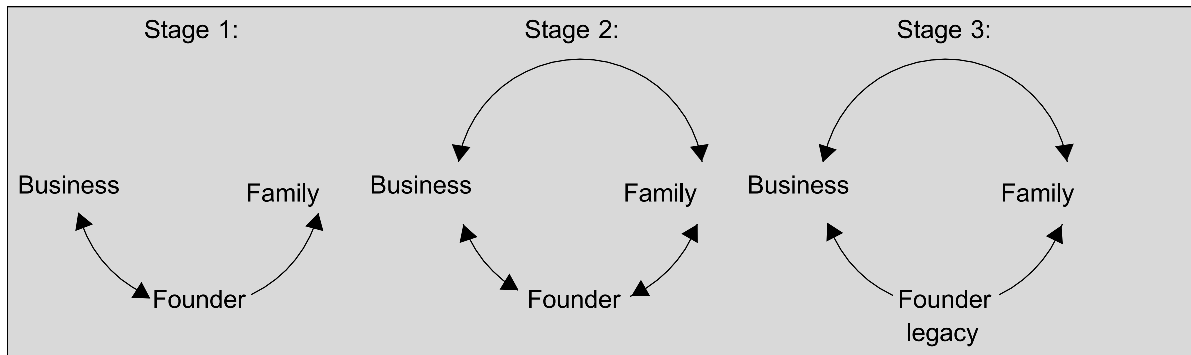


Figure 3: The evolution of the family business system

family, and there is a reciprocal influence between the founder and the business. In this stage, the interactions create only founder capital. If founder capital is distinctive, the family business will grow and become so successful that the founder's descendants also join the business. In stage two, all elements have a reciprocal influence on one another, and family capital is created and combined with founder capital. If the family capital is distinctive and aligned with founder capital, the family business will be ready for stage three. In this stage, the founder leaves the business, but the founder's legacy continues to influence both the family and the business. If the founder's legacy is an 'inheritance', the family business will continue for another generation. From then on, the challenge facing the family business is to create strong family capital across generations and to ensure that the family capital of the generations involved in the business remains aligned.

Transmission models

Given the proposition that the appropriate focus of the study of family business success is the investigation of how the culture created by the founder is sustained and adapted, this section will expand on how this might happen. Based on the evolution of family business presented in Figure 3, the analysis is broken up into three stages. At each stage, familiness is created, sustained and adapted through various channels. In this section, the basic frameworks are presented and discussed, while in the next section, the various channels are explained in more detail.

The aim of this discussion of the three stages is not to identify every single relationship and influence that exists between the founder, the business, the family and the environment. While realistic, such an approach would be too complex and would not contribute significantly to the aim of this section, which is to identify the main channels of creation, sustenance and adaptation of familiness.

Stage one begins when the founder establishes a business, and this stage is represented in Figure 4. The founder's actions are not only determined by his or her own personality and inclinations, but also by the particular environmental conditions at that stage. The environment also contains the external stakeholders of the business with whom the founder continually interacts.

The founder creates a culture (based on his or her 'theory of success'), which is transmitted to the rest of the business (channel 1) through various primary and secondary mechanisms explained by Schein (1983: 22). Once this initial culture is firmly established, it is transmitted to new employees (channel 2) through socialisation, recruitment and selection, anchoring values, and rites and rituals (Wiener 1988: 542; Harrison & Caroll 1991: 554). The founder also influences the values, beliefs and actions of the top management team (channel 3) by establishing his or her centrality in the top management team (Kelly, Athannasiou & Crittenden 2000: 27–40). The members of the top management team filter what they receive from the founder through their own mental models and indirectly transmit this to the rest of the business (channel 4) through the strategic choices they make (Hambrick & Mason 1984: 195–198). Since top management members are influenced by the culture and transmit the culture themselves, separating them from the culture is somewhat artificial. However, the distinction is maintained for the sake of analytical clarity, as top managers are unique from other family business employees in the sense that they are in closer and more regular contact with the founder, and are thus a distinctive transmission channel. The founder may also influence the business directly through what Rowe (2001: 81) refers to as strategic leadership – namely, bringing both managerial and visionary

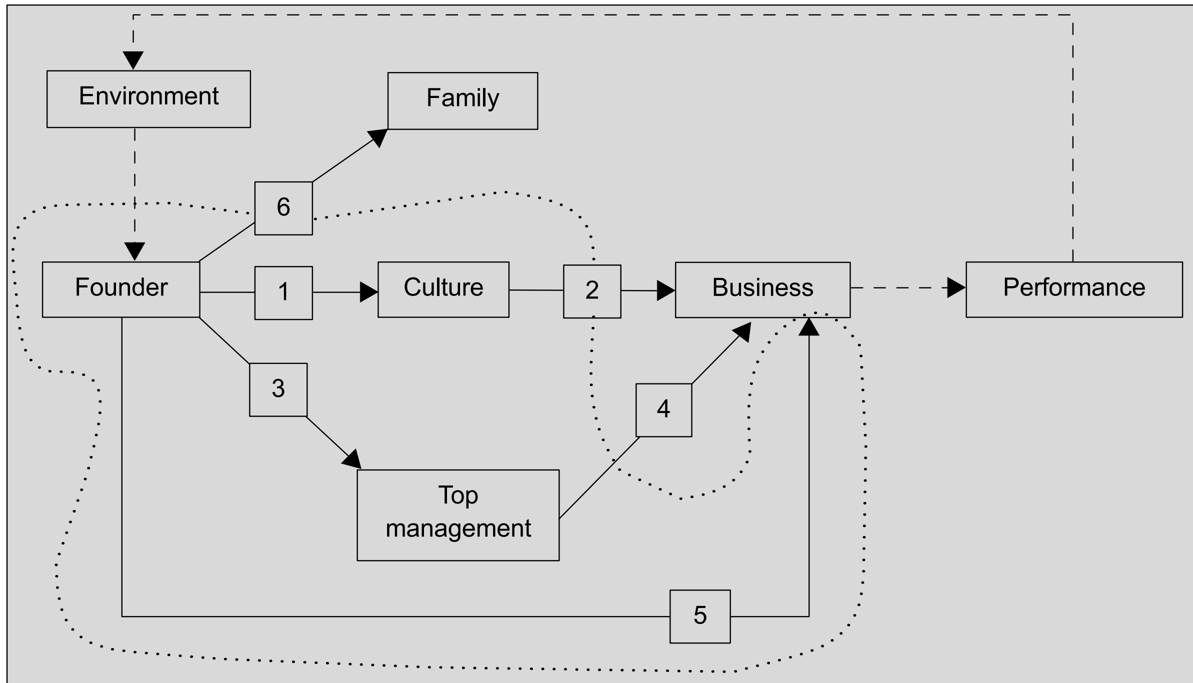


Figure 4: Creation and transmission of familiness in stage one

leadership directly to bear on the business (channel 5). Even though founder descendants are often too young to join the business at this stage, the founder already exerts his or her influence on them (channel 6), and shapes their identities, values, skills and behaviours through parental influence (Lansberg 1999: 78; Narva 2001: 3). In Figure 4, everything inside the area bounded by the dotted line represents the processes through which the founder introduces founder capital to the business.

In Figure 5, the channels of the second stage are explained. The processes that lead to the creation of founder capital (as indicated in Figure 4) are now simplified and encapsulated in one box. Many of the relationships from stage one still exist, but are either included in the concept of founder capital, or omitted in order to focus on the pertinent processes involved as the founder's descendants become an influential force in the family business.

What happens in this stage determines whether the business advances from being a founder-controlled

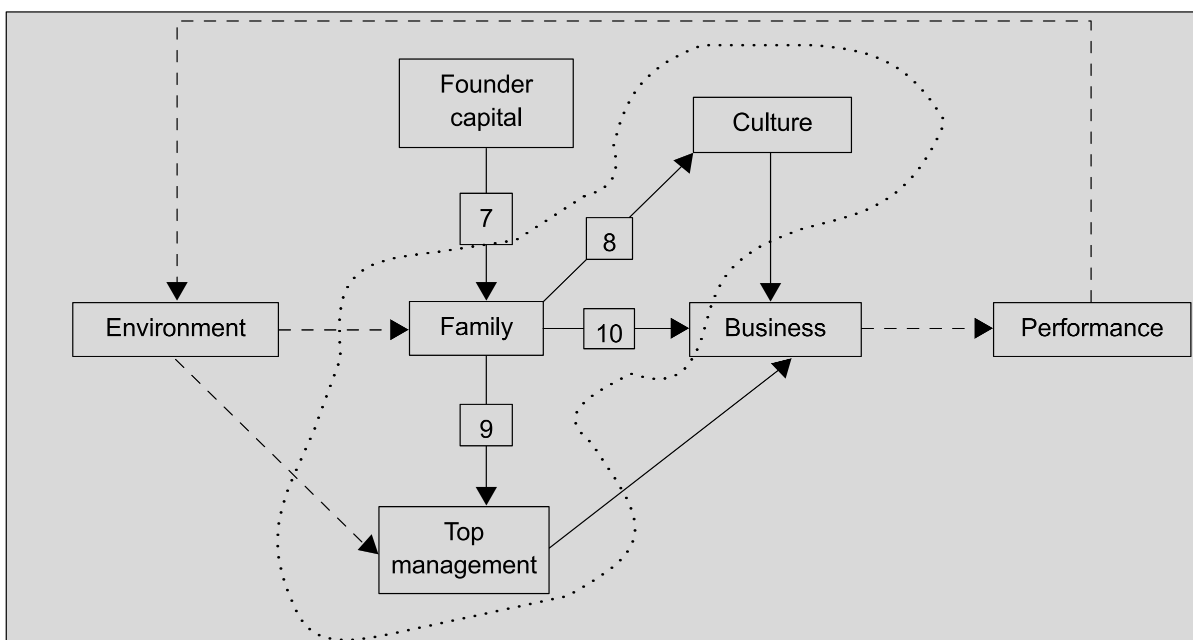


Figure 5: Creation and tansmission of familiness in stage two

business to a family business. The founder’s descendants (abbreviated as ‘family’) join the business in stage two. While employed, they learn from and are influenced by the founder (channel 7), transmitting his or her vision directly and indirectly (Fiegener, Brown, Prince & File 1996: 16–21). The family members absorb the founder’s influence, filter it through their mental models, and transmit this to the corporate culture (channel 8), top management (channel 9) and those parts of the business with which they are involved (channel 10) in ways that are usually similar to those of the founder. The founder’s descendants will probably find it more difficult to establish their authority and centrality in the business (Narva 2001: 10), and are likely to formalise and adapt some aspects of the culture. Top management is changed by the influence of the family and filters and transmits this to the rest of the business (as in stage one). The family’s values and actions are also influenced by the environment (namely, conditions and interactions with external stakeholders). In Figure 5, everything inside the area bounded by the dotted line represents the processes through which the founder’s descendants introduce family capital to the business.

It is important to reiterate that, as with Figures 4 and 5, Figure 6 does not indicate all possible influences and relationships. Some relationships, for example, the impact of the environment on the business in general, have been omitted in order to focus the attention on those relationships that are involved with the transmission of familiness. In Figure 6, the processes that lead to the creation of family capital are now simplified and encapsulated in one box. The focus in Figure 6 is the channels involved in the third stage of evolution, when the founder leaves the family business.

Once the founder leaves the business, founder capital remains in the business, but in a changed form as founder legacy. Founder legacy is absorbed in the culture of the business and continues to influence subsequent generations of the founder’s descendants (channel 11). The effect of founder legacy depends on whether it exists as a ‘hangover’ or an ‘inheritance’ (Ogbonna & Harris 2001: 25). Older generations of the family influence younger generations (channel 12). Younger generations filter and absorb this influence, and transmit this influence to the culture and the business in general.

Figures 4 to 6 explain how familiness is created, transmitted and adapted at each of the three stages of the evolution of a family business. The figures do not exhaust all the possible interactions between the founder, the family and the business, but identify the most important relationships.

Conclusion

A major implication of this research is that its theoretical contributions unify the various areas of research in the field of family business. The research of Dyer & Sanchez (1998: 287–295) indicates that most family business research tends to focus on the areas of succession, governance and survival issues. The conceptual models presented here offer a way of integrating these three

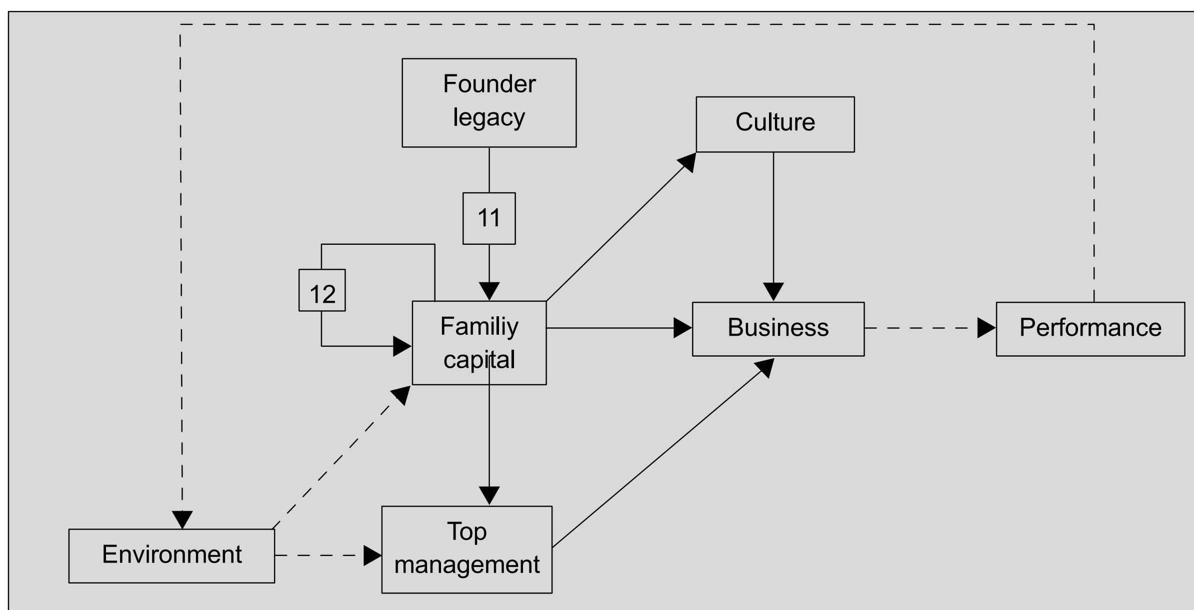


Figure 6: Creation and transmission of familiness in stage three

issues into a broader, coherent and interdependent whole. This coherent whole, in the form of the transmission models, includes areas that are often researched separately – areas such as culture and family psychology. Succession issues are clearly only a subset of the transmission process and fit in with the discussion of channels 6 and 7. Governance issues are also only a part of the broader transmission process and can be classified under channels 4 and 9. Since superior performance implies survival, survival issues are only a subset of the study of superior performance. In any case, an obsession with survival issues is inappropriate when it has been proven repeatedly, and across countries, that family businesses exhibit superior financial performance relative to non-family businesses and tend to outlive non-family businesses.

The transmission models also show how various areas of family business research cannot be separated from one another. Cultural issues (channels 1, 2 and 8), management and governance issues (channels 3, 4, 5 and 9), psychological issues (channels 6 and 10), succession issues (channels 6 and 7) and the role of individuals (channel 5) are shown to influence one another in the transmission models. For example, within the transmission models for stages one and two, it is clear that in order to understand channels 6 and 7 (succession issues), one first needs to understand the founder capital that precedes it (cultural and management issues channels 1 to 5). Conversely, to understand cultural and management issues (channels 8 and 9), one first needs to understand the preceding succession issues (channels 6 and 7). The conceptual models of the transmission of familiness therefore provide a unifying framework within which existing and future research on family business can be placed and the interdependence between different areas of research can be seen.

The transmission models could not operate in any family business if an artificial separation of family and business were attempted. The transmission models show how familiness plays a fundamental role in the long-term performance (good or bad) of a family business.

The conceptual models provide a clear framework for future empirical research of family businesses.

References

Anderson, R.C. & Reeb, D.M. 2003. ‘Founding-family ownership and firm performance: Evidence from the S&P 500’, *Journal of Finance*, 58(3). [Online] Available from: www.afajof.org/Pdf/forthcoming/FF_Ownership.pdf. Accessed: 20 December 2002.

Anderson, R.C. & Reeb, D.M. 2002. Founding-family ownership, corporate diversification and firm leverage. Working Paper. Washington: American University. [Online] Available from: www.cba.ua.edu/~dreeb. Accessed: 20 December 2002.

Balshaw, T. 2003. *Thrive: Making Family Business Work*. Cape Town: Human & Rousseau.

Barnes, L.B. & Hershon, S.A. 1976. ‘Transferring power in the family business’, *Harvard Business Review*, 53(4): 105–114.

Barney, J.B. 1986. ‘Organizational culture: Can it be a source of competitive advantage?’, *Academy of Management Review*, 11(3): 656–665.

Beckhard, R. & Dyer, W.G. 1983. ‘Managing continuity in the family owned business’, *Organizational Dynamics*, 5(1): 5–12.

Bhattacharya, U. & Ravikumar, B. 2001. ‘Capital markets and the evolution of family business’, *Journal of Business*, 74(2): 187–219.

Boeker, W. 1989. ‘Strategic change: The effects of founding and history’, *Academy of Management Review*, 32(3): 489–515.

Carlock, R.S. & Ward, J.L. 2001. *Strategic Planning for the Family Business: Parallel Planning to Unify the Family and Business*. New York: Palgrave.

Chandler, G.N. & Hanks, S.H. 1994. ‘Founder competence, the environment, and venture performance’, *Entrepreneurship Theory and Practice*, 18(3): 77–89.

Chua, J.H., Chrisman, J.J. & Sharma, P. 1999. ‘Defining the family business by behavior’, *Entrepreneurship Theory and Practice*, 23(4): 19–39.

Church, R. 1993. ‘The family firm in industrial capitalism: International perspectives on hypotheses and history’, *Business History*, 35(4): 17–43.

Churchill, N.C. & Hatten, K.J. 1987. ‘Non-market based transfers of wealth and power: A research framework for family businesses’, *American Journal of Small Business Management*, 11(3): 51–64.

Costa, G. & Gubitta, P. 2002. When the firm has the makings of the boss: Values, entrepreneurship, and management to bring up a family firm. Paper presented at the Family Business Network 13th Annual World Conference, 11–14 September 2002, Helsinki. [Online] Available from: www.fbnhelsinki2002.com/pdf/CostaGupitta.pdf. Accessed: 29 December 2002.

Daily, C.M. & Dollinger, M.J. 1993. ‘Alternative methodologies for identifying family versus non-family-managed businesses’, *Journal of Small Business Management*, April: 79–90.

Davis, P. 1983. ‘Realizing the potential of the family business’, *Organizational Dynamics*, 5(1):47–56.

Donckels, R. & Frohlich, E. 1991. ‘Are family businesses really different? European experiences from STRATOS’, *Family Business Review*, 4(2): 149–160.

Donnelley, R.G. 1964. ‘The family business’, *Harvard Business Review*, 42(3): 93–105.

Dyer, W.G. & Sanchez, M. 1998. ‘Current state of family business theory and practice as reflected in *Family Business Review*’, *Family Business Review*, 11(4): 287–295.

Erikson, T. 2002. ‘Entrepreneurial capital: The emerging venture’s most important asset and competitive advantage’, *Journal of Business Venturing*, 17: 275–290.

Fiegener, M.K., Brown, B.M., Prince, R.A. & File, K.M. 1996. ‘Passing on strategic vision: Favored modes of successor preparation by CEOs of family and non-family firms’, *Journal of Small Business Management*, 34(3): 15–26.

Gallo, M.A. & Sveen, J. 1991. ‘Internationalizing the family business: Facilitating and restraining factors’, *Family Business Review*, 4(2): 181–190.

Gersick, K.E., Davis, J.A., Hampton, M.M. & Lansberg, I. 1997. *Generation to Generation: Life Cycles of the Family Business*. Boston: Harvard Business School Press.

Gibson, K. 2002. ‘A case for the family-owned conglomerate’, *McKinsey Quarterly*, 4: 126–137.

Habbershon, T.G. & Williams, M. 1999. A resource based framework for assessing the strategic advantages of family firms. Wharton Enterprising Families Initiative, Working Paper no. 101. Pennsylvania: University of Pennsylvania.

Habbershon, T.G. & Williams, M. 2001. A unified systems theory of family firm performance. Wharton Enterprising Families Initiative, Working Paper no. 103. Pennsylvania: University of Pennsylvania.

Hambrick, D.C. & Mason, P.A. 1984. ‘Upper echelons: The organization as a reflection of its topmanagers’, *Academy of Management Review*, 9(2): 193–206.

Harrison, J.R. & Carroll, G.R. 1991. ‘Keeping the faith: A model of cultural transmission in formal organizations’, *Administrative Science Quarterly*, 36: 552–582.

Heck, R., Upton, N., Bellet, W., Dunn, B. & Parady, P. 1994. Family business as a field of study. Working paper, International Family Business Program Association. Durham: University of New Hampshire. [Online] Available from: www. genusresources.com. Accessed: 12 January2003.

Hoelscher, M.L. 2002. The relationship between family capital and family business performance: Collaboration and conflict as moderators. PhD thesis. Texas: Texas Tech University.

Hoy, F. & Verser, T.G. 1994. ‘Emerging business, emerging field: Entrepreneurship and the family firm’, *Entrepreneurship Theory and Practice*, 19(1): 923.

Jones, G. & Rose, M.B. 1993. ‘Family capitalism’, *Business History*, 35(4): 116.

Karakoulaki, H. 2002. Social capital vs family capital: Allies or opponents? Economic Geography Research Group, Working Paper 0211. Oxford: University of Oxford. [Online] Available from: www.geog.ox.ac.uk/research/wpapers. Accessed: 3 March 2003.

Kelly, L.M., Athanassiou, N. & Crittenden, W.F. 2000. ‘Founder centrality and the strategic behavior in the family-owned firm’, *Entrepreneurship Theory and Practice*, 25(2): 27–42.

Kepner, E. 1983. ‘The family and the firm: A coevolutionary perspective’, *Organizational Dynamics*, 5(1): 57–70.

Kim, D., Kandemir, S. & Cavusgil, S.T. 2003. ‘The role of family conglomerates in emerging markets: What Western companies should know’, *Thunderbird International Business Review* (forthcoming). [Online] Available from: www.globaledge.msu.edu/KnowledgeRoom. Accessed: 13 February 2003.

Lansberg, I. 1983. ‘Managing human resources in family firms: The problem of institutional overlap’, *Organizational Dynamics*, 5(1): 39–46.

Lansberg, I. 1999. *Succeeding Generations: Realizing the Dream of Families in Business*. Boston: Harvard Business School Press.

Levinson, H. 1971. ‘Conflicts that plague family businesses’, *Harvard Business Review*, 49(2): 90–98.

Liang, T.W. & Jek, G.J. 2000. The determinants of family business growth and performance in Singapore. Working Paper, Wharton SMU. [Online] Available from: www.smu.edu.sg/research/papers.htm. Accessed: 24 November 2002.

Litz, R.A. 1997. ‘The family firm’s exclusion from business school research: Explaining the void; addressing the opportunity’, *Entrepreneurship Theory and Practice*, 21(3): 55–71.

McConaughy, D.L., Matthews, C.H. & Fialko, A.S. 1997. Founding family controlled firms: Efficiency, risk and value. Paper presented at the 1997 USASBE Annual National Conference, 21–24 June 1997, San Francisco. [Online] Available from: www. usasbe.org/knowledge/proceedings/1997. Accessed: 22 December 2002.

McConaughy, D.L., Walker, M.C., Henderson, G.V. & Mishra, C.S. 1998. ‘Founding family controlled firms: Efficiency and value’, *Review of Financial Economics*, 7(1): 1–19.

Miller, R. 2000. ‘How culture affects mergers and acquisitions’, *Industrial Management*, September–October: 22–26.

Narva, R.L. 2001. Heritage and tradition in family business: How family controlled enterprises con-

nect the experience of their past to the promise of their future. Paper presented at the 2nd Annual Family Business Gathering, Stetson University, 23 February 2001, Florida. [Online] Available from: www.genusresources.com. Accessed: 15 February 2003.

Neubauer, F. & Lank, A.G. 1998. *The Family Business: Its Governance for Sustainability*. London: Macmillan.

Ogbonna, E. & Harris, L.C. 2001. 'The founder's legacy: Hangover or inheritance?', *British Journal of Management*, 12: 13–31.

Rowe, W.G. 2001. 'Creating wealth in organizations: The role of strategic leadership', *Academy of Management Executive*, 15(1): 81–94.

Ryan, C. 1995. 'Are family businesses better?', *Productivity SA,* March–April: 11–13.

Schein, E.H. 1983. 'The role of the founder in creating organizational culture', *Organizational Dynamics,* 5(1): 13–28.

Stafford, K., Duncan, K.A., Dane, S. & Winter, M. 1999. 'A research model of sustainable family businesses', *Family Business Review*, 12(3): 197–208.

Upton, N. & Petty, W. 2000. 'Venture capital investment and US family business', *Venture Capital,* 2(1): 27–39.

Vilaseca, A. 1995. Finance in the family business. Research Paper no. 287BIS, University of Navarra.

Wall, R.A. 1998. 'An empirical investigation of the production function of the family firm', *Journal of Small Business Management*, 36(2): 24–32.

Ward, J.L. 1987. *Keeping the Family Business Healthy: How to Plan for Continuing Growth, Profitability, and Family Leadership*. San Francisco: Jossey Bass.

Westhead, P. & Cowling, M. 1998. 'Family firm research: The need for a methodological rethink', *Entrepreneurship Theory and Practice*, 23(1): 31–56.

Wiener, Y. 1988. 'Forms of value systems: A focus on organizational effectiveness and cultural change and maintenance', *Academy of Management Review*, 13(4): 534–545.

Predicting turnover behaviour of direct salespeople

*P. Msweli-Mbanga**

The purpose of this article is to enhance the understanding of the turnover behaviour of distributors in direct selling organisations. The study examined how the antecedents and outcomes of five endogenous variables (intention to quit, organisational commitment, job satisfaction, performance and perceptions of the marketing mix) interrelate to explain and predict the turnover behaviour of direct salespeople. The key findings of this study are that organisational commitment is the best predictor of the turnover behaviour of distributors, and perceptions of the marketing mix the best predictor of job satisfaction. Of the three antecedents of performance (customer profile, perceptions of the marketing mix and organisational commitment), organisational commitment has the strongest predictive power.

Introduction

Direct salespeople are independent business owners who sell products directly to customers. There are two organisational formats in which direct selling can be implemented: single-level (SL) format and multilevel (ML) format (Brodie, Stanworth & Wotruba 2002). In the former, the salespeople focus on selling and achieving compensation based on their own sales. In multilevel organisations, however, direct salespeople are encouraged to build a sales organisation of people like themselves by using a unique coaching and training system called 'sponsoring'. In the sponsoring system, independent contractors share their knowledge and expertise with new entrants, who are still learning about the business. In return for this commitment, the sponsors are compensated for the sales efforts of their respective teams over and above the income they earn from their own retail sales.

Direct selling allows companies to minimise the overhead costs associated with retail sites and extensive management and salaried staff. This non-store method of retailing also offers global reach as companies such as Amway, Avroy Shlain, GNLD International and Herbalife expand their operations in overseas markets. For example, GNLD has over 200 000 distributors in more than 50 global markets (GNLD 2002), while Amway has more than three million independent business owners in 80 countries and territories worldwide (Amway 2002).

In the direct selling industry, turnover rates among direct salespeople of up to 100% have been reported in the literature (Granfield & Nicols 1975; Wotruba, Sciglimpaglia & Tyagi 1987; Wotruba 1990; Brodie & Stanworth 1998). Given the considerable investment that must be made in the recruitment and training of new direct salespeople, controlling the cost of turnover has become an important strategic issue that direct salespeople need to address.

It is not only the costs directly associated with the turnover that are at issue. High rates of salesforce turnover can also impact negatively on customers. Customers who are poorly serviced as a result of a lack of continuity in their relationship with their distributor can create unfavourable perceptions of a direct selling organisation. Raymond & Tanner (1994) argue that there are many cases of customers developing negative perceptions as a direct consequence of the frustration experienced when the distributors supplying them with their products leave the organisation.

Distributor turnover is also a concern because of the pivotal role distributors have in the structure of a direct selling organisation (Berry 1997; Biggart 1990). Not only are they the interface between the supplying organisation and the customer, but they also have an ongoing responsibility for training and motivating sales teams, and for implementing the marketing strategies of the supplying organisations.

The purpose of this study is to develop and empirically test a structural model that explains

* Professor P. Msweli-Mbanga is an associate professor of marketing in the School of Economics and Management, University of KwaZulu-Natal. E-mail: mbangap@ukzn.ac.za.

and predicts the turnover behaviour of distributors in direct selling organisations. More specifically, this study develops and tests a theoretical model that integrates the antecedents and outcomes of job satisfaction, in-role performance, and organisational commitment with the turnover behaviour of direct salespeople. The paper addresses two questions: (1) What are the antecedents of turnover behaviour in direct selling organisations? (2) How do these variables interrelate to explain and predict the turnover behaviour of direct salespeople?

The development of the structural model and the theoretical foundation from which the model is derived is discussed in the first section of this paper. The second section of the paper describes the methodology underlying the model tested in this study and the analytical procedures used to test it. The third section presents the findings of the research. The paper concludes by discussing the findings and their implications for direct selling practitioners.

Conceptual framework and hypotheses

Turnover behaviour

The terms 'retention' and 'turnover' are often used interchangeably in the literature on employee turnover behaviour. Price (1989: 462) defines turnover rate "as being both the entrance of new employees into the organisation and the departure of existing employees from the organisation". Muchinsky & Morrow (1980: 267) posit a similar definition of turnover as "an individual's voluntary termination of employment from an organisation". However, in direct selling organisations, turnover is viewed differently, because there is no formal quitting behaviour (Wotruba 1990). Distributors are either active, meaning that they regularly purchase products from a direct selling organisation and are involved in recruiting other participants, or they are inactive when they do not do so.

While Wotruba (1990) uses inactivity to measure the turnover behaviour of distributors, Sager (1991) raises a valid concern about the use of intention to quit and propensity to leave as surrogate measures of turnover. Sager argues that while intention to quit might be a consistent precursor of turnover, it cannot be viewed as being a perfect surrogate. Other authors (Lucas, Parasuraman, Davis & Enis 1987; Johnston, Futrell, Parasuraman & Sager 1988) also question the use of intention to quit as a sound predictor of actual turnover. Sager & Menon (1994) contend that a salesperson can intend to quit, but remain in a job indefinitely. They report that intention to search for other jobs is a stronger predictor of salespeople's turnover than intention to quit. Their argument is based on the fact that intention to leave is a volatile construct that can change from day to day. An employee may intend to leave, yet remain in the job for some reason.

On the other hand, Fishbein & Ajzen (1975) argue that a person's intention to perform a behaviour is the immediate determinant of the action. This contention is widely supported in literature (see, for example, Johnston, Varadarajan, Futrell & Sager 1987; Futrell & Parasuraman 1984; Naumann, Widmier & Jackson 2000; Hom & Kinicki 2001; Shields & Price 2002). On balance, therefore, this study will adopt intention to quit as a surrogate variable for measuring the turnover behaviour of distributors in direct selling organisations. While researchers adopt different perspectives in researching intention to quit, they share a number of common views. These relate to the consideration of key constructs and antecedents of turnover such as job satisfaction, organisational commitment, performance and personal characteristics. The discussion that follows is structured according to these key constructs.

In this study, seven variables (intention to quit, job satisfaction, organisational commitment, distributor performance, customer profile, perceptions of the marketing mix and perceived quality) are hypothesised to interact as depicted in Figure 1. Theoretical justification of how these variables interact is provided in the following section.

Job satisfaction, organisational commitment and turnover

Earlier work on job satisfaction by Churchill, Ford & Walker (1974), which examined the extrinsic and intrinsic nature of a job, shaped subsequent research on the job satisfaction of salespeople. Churchill et al. (1974: 255) defined job satisfaction as "all characteristics of the job itself and the work environment which (industrial) salesmen find rewarding, fulfilling and satisfying, or frustrating and unsatisfying". Lucas et al. (1987) concur with Churchill and his colleagues in viewing job satisfaction as comprising two components: (1) extrinsic job satisfaction (namely, satisfaction with pay, job security, working condition, fringe benefits, recognition and status), and (2) intrinsic job satisfaction (namely, employees' satisfaction with the work itself and the opportunities provided for personal growth).

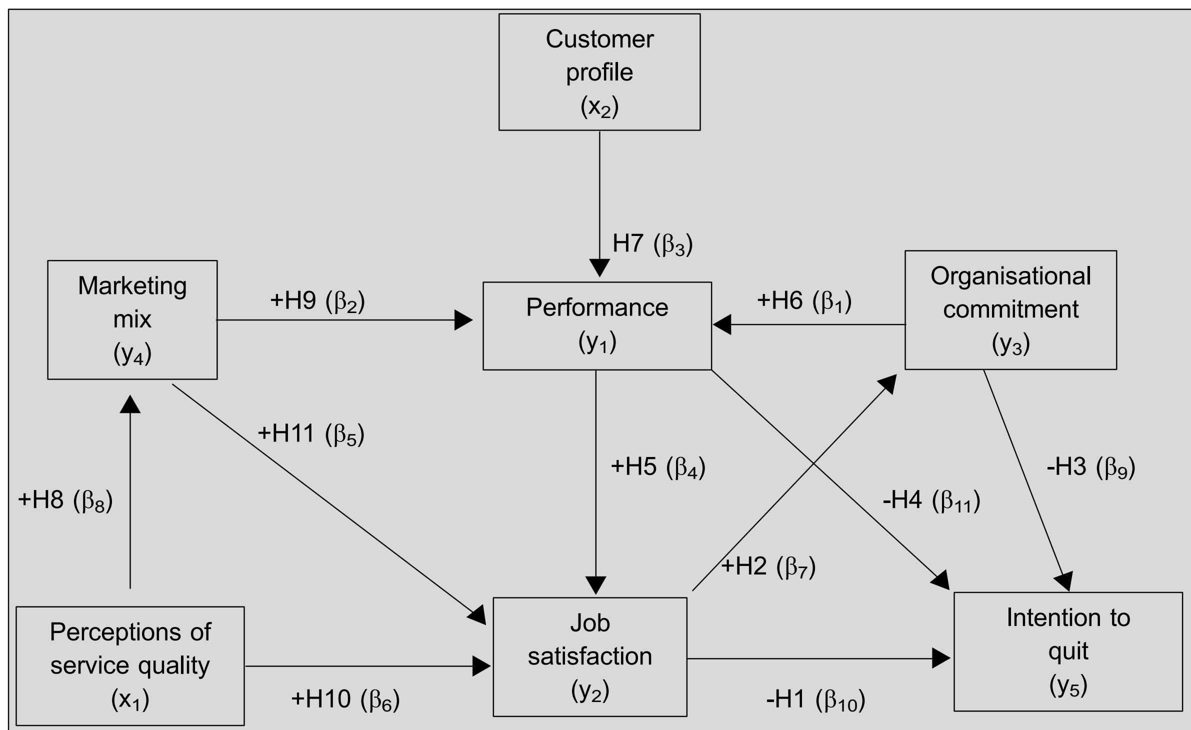


Figure 1: The structural model – hypothesised relationships

The idea of viewing job satisfaction as comprising two components originated from Herzberg's Two-Factor Theory (also known as motivator-hygiene theory). Herzberg's Two-Factor Theory contends that job satisfaction and dissatisfaction actually derive from two contrasting sources, whereby satisfaction is a result of the presence of motivator factors, and dissatisfaction the result of the lack of hygiene factors (Herzberg, Mausner & Snyderman 1959). The motivators are promotions opportunities, opportunity for personal growth, recognition, responsibility and achievement, and the hygiene factors are quality of supervision, company pay policies, physical working conditions, relation with others and job security.

Considerable research has been carried out concerning the relationship between job satisfaction and turnover, with and sometimes without the organisational commitment variable. In their meta-analytical study, Mathieu & Zajac (1990) found that organisational commitment has been reported to mediate the relationship between job satisfaction and turnover. More conclusive evidence of the relationship between job satisfaction and organisational commitment has been provided by Sager (1990), as well as by Cravens, Low & Moncrief (2001). The authors found that the organisational commitment of employees increased as job satisfaction increased. Moreover, organisational commitment was reduced when intention to quit increased. Sager (1990) and Cravens et al. also found that job satisfaction is not directly linked to intention to quit. Instead, organisational commitment is an intervening variable between job satisfaction and intention to quit.

Another perspective regarding the relationship between job satisfaction and intention to quit was put forward by MacKenzie, Podsakoff & Ahearne (1998), however, who found that job satisfaction is directly linked to intention to quit, implying that turnover is likely to be reduced when employees are satisfied with their jobs. In line with this argument, Podsakoff, Ahearne & MacKenzie (1997) suggested that performance could be increased and intention to quit would be reduced by making the work context more pleasant and supportive.

Given the varied views on the link between job satisfaction, organisational commitment and turnover, this study seeks to clarify the relationship between these three variables. The study also seeks to empirically estimate the strength of the job satisfaction and turnover relationship, with organisational commitment as an intervening variable. As such, the following hypotheses are advanced:

Hypothesis 1: Job satisfaction has a negative relationship with intention to quit.

Hypothesis 2: Job satisfaction has a positive relationship with organisational commitment.

Hypothesis 3: Organisational commitment has a negative relationship with intention to quit.

Performance and turnover

Performance, as defined by direct sales managers, consists of personal productivity measured in terms of sales volume, sales volume generated by one's network, and the number of other salespeople recruited. This indicates that recruiting salespeople in direct selling is directly linked to performance. The literature provides evidence of the link between intention to quit and the performance of salespeople (McNeilly & Russ 1992).

While sales performance is a widely researched variable (Churchill, Ford, Hartley & Walker 1985; O'Reilly & Chatman 1986; Williams & Anderson 1991; Msweli-Mbanga 2001), its antecedents and outcomes in the direct selling context still need to be examined on a wider scale. Wotruba's (1990) study, which examined the relationship between performance and the turnover behaviour of direct salespeople, showed a small (r = -.12) but significant correlation between the two variables. The hypothesis that performance moderates the relationship between job satisfaction and turnover was not supported in Wotruba's (1990) study. Drawing from Wotruba's findings, a fourth hypothesise was formulated as follows:

Hypothesis 4: Performance is negatively related to intention to quit.

Several researchers have constructed structural models that posit that job performance is related to turnover through job satisfaction as an intervening variable (Babakus, Cravens, Johnston & Moncrief, 1996). Mackenzie et al. (1998), among others, support the proposition that performance is an antecedent of job satisfaction. A fifth hypothesis was thus formulated as follows:

Hypothesis 5: Performance is positively related to job satisfaction.

The literature has not been conclusive about the relationship between performance and organisational commitment. For example, a study by MacKenzie et al. (1998) did not support the hypothesis that performance is positively related to organisational commitment. Earlier studies by Bluedorn (1982) and Porter & Steers (1973) report a significant positive relationship between performance and organisational commitment. Indeed, distributors who display high levels of organisational commitment are expected to generate more sales and build bigger networks of salespeople. This leads to the following hypothesis:

Hypothesis 6: Organisational commitment is positively associated with performance.

Hypothesis 7 is based on Biggart's (1990) study, which examined the difference between traditional bureaucratic firms and direct selling firms. Biggart (1990) contends that direct selling firms differ fundamentally from bureaucratic firms in that they pursue profit by making social networks serve business ends. This indicates that friends and family are not just personal emotional contacts – they are regarded as serious business contacts. It is therefore reasonable to think that if distributors hold favourable perceptions of the marketing mix elements of their offering, they are more likely to generate more business from their friends and family. This rationale is the basis for the following hypothesis:

Hypothesis 7: The higher the proportion of friends and family in a distributor's customer profile, the higher the performance.

Perception of the marketing mix, perceived quality and customer profile

Value creation has become the central purpose of marketing (Zeithaml 1988, 2000; Lam, Shankar, Erramilli & Murthy 2004). As such, marketers are constantly looking for new ways to improve the perceived value of their offerings. These efforts have, to a large extent, created 'value conscious' consumers, and direct sales consumers are no exception. Direct salespeople are also looking for ways of maximising the value they provide to their consumers.

There is sufficient evidence in the literature to suggest that customers' value perceptions can be increased by enhancing their perceptions of the marketing mix, as well as their perceptions of service quality (Zeithaml 1988, Bolton & Drew 1991; Grewal & Monroe 1998; Kashyap & Bojanic 2000; Anderson, Fornell & Lehmann 1994). This study postulates that if distributors hold favourable perceptions of the organisation's marketing mix, as well as the firm's service quality, they will exhibit

higher organisational commitment, performance and job satisfaction. The following hypotheses are thus stated:

Hypothesis 8: Perceptions of service quality are positively related to perceptions of the marketing mix.

Hypothesis 9: There is a positive relationship between distributors' perceptions of the marketing mix and distributor performance.

Hypothesis 10: Distributors' perceptions of service quality are positively related to their job satisfaction.

Hypothesis 11: There is a positive relationship between distributors' perceptions of the marketing mix and their job satisfaction.

Table 1 illustrates the 11 hypotheses discussed in the foregoing section.

Table 1: Summary of hypotheses

H1:	Job satisfaction → (-) intention to quit
H2	Job satisfaction → (+) organisational commitment
H3:	Organisational commitment → (-) intention to quit
H4:	Performance → (-) intention to quit
H5:	Performance → (+) job satisfaction
H6:	Organisational commitment → (+) performance
H7:	Customer profile → performance
H8:	Perceptions of quality → (+) marketing mix
H9:	Marketing mix → (+) performance
H10:	Perceptions of quality → (+) job satisfaction
H11:	Marketing mix → (+) job satisfaction

Method

Sample

Approximately 40 direct selling firms are affiliated to the Direct Selling Association of South Africa. Of these, only the 18 firms that had been in existence for more than five years were approached to participate in the study, and only three were willing to do so. Each of these companies was asked to provide 1000 randomly selected names and addresses of their distributors operating in South Africa. A total of 3000 questionnaires was mailed to the target respondents with self-addressed stamped envelopes. Five hundred questionnaires were returned, which represented a response rate of 16.7%.

Two weeks before the questionnaires were mailed, a pre-test was conducted using a sample of 50 distributors. Minor changes of terminology were made as a result.

Variables and measures

Intention to quit. Intention to quit was measured by asking respondents to indicate their likelihood to quit in the next three months as (1) very low, (2) low, (3) high, (4) very high. A single item scale was used in the absence of any appropriate multiple item instruments. Past studies have used single item scales to measure likelihood to quit (see, for example, Wotruba 1990 and Tyagi & Wotruba 1993).

Job satisfaction. Job satisfaction was measured using the shorter version of the Industrial Salesperson Job Satisfaction scale (INDSALES) developed by Comer, Machleit & Legace (1989). The shorter version of the INDSALES scale consists of 28 items measuring seven dimensions of job satisfaction. For the purposes of this study, six dimensions were used, and the 'promotion and advancement' dimension was omitted because it is irrelevant in the direct selling context. Distributors are not employees of their host direct selling organisations, and promotion in the sense described in the scale is thus not applicable.

Organisational commitment. Mowday, Steers & Porter's (1979) scale was used to measure organisational commitment. The shortened nine-item version of the Organisational Commitment Questionnaire (OCQ) was utilised, using only the positive item scores, as suggested by the authors.

Distributor performance. Distributor performance was measured using three items similar to those used by direct selling organisations to assess the performance of their salespeople: (1) the number of distributors recruited during the preceding 12 months – the responses were scored on a scale of 1–3, (1 = 0–10; 2 = 11–20; and 3 = 21+); (2) their average monthly turnover including the turnover of

their team – the responses were scored on a scale of 1–3, (1 = R0–R5000; 2 = R5001–R10 000; and 3 = R10 000+); (3) their monthly personal sales – the responses were scored on a scale of 1–3; (1 = 'low performers'; 2 = 'average performers'; and 3 = 'high performers'). The responses with a score of 1 were recorded as low performers, those with a score of 2 as average performers, and those with a score of 3 as high performers.

Customer profile. The variable was measured by asking respondents to respond to two questions: (1) What percentage of your monthly personal sales are made to friends and family? The responses were scored on a scale of 1–6 (1 = 0–10%; 2 = 11–20%; 3 = 21–30%; 4 = 31–40%; 5 = 41–50%; and 6 = 51%+). (2) Of your total group sales, what percentage is sold to friends and family? Again, the responses were scored on a scale of 1–6 (1 = 0–10%; 2 = 11–20%; 3 = 21–30%; 4 =31–40%; 5 = 41–50%; and 6 = 51%+).

Perceptions of the marketing mix. Four items were used to measure respondents' perceptions of the marketing mix construct: (1) I am happy with the product range offered by my organisation; (2) The product range addresses the needs of my customers; (3) I am happy with how the products are promoted in my organisation; (4) I am happy with the price of the products in my organisation. The items were measured on a five-point Likert scale, ranging from 'strongly disagree' to 'strongly agree', with 1 indicating negative perceptions of the marketing mix and 5 indicating positive perceptions.

Perceived quality. Cronin & Taylor's (1992) SERVPERF scale was used to rate salespeople's service quality perceptions along five dimensions: tangibility, reliability, responsiveness, assurance and empathy. Item scores were averaged across all five dimensions to form an overall service quality score.

Analytical procedures

Based on the hypothesised relationships in Figure 1, four equations have been constructed, linking two exogenous constructs to five endogenous constructs as follows:

$y_1 = \beta_1 y_3 + \beta_2 y_4 + \beta_3 x_2 + E_1$ (Equation 1)

$y_2 = \beta_4 y_1 + \beta_5 y_4 + \beta_6 x_1 + E_2$ (Equation 2)

$y_3 = \beta_7 y_2 + E_3$ (Equation 3)

$y_4 = \beta_8 x_1 + E_5$ (Equation 4)

$y_5 = \beta_9 y_3 + \beta_{10} y_2 + \beta_{11} y_1 + E_4$ (Equation 5)

Where:

y_1 = performance
y_2 = job satisfaction
y_3 = organisational commitment
y_4 = perception of the marketing mix
y_5 = intention to quit
x_1 = perception of quality
x_2 = customer profile
$E_1 E_2$ are error terms for equations 1–5
β_1–β_{11} are the path coefficients for each hypothesised relationship.

The generalised least squares (GLS) method in Amos version 3.6 was employed to estimate the path coefficients in the model. The GLS method is often used when there are unequal variances among the error terms (Draper & Smith 1981). In estimating the model parameters, the identification problems were checked, that is, the inability of the proposed model to generate unique estimates, usually as a result of incorrectly specifying the model (Maruyama 1998). The model had no identification problems. Next, the results were examined for Heywood cases and other offending estimates. Offending estimates occur when the estimated error terms are negative, or when the standardised regression weights are greater than 1 (Hair, Anderson, Tatham & Black 1998). As can be seen in Table 3, there are no offending estimates.

Empirical results

The means, standard deviations and reliabilities of the variables are given in Table 2.

The scores on all items for each variable were averaged to obtain single dimension scores. This was done because the total number of measured items was large and it was impractical to incorporate all the items in the model. Furthermore, multiple measures represented by a single score were applied in previous research using structural equation modelling (see Sager & Menon 1994; Bagozzi 1980; Mackenzie et al. 1998).

The mean scores for the variables indicate that, in general, the respondents exhibit a low intention to quit (1.61), relatively high job satisfaction (3.96), favourable perceptions of quality (4.51), low performance (1.5) and very high organisational commitment (6.32). A relatively low mean score (2.47) for customer profile indicates that a large proportion of their customers are *not* from among their friends and family. The respondents scored high on perceptions of the marketing mix (4.38).

Table 2: Means, standard deviations and reliabilities

Variables	No. of items	Measurement/ Likert scale	Mean	Std dev.	Alpha
Intention to quit	1	Four-point Likert	1.61	.96	—
Organisational commitment	8	Seven-point Likert	6.32	.82	.91
Job satisfaction	20	Five-point Likert	3.96	.59	.73
Perceptions of quality	19	Seven-point Likert	4.51	.48	.70
Performance	3	Three-point Likert	1.50	.61	.60
Customer profile	2	Six-point Likert	2.47	1.38	—
Marketing mix	4	Five-point Likert	4.38	.59	.74

Cronbach's alpha was used to assess the internal consistency reliability of the measuring scales. Table 2 indicates that the internal consistency of the variables, measured by Cronbach's alpha, are within the benchmark of .70 as suggested by Nunnaly & Bernstein (1994), and .60 for exploratory research as suggested by Hair et al. (1998). The single item measures were fixed to unity, because it is not empirically possible to estimate their reliabilities (Hair et al. 1998). A correlation matrix of the seven variables was chosen as data input. According to Arbuckle (1997), a correlation matrix is preferred to a covariance matrix if the objective is to explore the pattern of interrelationships between a set of variables.

Table 3 reports standardised parameter estimates, standard error and critical ratio.

Table 3 shows that the regression weights were standardised to allow for comparison of the effect of each exogenous variable on the endogenous variable. The standardisation process was carried out by transforming the input data into new measurement variables with a mean of 0 and a standard deviation of 1.

Maruyama (1998) suggests, as a simple rule of thumb, that a parameter is significantly different from zero if it is greater than twice its standard error in absolute value. However, Bentler (1980) contends that all parameters whose estimates are large compared to their standard errors are significant. In concurrence with Bentler's (1980) viewpoint, structural equations that contain insignificant coefficients were identified and removed.

Unlike the findings reported by Hom, Katerberg & Hulin (1979) and MacKenzie et al. (1998), which showed a statistically significant relationship between job satisfaction and intention to quit, the results of this study showed that the job satisfaction intention to quit path was insignificant. Hypothesis 1 was thus not supported. Three additional paths were not significant – the performance intention to quit path, the perceptions of the marketing mix performance path, and the customer profile performance path. Hypotheses 4, 7 and 9 were thus not supported. The model was re-specified without the insignificant paths, and the parameter estimates of the revised model are presented in Table 3.

Table 3: Parameter estimates for the revised structural model

Hypotheses	Standardised parameter estimates	Standard error (SE)	Critical ratio (CR)
H2: Job satisfaction → (+) organisational commitment	0.809	0.124	10.621
H3: Organisational commitment → (-) intention to quit	-0.388	0.071	-5.785
H5: Performance → (+) job satisfaction	-0.151	0.049	-2.350
H6: Organisational commitment → (+) performance	0.752	0.074	7.580
H8: Perceptions of quality → (+) marketing mix	0.145	0.081	1.986
H10: Perceptions of quality → (+) job satisfaction	0.157	0.081	2.189
H11: Marketing mix → (+) job satisfaction	0.651	0.083	7.938

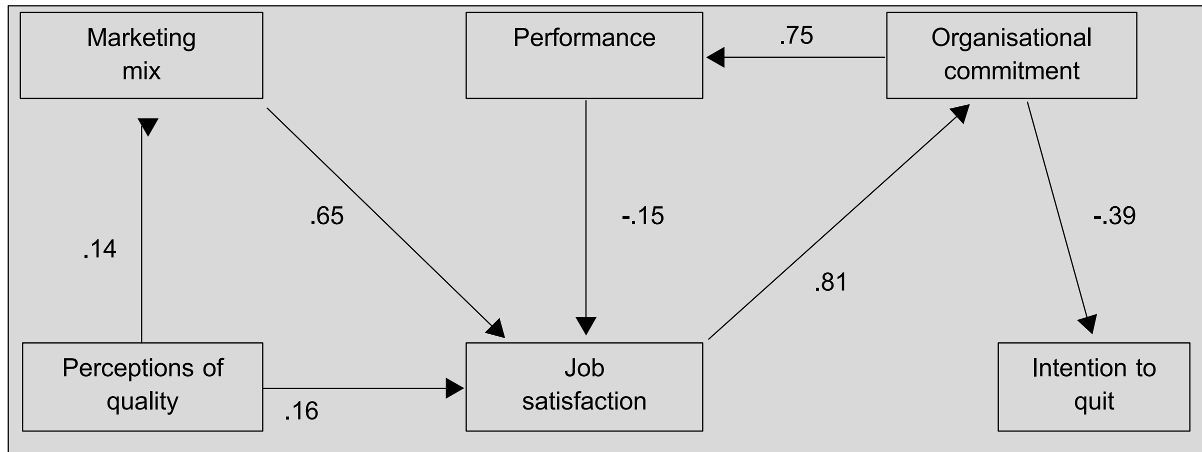


Figure 2: Path diagram for the revised distributor retention model

Figure 2 presents a summary of the findings of this study. As expected, the higher the organisational commitment, the lower the likelihood to quit, lending support to Hypothesis 3.

Perceptions of the marketing mix are the strongest predictor (0.65) of job satisfaction. What seemed somewhat surprising is the negative direction of the performance/job satisfaction relationship. This is contrary to the hypothesis of this study, which postulates that distributor performance has a positive association with job satisfaction.

The chi-square value of 98.13 with 34 degrees of freedom is statistically significant at the .000 significance level. The structural equation modelling method of testing for model fit differs from the customary desire to find statistical significance (Hair et al. 1998). As argued in the literature (Hair et al. 1998; Maruyama 1998), the goal is to find non-significant differences between actual and predicted model. Thus, a chi-square value with a significance level greater than 0.05 or 0.01 indicates that the actual and predicted input matrix are not statistically different (Hair et al. 1998).

Although the chi-square measure of this study may suggest that our data do not fit the proposed model, a chi-square measure has been criticised for its sensitivity to sample size. Other goodness-of-fit (GIF) measures therefore need to be taken into consideration when evaluating structural equation modelling. The GFI value of 0.91 and root square error of approximation (RMSEA) value of 0.08 reported in Table 4 are at a marginal acceptance level.

Table 4: Goodness-of-fit measures for revised model

Goodness-of-fit measures	
Chi-square	98.13
Degrees of freedom	34
Significance level	0.000
Goodness-of-fit index (GFI)	0.91
Adjusted goodness-of-fit index (AGFI)	0.90
Root square error of approximation (RMSEA)	0.08

Discussion and implications

The purpose of this article has been to enhance the understanding of the intention to quit among distributors in direct selling organisations. As depicted in Figures 1 and 2, five endogenous variables were the focus of this study: (1) intention to quit, (2) job satisfaction, (3) performance, (4) organisational commitment, and (5) perceptions of the marketing mix. The study examined how the antecedents and outcomes of these variables interrelate to explain and predict the intention to quit of direct salespeople.

Seven of the 11 hypothesised relationships were statistically significant. The four exceptions are: (1) the relationship between job satisfaction and intention to quit; (2) the relationship between performance and intention to quit; (3) the relationship between perceptions of the marketing mix and distributor performance; and (4) the relationship between customer profile and performance. The findings of this study show that job satisfaction does not directly influence intention to quit (see Figure 2). Instead, organisational commitment mediates the relationship between job satisfaction and intention to quit. This finding implies that improving organisational commitment is more likely to reduce the intention to quit among direct salespeople. Organi-

sational commitment was also found to be the strongest predictor of the intention to quit of salespeople.

The findings clearly show that committed and loyal distributors are more likely to stay in their organisations than uncommitted people who experience high levels of job satisfaction. Management in direct selling organisations can look at improving the processes that underlie job satisfaction, such as motivation, supervision, satisfaction with income, and relationship with fellow distributors as a means of enhancing organisational commitment.

The results also show that in the direct selling context, performance is not directly linked to the intention to quit. Instead, performance has been found to be a strong predictor of job satisfaction. The negative relationship between the two variables indicates that as distributor performance increases, job satisfaction decreases. This finding is inconsistent with the positive relationship observed by other authors (MacKenzie et al. 1998; Babakus et al. 1996). This finding suggests the need for further research to identify aspects of performance that lead to lower job satisfaction. As Babakus et al. (1996) suggest, job satisfaction among direct salespeople can be increased by increasing the salesperson's perception of organisational support through training and coaching. Indeed, distributor performance is directly linked to the extent to which distributors are exposed to business training (Berry 1997). It is worthwhile to further investigate the link between training and coaching, distributor performance and job satisfaction.

Consistent with our expectations, perceptions of service quality are positively related to perceptions of the marketing mix. Perceptions of the marketing mix also turned out to be the best predictor of job satisfaction. The major implication of these findings is that winning the support of customers requires that distributors should be 'sold to' the company's products and its marketing plans. This explanation is consistent with research in internal marketing (Grönroos 1994), which shows that internal customers must buy in if external customers are likewise to buy. One would then expect to have a direct link between perceptions of the marketing mix and distributor performance, as postulated in this study. Our finding of a low correlation between distributor performance and perceptions of the marketing mix may be explained by the fact that performance in direct selling involves more than selling to external customers. Performance also includes the number of people recruited and the size of the sales team built. This suggests that performance in direct selling demands more than sales competence. It also requires leadership and management skills because, in essence, direct salespeople are more than salespeople – they are independent contractors who own their own businesses.

It is noteworthy that organisational commitment is not only a predictor of the intention to quit among distributors in direct selling organisations. It is also the strongest predictor of performance. This implies that for management to reduce turnover and improve performance, they have to focus on organisational commitment. There is a need for future research to identify additional variables that impact on both commitment and performance. This would shed more light on the turnover behaviour of direct salespeople. The absolute fit measures of this study indicate that the model is marginally acceptable. However, further research is required to validate this model and test it in different direct selling contexts.

Summary

Direct selling has been defined as a face-to-face approach to retailing and distribution of consumer goods away from a fixed retail location. This form of retailing allows firms to gain entrance to global markets while avoiding the excessive overhead costs associated with retail sites and extensive management of salaried staff. The problem associated with the direct selling industry is the high turnover rate of direct salespeople. Given the considerable investment made in the recruitment and training of new direct salespeople, improving retention rates has become a key strategic issue that direct selling organisation have had to address. The costs directly associated with low retention rates are not the only issue. High salesforce turnover rates can also impact negatively on the organisation's customers. Customers who are poorly serviced as a result of a lack of continuity in their relationship with their salesperson can create unfavourable perceptions of a direct selling organisation.

The purpose of this study is to identify the antecedents of turnover behaviour in direct selling organisations. The study examined how the antecedents and outcomes of five variables (intention to quit, organisational commitment, job satisfaction, performance, and perceptions of the marketing mix) interrelate to explain and predict the turnover behaviour of direct salespeople. The theoretical-based hypotheses of this study are as follows:

H1: Job satisfaction has a negative relationship with intention to quit.
H2: Job satisfaction has a positive relationship with organisational commitment.
H3: Organisational commitment has a negative relationship with intention to quit.
H4: Performance is negatively related to intention to quit.
H5: Performance is positively related to job satisfaction.
H6: Organisational commitment is positively associated with performance.
H7: The higher the proportion of friends and family in the distributor customer profile, the higher the performance.
H8: Perceptions of service quality are positively related to perceptions of the marketing mix.
H9: There is a positive relationship between distributors' perceptions of the marketing mix and distributor performance.
H10: Distributors' perceptions of service quality are positively related to their job satisfaction.
H11: There is a positive relationship between distributors' perceptions of the marketing mix and their job satisfaction.

A total of 3000 questionnaires was mailed to the target respondents, and 500 questionnaires were returned, which represented a response rate of 16.7%. Intention to quit was measured using one item measured in a four-point Likert scale. Job satisfaction was measured using the shorter version of the Industrial Salesperson Job Satisfaction scale. An eight-item scale adapted from Mowday et al. (1979) was used to measure organisational commitment. Performance was measured using three items adapted from performance evaluation tools used by direct sales managers. Service quality was measured using Cronin & Taylor's (1992) SERVPERF scale.

The generalised least squares method was employed to estimate the path coefficients in the proposed model. Seven of the 11 hypothesised relationships were statistically significant (H2, H3, H5, H6, H8, H10 and H11). The findings show that organisational commitment mediates the relationship between job satisfaction and intention to quit. This finding implies that improving organisational commitment is likely to reduce the intention to quit among direct salespeople. Organisational commitment was also found to be the strongest predictor of the intention to quit, and perceptions of the marketing mix the best predictor of job satisfaction.

Overall, the findings indicate that the higher the salespeople's perceptions of service quality, the higher their job satisfaction. High job satisfaction is more likely to result in high organisational commitment, which leads to low intention to quit. Therefore, if direct selling management improves service quality, salespeople would be more likely to experience increased job satisfaction and organisational commitment and would be more likely to stay with their organisations. Furthermore, high organisational commitment is more likely to improve sales performance.

Acknowledgements

Thanks are due to Thomas Wotruba and two anonymous reviewers for their helpful comments.

References

Amway 2002. Explore the Amway Opportunity. Available at: www.amway.com

Anderson, E.W., Fornell, C. & Lehmann, D.R. 1994. 'Customer satisfaction, market share, and profitability: findings from Sweden', *Journal of Marketing,* 58(July): 53–66.

Arbuckle, J.L. 1997. *Amos User's Guide Version 3.6* USA: Small Waters Corporation.

Babakus, E., Cravens, D.W., Johnston, M. & Moncrief, W.C. 1996. 'Examining the role of organisational variables in the salesperson job satisfaction model', *Journal of Personal Selling and Sales Management,* XVI(3): 33–46.

Bagozzi, R.P. 1980. 'Performance and satisfaction in an industrial salesforce: an examination of their antecedents and simultaneity', *Journal of Marketing*, 44(Spring): 65–77.

Bentler, P.M. 1980. 'Multivariate analysis with latent variables: causal modelling', *Annual Review of Psychology,* 31: 419–456.

Berry, R. 1997. *Direct Selling – From Door to Door to Network Marketing*. Oxford: Butterworth Heinemann.

Biggart, N.W. 1990. *Charismatic Capitalism: Direct Selling Organisations in America.* Chicago: University of Chicago Press.

Bluedorn, A.C. 1982. 'A unified model of turnover from organisations', *Human Relations,* 35(2): 135–153.

Bolton, R.N. & Drew, J.H. 1991. 'A multistage model of customers' assessments of service quality and value', *Journal of Consumer Research,* 17(March): 375–384.

Brodie, S. & Stanworth, J. 1998. 'Independent contractors in direct selling', *International Small Business Journal,* (April–June), 16(3): 95–101.

Brodie, S., Stanworth, J. & Wotruba, T. 2002. 'Comparisons of salespeople in multilevel vs single-level direct selling organisations', *Journal of Personal Selling and Sales Management,* XX11 (2): 1–18.

Churchill, G.A., Ford, N.M. & Walker, O.C. 1974. 'Measuring the job satisfaction of industrial salesmen', *Journal of Marketing Research,* 11: 254–260.

Churchill, G.A., Ford, N.M., Hartley, S.W. & Walker, O.C. 1985. 'The determinants of salesperson performance: a meta-analysis', *Journal of Marketing Research*, 22: 103–118.

Comer, J.M., Machleit, K.A. & Legace, R.R. 1989. 'Psychometric assessment of a reduced version of INDSALES', *Journal of Business Research,* 18: 291–302.

Cravens, D.W., Low, G.S. & Moncrief, W.C. 2001. 'The role of satisfaction with territory design on the motivation, attitudes, and work outcomes of salespeople', *Journal of the Academy of Marketing Science,* 29(20): 165–178.

Cronin, J.J. & Taylor, S.A. 1992. 'Measuring service quality: a re-examination and extension', *Journal of Marketing,* 56(July): 55–68.

Draper, N.R. & Smith, H. 1981. *Applied Regression Analysis*. New York: John Wiley and Sons.

Fishbein, M. & Ajzen, I. 1975. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research Reading*. MA: Addison-Wesley.

Futrell, C.M. & Parasuraman, A. 1984. 'The relationship of satisfaction and performance to salesforce turnover', *Journal of Marketing,* 48(Fall): 33–40.

GNLD 2002. A world of opportunity. Available at: www.gnld.com.

Granfield, M. & Nicols, A. 1975. 'Economic and marketing aspects of the direct selling industry', *Journal of Retailing*, 51(1): 33–50.

Grewal, D. & Monroe, K.B. 1998. 'The effects of price-comparison advertising on buyers', *Journal of Marketing,* 62(2): 46–66.

Grönroos, C. 1994. 'From marketing mix to relationship marketing: towards a paradigm shift in marketing', *Management Decision,* 32(2): 4–20.

Hair, J.F. (Jr), Anderson, R.E., Tatham, R.L. & Black, W.C. 1998. *Multivariate Data Analysis*. 5th edition. New Jersey: Prentice Hall.

Herzberg, F., Mausner, B. & Snyderman, B.B. 1959. *The Motivation to Work*. 2nd edition. USA: John Wiley and Sons.

Hom, P.W., Katerberg, R. & Hulin, C.L. 1979. 'Comparative examination of three approaches to the prediction of turnover', *Journal of Applied Psychology,* 64(3): 280–290.

Hom, P.W. & Kinicki, A.J. 2001. 'Toward a greater understanding of how dissatisfaction drives employee turnover', *Academy of Management Journal,* 44(5): 975–987.

Johnston, M.W., Varadarajan, P.R., Futrell, C.M. & Sager, J. 1987. 'The relationship between organisational commitment, job satisfaction and turnover among new salespeople', *Journal of Personal Selling and Sales Management*, VII(Nov): 29–38.

Johnston, M.W., Futrell, C.M., Parasuraman, A. & Sager, J. 1988. 'Performance and job satisfaction effects on salesperson turnover: a replication and extension', *Journal of Business Research*, 16(1): 67–83.

Kashyap, R. & Bojanic, D.C. 2000. 'A structural analysis of value, quality, and price perceptions of business and leisure travellers', *Journal of Travel Research,* 39: 45–51.

Lam, S.Y., Shankar, V., Erramilli, M.K. & Murthy, B. 2004. 'Customer value, satisfaction, loyalty, and switching costs: an illustration from a business-to-business service context', *Journal of the Academy of Marketing Science,* 32(3): 271–299.

Lucas, G.H. (Jr), Parasuraman, A., Davis, R.A. & Enis, B.M. 1987. 'An empirical study of salesforce turnover', *Journal of Marketing*, (July): 34–59.

MacKenzie, S.B., Podsakoff, P.M. & Ahearne, M. 1998. 'Some possible antecedents and consequences of in-role and extra-role salesperson performance', *Journal of Marketing*, 62(July): 224–247.

Maruyama, G.M. 1998. *Basics of Structural Equation Modelling*. London: Sage.

Mathieu, J.E. & Zajac, D.M. 1990. 'A review and meta-analysis of the antecedents, correlates, and consequences of organisational commitment', *Psychological Bulletin*, 18(2): 171–197.

McNeilly, K.M. & Russ, F. 1992. 'The moderating effect of sales performance on relationships involving antecedents of turnover', *Journal of Personal Selling and Sales Management,* XII(1): 9–20.

Mowday, R.T., Steers, R.M. & Porter, L.W. 1979. 'The measurement of organisational commitment', *Journal of Vocational Behaviour,* 14(2): 224–247.

Msweli-Mbanga, P. 2001. 'Modelling distributor performance in network marketing organisations', *South African Journal of Business Management*, 32(3.): 33–40.

Muchinsky, P.M. & Morrow, P.C. 1980. 'A multidisciplinary model of voluntary employee turnover', *Journal of Vocational Behaviour*, 17: 263–290.

Naumann, E., Widmier, S.M. & Jackson, D.W. 2000. 'Examining the relationship between work attitudes and propensity to leave among expatriate salespeople', *Journal of Personal Selling and Sales Management,* XX(4): 227–241.

Nunnaly, J.C. & Bernstein, I.R. 1994. *Psychometrics Theory.* 3rd edition. New York: McGraw-Hill.

O'Reilly, C. & Chatman, J. 1986. 'Organisational commitment and psychological attachment: the effects of compliance, identification, and internalisation on prosocial behaviour', *Journal of Applied Psychology*, 71(3): 492–499.

Podsakoff, P.M., Ahearne M. & MacKenzie, S.B. 1997. 'Organisational citizenship behaviour and the quantity and quality of work group performance', *Journal of Applied Psychology*, 82(2): 262–270.

Porter, L.W. & Steers, R.M. 1973. 'Organisational, work,

and personal factors in employee turnover and absenteeism.' *Psychological Bulletin*, 80(2): 151–176.

Price, J.L. 1989. 'The impact of turnover on the organisation.' *Work and Occupations,* 16(4): 461–473.

Raymond, M. & Tanner, J.F. (Jr) 1994. 'Maintaining customer relationship in direct sales: stimulating repeat purchase behaviour,' *Journal of Personal Selling and Sales Management,* 14(3): 67–76.

Sager, J.K. 1990. 'How to retain salespeople.' *Industrial Marketing Management,* 19: 155–166.

Sager, J.K. 1991. 'A longitudinal assessment of change in salesforce turnover', *Journal of the Academy of Marketing Science,* 19(1): 25–36.

Sager, J.K. & Menon, A. 1994. 'The behavioural intentions in turnover of salespeople', *Journal of Business Research,* 2: 179–188.

Shields, M.A. & Price, S.W. 2002. 'Racial harassment, job satisfaction and intentions to quit: evidence from the British nursing profession', *Economica,* 69: 295–326.

Tyagi, P.K. & Wotruba, T.R. 1993. 'An exploratory study of reverse causality relationships among salesforce turnover variables', *Journal of the Academy of Marketing Science,* 21(2): 143–153.

Williams, L.J. & Anderson, S.E. 1991. 'Job satisfaction and organisational commitment as predictors of organisational citizenship and in-role behaviours.' *Journal of Management*, 17(3): 601–617.

Wotruba, T.R., Sciglimpaglia, D. & Tyagi, P.K. 1987. 'Toward a model of turnover in direct selling organisations', *AMA Winter Education Conference Marketing Theory Proceedings*, 348–353.

Wotruba, T.R. 1990. 'The relationship of job image, performance, and job satisfaction to inactivity-proneness of direct salespeople', *Journal of the Academy of Marketing Science,* 18(2): 113–121.

Zeithaml, V.A. 1988. 'Consumer perceptions of price, quality, and value: A means end model and synthesis of evidence', *Journal of Marketing,* 52(July): 2–22.

Zeithaml, V.A. 2000 'Service quality, profitability, and the economic worth of customers', *Journal of the Academy of Marketing Science,* 28(1): 67–85.

Consumer spending patterns in the informal retail trade sector of South Africa

D.H. Tustin[*]

Strategically speaking, retailing occupies a two-way position in a market economy. On the one hand, it is the last link in the marketing channel from producer to wholesaler to retailer. In this regard, retail mainly performs a distributive function. However, the role of the retailer is further extended to include an informative function. In performing the latter function, retail acts as an agency through which the consumer sends signals back through the channel to the manufacturer by a way of a specific reaction to the offering of goods in the market place. Consumers' reaction to product offerings and consumer behaviour is directly measurable from consumer spending patterns, among other indicators. Furthermore, quantifying consumer spending is of particular use in analysing demand for groups of goods and services and for tracking the spending trends of different consumer units. Although national data on consumer spending patterns in the formal sector are readily available, there are a dearth of disaggregated data on consumer spending patterns in the informal economy. This is particularly evident in the informal retail trade industry of South Africa, despite consensus that the informal retail trade industry is growing rapidly and is claiming a growing share of the total retail trade industry. These structural changes have resulted in changes in consumer demand that differ by product, region and population group.

Introduction

An analysis of consumer spending patterns in the informal retail trade industry of South Africa in particular serves to promote a better understanding of the nature of an industry showing a growing potential for generating future jobs and household income. Private business planners in the retail and other sectors of the South African economy, as well as government, could benefit from such an analysis, as it allows for the tracking of different spending trends among different types of consumer units. It also reflects the welfare of particular segments of the population (by socioeconomic group) and could well be used for business and government policy decision-making purposes. The generation of information on informal retail trade activities is also crucial for effective and efficient decision-making regarding the allocation of resources in an area. Finally, the information could be useful in selecting new market baskets of goods for the consumer price index to determine the relative importance of consumer price index components and to drive new cost weights for the market baskets.

However, it should be borne in mind that it is very difficult to obtain accurate information on the informal retail activities (especially on long-term trends) as businesses in this sector generally show relatively low survival rates (Ligthelm & Masuku 2003: 17). Data gathering in the informal retail trade industry is further complicated by the fact that entrepreneurs choose to operate in the informal sector precisely to avoid official business registration.

Objective and outline

Structural changes in the retail industry of South Africa point to a growing informal retail trade sector. This trend is also evident world-wide (Moran & McCully 2001; Bonnie & Van Dyck 2003). Against this background, the objective was to investigate the growing nature of the informal retail trade sector in South Africa by means of a demand-side analysis capturing the long-term trends of household consumer spending in South Africa with specific reference to the retail industry. A long-term analysis of this nature reflects the interaction of many factors that influence business decision-making. Among these, increasing affluence, changing demographics and changing tastes and lifestyles are particularly important.

[*] Dr D.H. Tustin is a senior researcher at the Bureau of Market Research, University of South Africa. E-mail: tustidh@unisa.ac.za.

The methodology used to identify household expenditure patterns in the informal retail trade industry entailed three phases. Firstly, a derived private consumption expenditure model was constructed. Following this, a retail trade indexing method was designed to determine household expenditure trends by product group for the formal retail sector. Finally, three experimental survey groups were used to identify household consumption expenditure trends within the informal retail trade sector.

Defining the informal retail trade industry

The difference between the formal and informal economy is best described by the commonly used definition applied to explain the working of the informal economy. The informal sector (as heterogeneous as the formal sector) is defined as all unregistered activities that do not contribute to the officially calculated (or observed) gross national product (Feige 1989, 1994; Frey & Pommerehne 1984). Smith (in Schneider 2002: 3) defines it as "market-based production of goods and services, whether legal or illegal, that escapes detection in the official estimates of gross domestic product (GDP)". The informal sector thus includes all economic activities that would generally be taxable were they reported to the state (tax) authorities. According to the 15th International Conference of Labour Statistics, the broader concept of an informal economy includes domestic workers, casual workers, agricultural workers and informal (retail trade) sales (Bono 2003: 2). In South Africa, informal activities are also evident in the small-scale manufacturing, construction, transport (minibus and taxi industry) and personal services (in other words, nannies, cleaners) industries. According to Schneider (2002: 6), the total informal economy of South Africa represented approximately 28.4% of the gross national product (GNP) in 2000. Table 1 reflects a comparison of the size of the informal economy across Africa in 2000.

On the basis of the computation by Schneider (2002: 6), the informal sector of South Africa represented approximately 27.7% of the gross domestic product (GDP) of R888 billion in 2000 (SARB 2004). This means that the informal sector was worth approximately R250 billion in 2000. However, Table 1 fails to reflect sector data, arguably because there are limited data on the size of the informal economy by individual sector type. To fill this information gap, the extremely diverse retail infrastructure of South Africa is analysed from a demand-side perspective. Before proceeding with this endeavour, a broad definition of the retail industry of South Africa is provided, followed by a definition of the informal retail trade sector of South Africa.

Table 1: The size of the informal economy for selected African nations (2000)

Africa	% of GNP
Zimbabwe	59.4
Tanzania	58.2
Nigeria	57.9
Zambia	48.9
Benin	45.2
Senegal	43.2
Uganda	43.1
Niger	41.9
Mali	41.0
Ethiopia	40.3
Malawi	40.3
Mozambique	40.3
Cote d'Ivoire	39.9
Madagascar	39.6
Burkina Faso	38.4
Ghana	38.4
Tunisia	38.4
Morocco	36.4
Egypt	35.1
Algeria	34.1
Botswana	33.4
Cameroon	32.8
South Africa	**28.4**
Average	42.0

Source: Schneider (2002: 6)

The formal retail trade sector includes all business enterprises that derive more than 50% of their turnover from sales of goods to the general public for household use (Statistics SA 2003: 12). On the basis on this definition, the retail trade sector includes general dealers (grocers and other dealers in foodstuffs and general department stores), butchers, bottle stores, dealers in clothing, footwear and textiles, dealers in furniture and household requisites, bookstores and stationers, jewellers, chemists and dealers in miscellaneous goods.

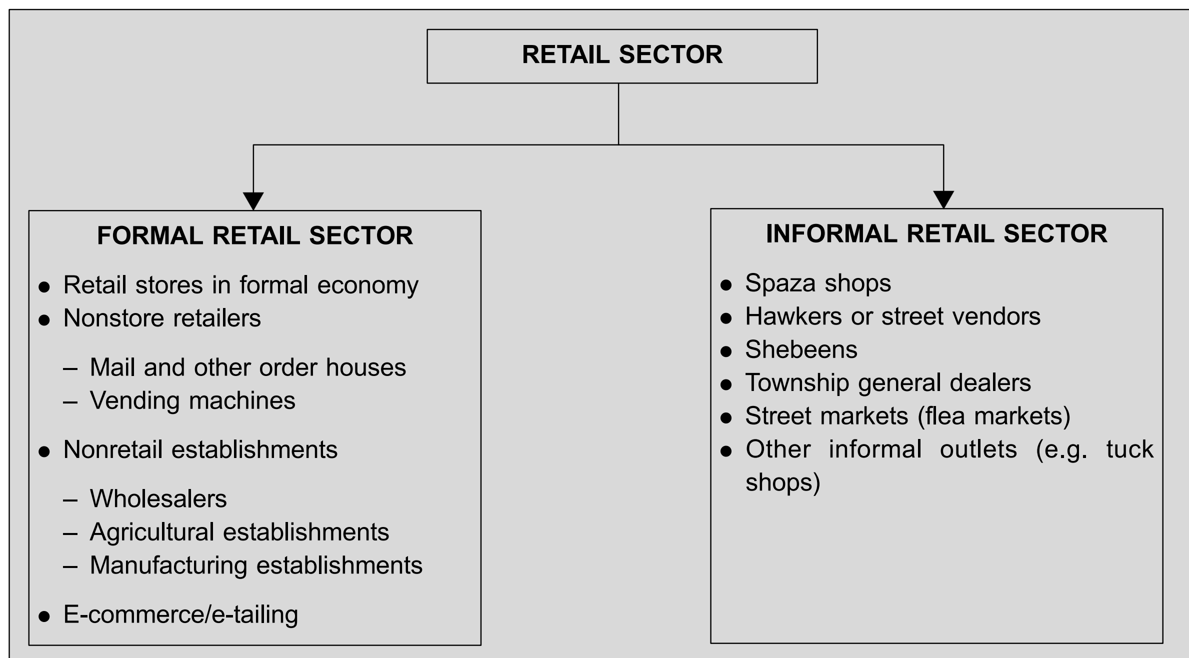


Figure 1: Classification of the retail industry

These retail outlets have established themselves mainly in urban areas, targeting middle- to upper-income households. However, there are also retailers who specifically target the bottom end of the market, which is also serviced by rural grocers, spaza shops and hawkers, among others. The diverse nature of the retail infrastructure in South Africa is further reflected in mail and other order houses, vending machines and nonretail establishments (for example, wholesalers, agricultural and manufacturing establishments) selling directly to consumers, as well as e-commerce. On the basis of the nature of the business, the retail infrastructure in South Africa can be classified into formal and informal retail. This classification is reflected in Figure 1.

On the basis of Figure 1, the informal retail industry of South Africa comprises (Ligthelm & Masuka 2003: 5; Euromonitor International 2002):

- Spaza outlets, prevalent in townships and located in residential homes. These outlets tend to operate in traditional urban black areas and carry a greater range of products than hawkers (see definition below), including canned goods, margarine, tea and coffee, and (with the advent of electricity) frozen chicken and dairy products.
- Shebeens or informal liquor outlets.
- Hawkers or street vendors, who operate from a temporary or permanent structure on a street or at a taxi rank or train station. In general, hawkers seek trading positions in urban areas of high pedestrian and commuter traffic, selling personal care products, basic groceries and fruit and vegetables.
- Township general dealers, who are stand-alone businesses with a brick or mortar superstructure, often located in a business area, but possibly also in residential sections of townships. These businesses generally carry a wider product range than spazas and have more fixtures and fittings, allowing self-catering for clients.
- Flea markets that consist of handmade arts and crafts.

Growth of the informal retail trade

Measuring the growth in informal trade requires the construction of a derived private consumption expenditure model. This model is shown in Figure 2.

The derived private consumption model was constructed according to principles similar to those applied by the South African Reserve Bank (SARB) in constructing national accounts data for South Africa. As a point of departure, the model features current income of households defined as net national income of households and transfers received by households from general government, incorporated enterprises and abroad. Personal disposable income is derived after deducting direct taxes from current personal income. Personal

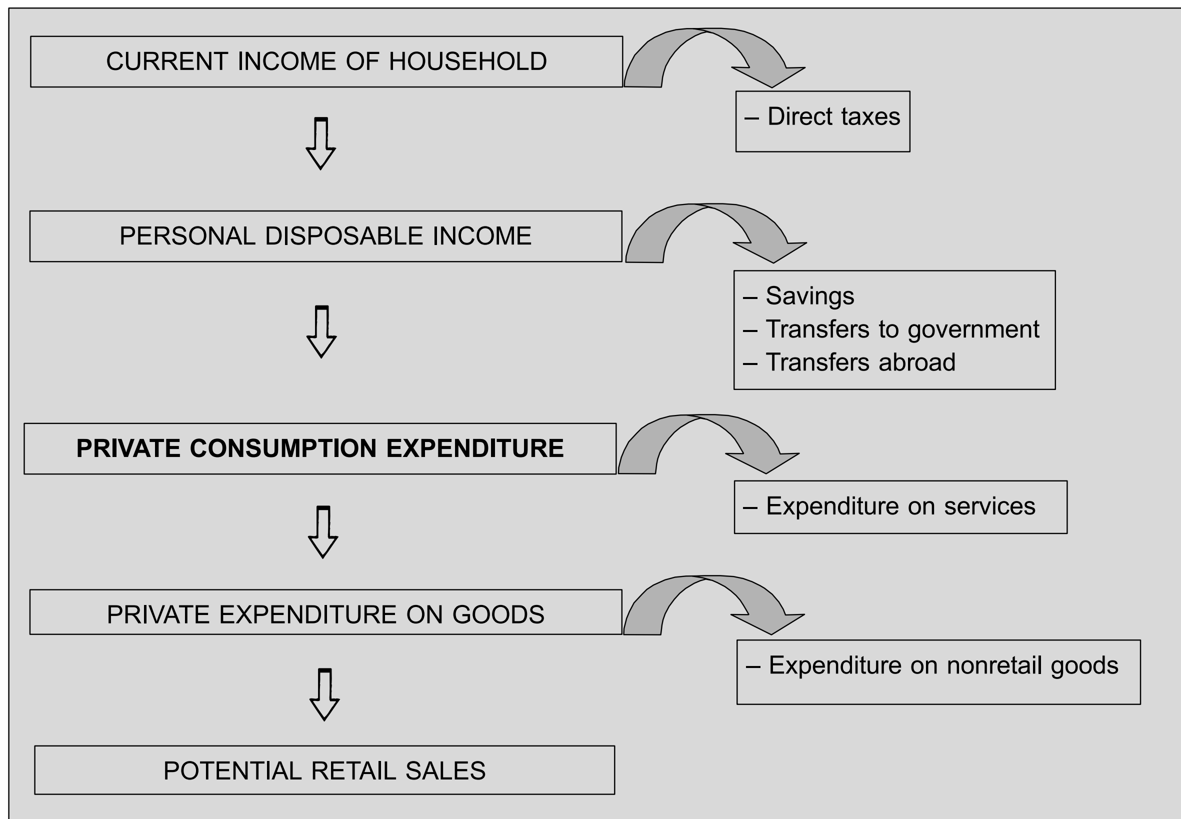


Figure 2: Derived private consumption expenditure model

disposable income is spent on transfers to the general government and the rest of the world, personal savings and private consumption expenditure. The latter is computed by subtracting transfers and personal savings from personal disposable income. Private consumption expenditure refers to all amounts spent on goods and services. However, retail trade, as defined by the Standard Industrial Classification (SIC) (CSS 1993), is concerned only with goods (durables, semidurables and nondurables). Consumer services are offered by institutions other than retail establishments. By deducting services from total private consumption expenditure, private expenditure on goods only is derived. However, expenditure on goods includes goods not classified as retail items by Statistics South Africa. These include personal transport equipment (such as motor cars, motor cycles, bicycles and caravans); motor tyres, parts and accessories; petroleum products; and fuel and power. When expenditure on these nonretail items is deducted from expenditure on goods, potential retail sales are derived. Potential retail sales include retail trade in the formal and informal sectors of the economy. Clearly, the derived private consumption expenditure model is not equipped to measure expenditure patterns for the informal retail trade sector only. This problem can be partially overcome by using available secondary data published by Statistics South Africa on formal retail trade sales with which growth in the informal retail sector can be estimated, as shown in Figure 3.

Retail trade consumer spending patterns by product group

To analyse consumer spending patterns by product group firstly required an investigation into total private consumption expenditure. Total private consumption expenditure data were further disaggregated to determine consumer spending patterns by product group for the formal retail sector only. Finally, consumer spending patterns in the formal retail sector allowed for a final analysis of consumer spending patterns by product group in the informal retail trade sector.

Private consumption expenditure

Private consumption expenditure, as measured by SARB, is defined broadly as the total final expenditure of all households (Investec 2003). This component of the economy represents approximately 62.0% of the gross domestic production of

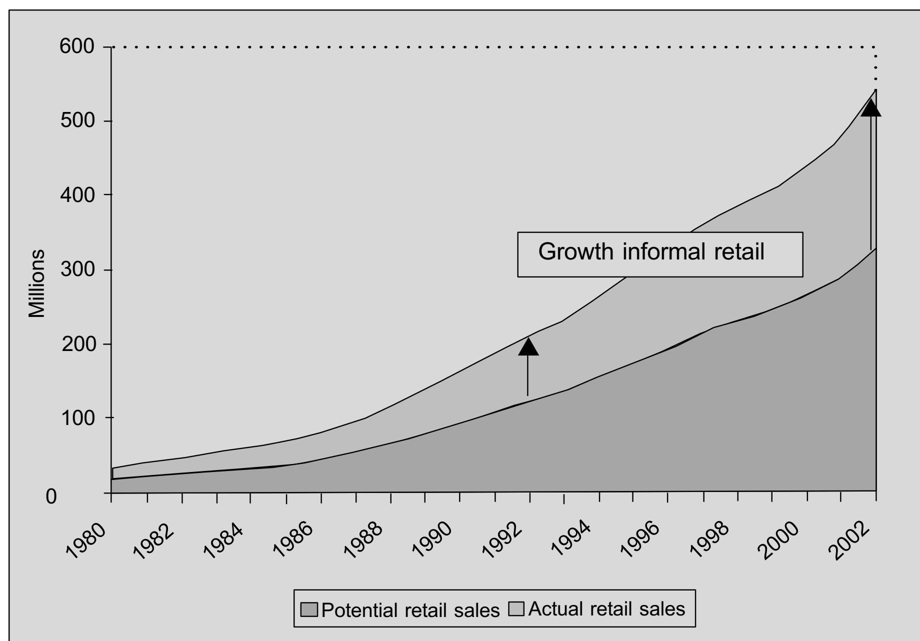


Figure 3: Potential and actual retail sales at current prices (1980–2002)

South Africa (SARB 2004). According to SARB (2004: S116), the components of private consumption expenditure include:

- Durable goods (furniture, household appliances, personal transport equipment, recreation and entertainment goods)
- Semidurable goods (clothing and footwear, household textiles, furnishings, glassware, motor tyres, parts and accessories)
- Nondurable goods (food, beverages and tobacco, household fuel and power, household consumer goods, medical and pharmaceutical products, petroleum products)
- Services (rent, household services, medical services, transport and communication services, recreation, entertainment and educational services).

In the following section, a broad overview is provided of recent expenditure trends for these four private consumption expenditure components.

Structural changes in private consumption expenditure

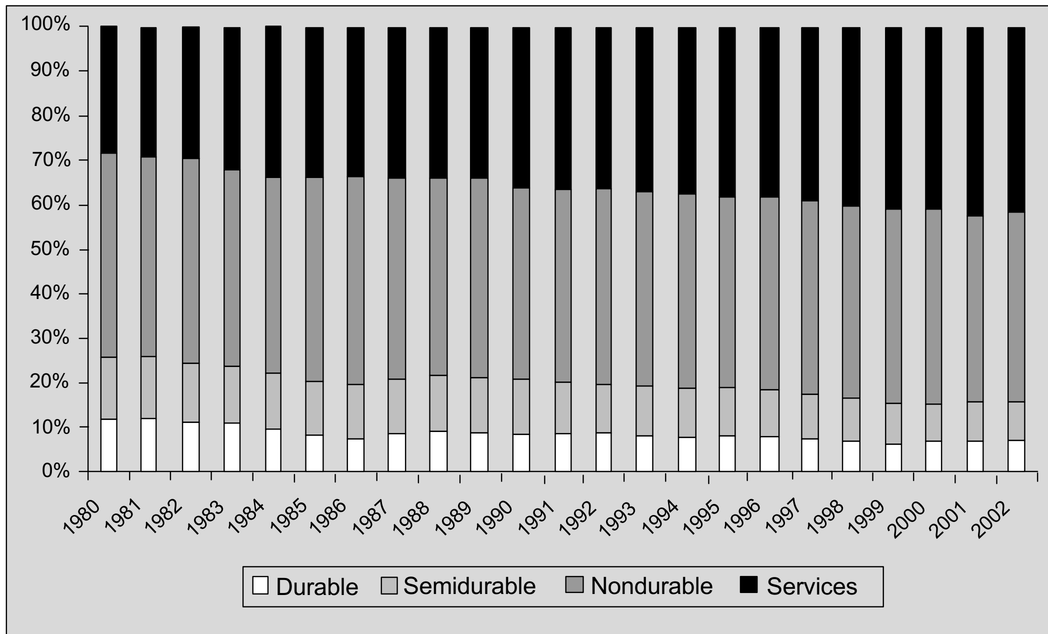
A major structural change witnessed in consumption spending patterns over the past few years is the change in spending patterns of South African consumers towards services, as is the case in other countries (Moran & McCully 2002). These structural changes are reflected in Figure 4.

It is evident from Figure 4 that over the long term, growth in the services component of household consumption expenditure outpaced semidurable, nondurable and durable consumption expenditure. The shift towards services can be attributed partly to the increasing specialisation of functions in a modern economy in which households prefer to use services than to perform tasks themselves. A further reason is the perceived poor quality of certain government services, driving a greater demand for private security, medical and education services. Furthermore, certain services that are relatively new to South Africa, such as gambling and cellular phones, have experienced a dramatic growth in demand over the past few years.

The poor performance of durable goods consumption during the past few years may have much to do with changes in interest rates. High interest rates imply high financing costs for products such as motor vehicles, for which few consumers can afford to pay cash.

The consumption of semidurable goods has been partly buoyed by increased home-ownership and cheap imports replacing domestically produced goods. The financial empowerment of previously economically disadvantaged individuals is also serving to support consumer spending in the semidurable goods category.

Nondurable household consumption, for example food, is more strongly linked to population growth. Given that population growth rates are declining, it



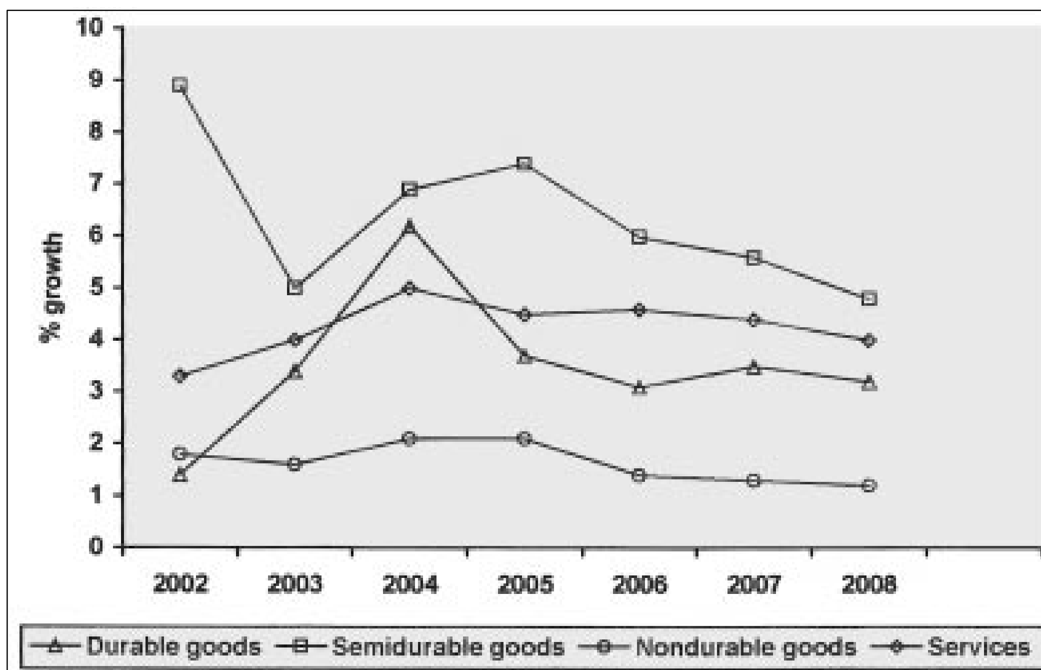
Source: SARB (2004)

Figure 4: Private consumption expenditure trends (1980–2002)

is not surprising that this component has experienced mediocre performance. Furthermore, beverages and tobacco products are experiencing substantial cost increases as a result of increased excise duties, while sales may also be dropping as a result of the displacement of funds to gambling.

To provide a broad overview of future trends within each of the expenditure categories, Figure 5 reflects a breakdown of anticipated expenditure up to 2008.

It is clear from Figure 5 that spending on durables, semidurables and services will continue to increase at the expense of nondurables. Consumer expenditure on durable goods is expected to accelerate by about 3.5% per annum, driven by a lower cost of borrowing, a modest inflation rate and continued financial empowerment of previously disadvantaged individuals. The rising trend in house prices and increased home ownership is anticipated to



Source: Investec (2004)

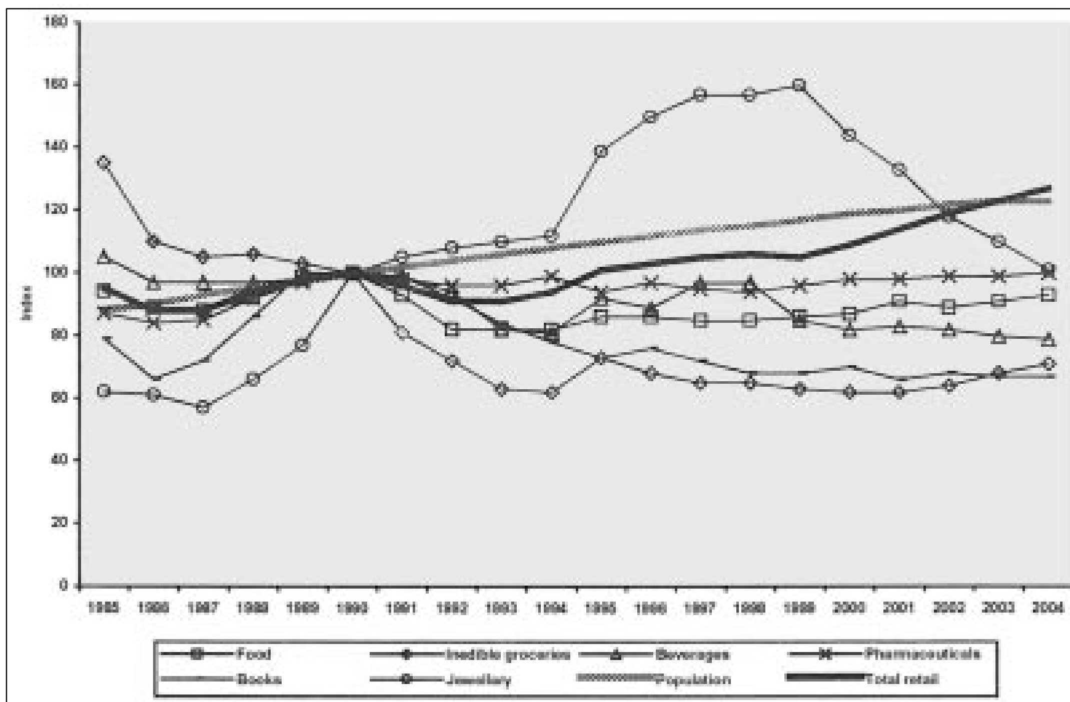
Figure 5: Growth of private consumption expenditure categories (2002–2008)

boost expenditure on services. Spending on transport, medical services (as a result of the prevalence of HIV/AIDS) and cell-phone contracts is likely to underpin growth in expenditure on services over the forecast period.

Formal retail consumer spending

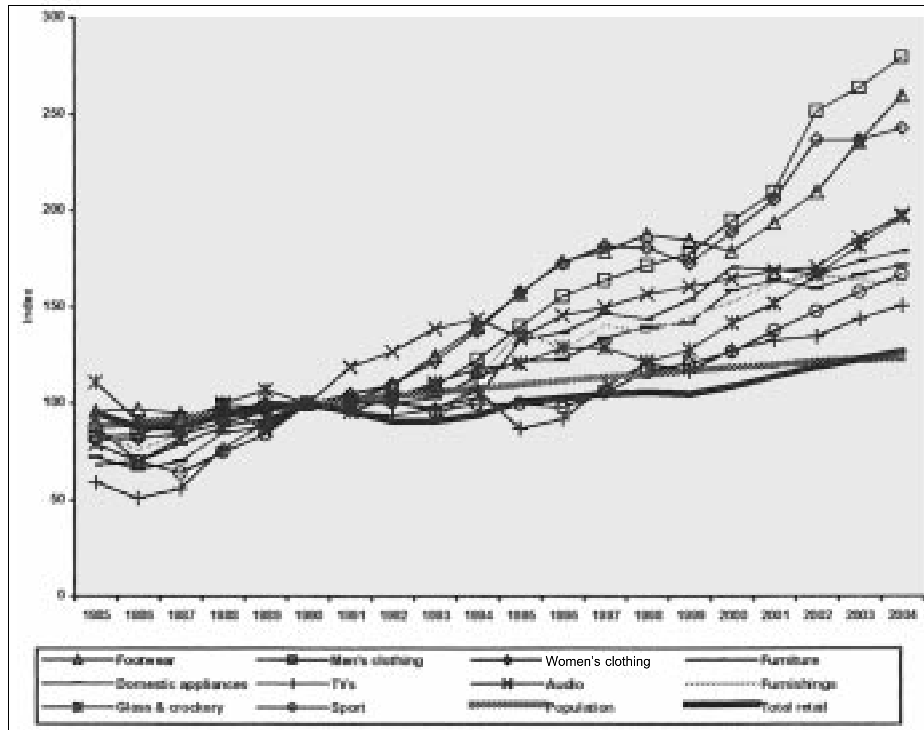
It was evident from the derived private consumption expenditure model that expenditure on services and nonretail trade goods must be subtracted from private consumption expenditure on goods to arrive at potential retail sales (formal and informal retail sales). As a major component of potential retail sales, formal retail sales currently constitute approximately 65% of potential retail sales. However, in 1980, formal retail sales made up approximately 85% of potential retail sales. This structural change in the formal retail sector of South Africa reflects clear long-term changes in consumer expenditure patterns. The changes can be attributed largely to the growing informal retail trade. However, it is evident from Statistics South Africa retail trade data (reflecting figures only for formal retail trade) that consumer spending patterns by product group also show clear trends when analysed over time (Statistics South Africa 2004). These long-term changes in the consumer demand for products sold by formal retail are summarised in Table 2. It should be noted that a so-called 'formal retail trade index' was constructed to determine the long-term changes in consumer demand for formal retail products. The index method used population figures and total formal retail sales figures as a yardstick to measure the relative performance of individual retail trade product groups over time. The base year used for indexing these variables was 1990 (1990 = 100). The outcome of these calculations, reflected in Table 2, is further substantiated by the information reflected in Figures 6 and 7, which provide a more detailed overview of changing consumer demand for products sold by the formal retail industry. Table 2 reflects those products that showed below-average and above-average formal retail trade sales performances over time. These trends typically reflect consumer demand by product group, and ultimately the changes in consumer expenditure over time.

The formal retail trade sales figures of Statistics South Africa do not reflect separate data on telecommunication (cellphone) expenditure. However, this retail sales item could also be classified under 'above average' consumer demand, based on increased cellphone ownership over the past few years. In June 2004, a full 18.7 million South Africans were mobile phone users, while the market is anticipated to increase to 19 million by 2006 (Cellular Online 2004).



Sources: Statistics South Africa (2004); Tustin (2004)

Figure 6: Below average consumer demand for formal retail trade products (1985–2004)



Sources: Statistics South Africa (2004); Tustin (2004)

Figure 7: Above average consumer demand for formal retail trade products (1985–2004)

Table 2: Long-term changes in consumer demand for formal retail products (1985–2004)

CONSUMER DEMAND CHANGES		
Below average demand	**Above average demand**	**Fairly unchanged demand**
Food	Footwear	Textiles
Inedible groceries	Men's clothing	Hardware
Beverages	Ladies' clothing	
Pharmaceuticals	Furniture	
Books	Domestic appliances	
Jewellery	Television sets	
	Audio	
	Other domestic furnishings	
	Glass and crockery	
	Sport and recreation	

There are various explanations for the structural changes in consumer spending patterns in the formal retail trade sector. For example, decreased consumer demand for food products is due mainly to the maturity of the food retail sector, which currently offers limited opportunity for growth. Furthermore, high customs and excise duties over the past few years, as well as increased expenditure on services, has negatively impacted on the formal sales of retail products. In addition, high priced medical products and the growing acceptance of traditional medicine have resulted in moderate sales of pharmaceutical products, while high import prices of books, online purchases (not reflected in recorded sales figures) and aliteracy (among people that can read, but do not) have

Table 3: Percentage distribution of household expenditure by product and retail outlet type

Area / Product	Gauteng (2000)			Durban (2002)			Cape Peninsula (2001)		
	Formal	Informal	Total	Formal	Informal	Total	Formal	Informal	Total
	%	%	%	%	%	%	%	%	%
Food	86.9	13.1	100.0	88.7	11.3	100.0	85.9	14.1	100.0
Clothing, footwear and accessories	96.0	4.0	100.0	94.9	5.1	100.0	89.0	11.0	100.0
Fuel and light	70.2	29.8	100.0	66.3	33.7	100.0	65.4	34.6	100.0
Medical and dental: medicine	95.6	4.4	100.0	92.7	7.3	100.0	96.8	3.2	100.0
Education: stationery and books	94.4	5.6	100.0	94.3	5.7	100.0	95.4	4.6	100.0
Recreation: equipment etc.	97.6	2.4	100.0	98.4	1.6	100.0	97.7	2.3	100.0
Furniture and household equipment	94.1	5.9	100.0	94.8	5.2	100.0	92.5	7.5	100.0
Alcoholic beverages	59.9	40.1	100.0	74.1	25.9	100.0	77.4	22.6	100.0
Cigarettes and tobacco	93.2	6.8	100.0	82.5	17.5	100.0	61.4	38.6	100.0
Washing and cleaning materials	94.7	5.3	100.0	96.7	3.3	100.0	95.1	4.9	100.0
Personal care	84.9	15.1	100.0	94.2	5.8	100.0	92.7	7.3	100.0
Reading matter and stationery	68.9	31.1	100.0	82.4	17.6	100.0	84.7	15.3	100.0
Holiday/weekend: refreshments	100.0	0.0	100.0	83.7	6.3	100.0	98.5	1.5	100.0

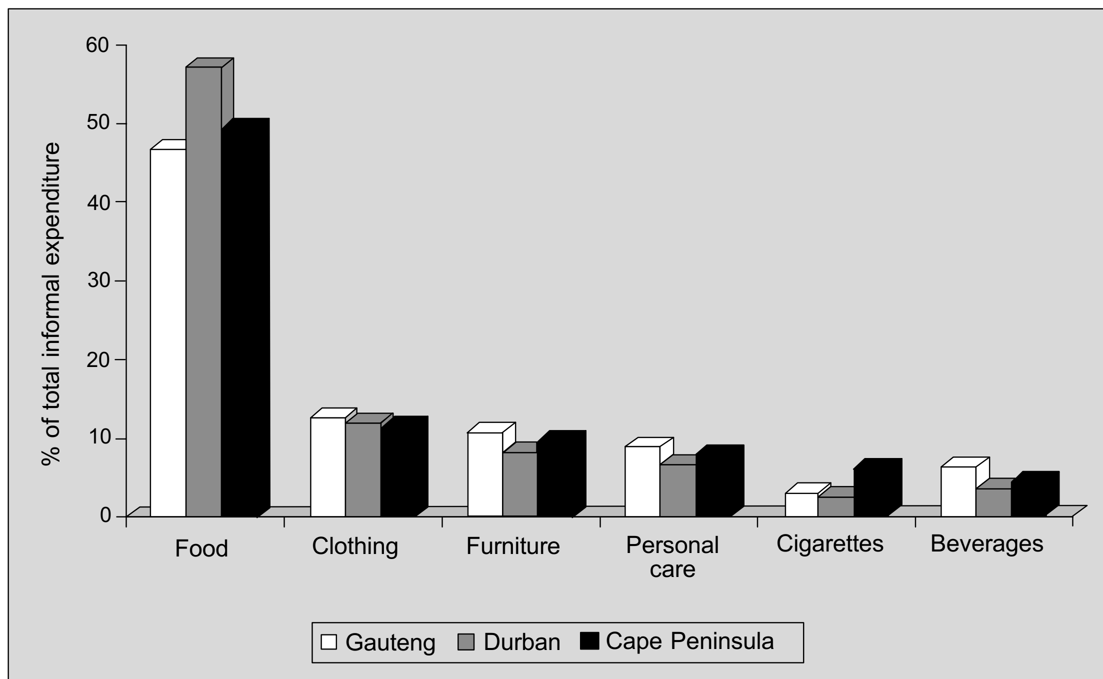
Sources: Martins (2001, 2002, 2003)

contributed to a slowdown in book purchases. Finally, changes in the discretionary purchasing power of consumers regarding services (for example, cellphones) are probably the reason for lower consumer demand for jewellery products.

More lively demand for furniture, domestic appliances, television sets, audio equipment, domestic furnishings and glass and crockery products can be attributed largely to increased home ownership, especially among Africans. Purchases of some of these retail products are also affected by interest rate levels. In addition, low-priced footwear and clothing items have resulted in a recent increase in demand for such items, while the internationalisation of sports events has sparked more lively consumer spending on sport and recreation-related products.

To avoid confusion in analysing the information in Figures 6 and 7, it should be noted that the growing informal retail trade sector has not only managed to attract consumer spending for those formal retail products reflecting long-term below average consumer demand, but increased expenditure at informal retail trade outlets is also evident for those products that show above average consumer demand in the formal retail sector. Evidence from income and expenditure surveys conducted by the Bureau of Market Research (BMR) (Martins 2001, 2002, 2003) proves this notion. Table 3 shows the percentage distribution of household expenditure on products sold via formal and informal retail outlets. For the purposes of comparison, survey data for three experimental survey groups (Gauteng, Durban and the Cape Peninsula) are shown in Table 3. The BMR conducted surveys in these areas in 2000 (Gauteng), 2001 (Cape Peninsula) and 2002 (Durban).

Table 3 shows relatively lively demand for products sold by informal retail outlets, including food (sold by spaza shops and township general dealers), beverages (sold by shebeens and township general dealers), clothing (sold at flea markets), furniture (sold at craft markets), cigarettes (sold by spaza shops and township general dealers) and personal



Sources: Martins (2001, 2002, 2003)

Figure 8: Relative composition of the informal retail trade product basket

care items (sold by spaza shops and township general dealers). To obtain a more detailed picture of the informal retail trade sector, the information reflected in Table 3 was further disaggregated to determine:

- The relative composition of the informal retail trade product basket (Figure 8)
- The average per capita expenditure by outlet and population group (Table 4).

Figure 8 reflects the relative share of products purchased by households from informal retail outlets located in the experimental survey areas. The product basket for each experimental area represents at least 85% of all products purchased by consumers. It is evident that consumer demand in the informal retail sector is the highest for food products – approximately half of the total consumer demand for products in the informal retail sector – followed by clothing, furniture and personal care.

To further explore average consumer spending by population group, the information contained in Table 4 shows the average per capita expenditure at informal retail outlets for the three experimental survey groups.

The average per capita expenditure at informal retail outlets reflects clear differences by population group. In Gauteng, the per capita expenditure figures for Africans and Asians show that just over R1 000 per annum is spent at informal retail outlets. In Durban, the per capita expenditure of Asians at informal retail outlets is slightly higher than for other

Table 4: Average per capita expenditure at informal retail outlets

Population group / Area	Africans	Asians	Coloureds	Whites	Total
	Per capita expenditure (Rand)				
Gauteng	1 297.81	1 056.85	811.33	434.48	1 060.77
Durban	772.01	937.09		246.16	737.19
Cape Peninsula	1 098.99	973.16	1 187.44	602.65	1 026.85
South Africa					**795.08**

population groups. In the Cape Peninsula, per capita expenditure at informal retail outlets is the highest among coloureds.

The average annual per capita expenditure of individuals at informal retail outlets calculated for South Africa as a whole (R795.08 per annum) was computed on the assumption that sales in the informal retail trade sector constitute approximately 12.5% of the potential retail trade sales. On the basis of this estimate, approximately R36 billion is channelled annually through the informal retail trade sector of South Africa.

Although the experimental surveys do not reflect per capita figures for South Africa as a whole, they at least provide some indication of the extent of private consumption expenditure taking place at informal retail outlets. Because geographic surveys of this nature are not conducted frequently, mainly because of cost implications, calculating growth rates for consumer spending at informal retail outlets is almost impossible. Moreover, most income and expenditure surveys of the past have not recorded the activities of the informal retail trade because of its relatively limited economic performance. However, based on information used to construct the derived consumer spending model, it was possible to compute an average annual growth rate for the informal retail trade. According to this calculation, the potential retail trade sector recorded an average annual growth rate of 13.9% from 1990 to 2002. During this period, the average annual growth in formal and informal retail trade sales increased by 9.7% and 14.7% respectively.

Conclusion

Quantifying consumer spending is an important part of tracking consumer behaviour. This is especially important for sustaining a competitive advantage in today's retailing environment. A demand-side analysis of the retail industry has the potential to generate valuable information that will allow retailers, researchers, industry specialists and government to better understand the structural changes occurring in the retail sector of South Africa. As indicated, these structural changes have resulted in changes in consumer demand for goods and services in the retail industry, which is currently showing growing potential for future demand within the informal retail trade industry in particular. These structural changes also present new business opportunities for business planning expansion for entrepreneurs seeking new business opportunities. Seizing these opportunities requires, among other things, sound monitoring of consumer and business trends that might impact on and necessitate adjustments to strategic business planning.

References

Bonnie, S.M. & Van Dyck, G. 2003. Branding without labels – dynamics of the African traditional market. Paper presented at 5th PAMRO annual meeting. Johannesburg: SAARF [Online]. Available: www.pamro.org. Accessed: January 2004.

Bono, M. 2003. *Participation in the South African Informal Economy: An Indicator of Poverty?* University of Michigan.

Cellular Online. 2004. *Mobile Stats*. [Online.] Available: www.cellular.co.za. Accessed: November 2003.

Central Statistical Service (CSS). 1993. *Standard Industrial Classification of all Economic Activities*. 5th edition. Pretoria: CSS.

Euromonitor International. 2002. *Retail Trade International: South Africa,* Executive summary: retail sales in the macroeconomic context. November. [Online]. Available: www.euromonitor.com/Retail_-Trade_International_-_South_Africa. Accessed: January 2004.

Feige, E.L. (ed.) 1989. *The Underground Economies: Tax Evasion and Information Distortion.* Cambridge: Cambridge University Press.

Feige, E.L. 1994. 'The underground economy and the currency enigma', Supplement to *Public Finance/ Finance Publiques,* 49: 119–136.

Frey, B.S. & Pommerehne, W. 1984. 'The hidden economy: state and prospect for measurement', *Review of Income and Wealth,* 30(1): 1–23.

Investec. 2003. *SA Economic Research 2003/07: Outlook for Consumption Expenditure.* Second quarter. Johannesburg.

Investec. 2004. *SA Economic Research: Prospects 2003–2008 – Consumer Expenditure Outlook.* January. Johannesburg.

Ligthelm, A.A. & Masuku, T. 2003. Size, structure and profile of the informal retail sector in South Africa. Research report no. 323. Pretoria: Unisa, Bureau of Market Research.

Martins, J.H. 2001. Total expenditure of households in Gauteng by expenditure item and type of outlet, 2000. Research report no. 289. Pretoria: Unisa, Bureau of Market Research.

Martins, J.H. 2002. Expenditure of households in the Cape Peninsula by expenditure item and type of outlet, 2001. Research report no. 304. Pretoria: Unisa, Bureau of Market Research.

Martins, J.H. 2003. Total expenditure of households in the Durban Metropolitan Area by expenditure item and type of outlet, 2002. Research report no. 321. Pretoria: Unisa, Bureau of Market Research.

Moran, L.R. & McCully, C.P. 2002. *Trends in Consumer Spending, 1959–2000*. Washington: Bureau of Economic Analysis, US Department of Commerce.

Schneider, F. 2002. Size and measurement of the informal economy in 110 countries around the world. Linz, Austria: Johannes Kepler University of Linz, Department of Economics.

South African Reserve Bank (SARB). 2004. *Quarterly Bulletin time series*. [Online]. Available: www.reservebank.co.za. Accessed: January 2004.

South African Reserve Bank (SARB). Economic Department. 2004. *Quarterly Bulletin,* 229, September. Pretoria.

Statistics South Africa. 2003. Retail trade sales. Statistical Release P6242.1, S3, December. Pretoria: Statistics South Africa.

Statistics South Africa. 2004. Time series data. [Online]. Available: www.statssa.gov.za. Accessed: January 2004.

Tustin, D.H. 2004. Forecast of economic indicators and formal retail sales by product group for 2004. Research report no. 324. Pretoria: Unisa, Bureau of Market Research.

Evaluation of the need to introduce a system of group taxation in South Africa

*J.S. Wilcocks & S.N. Middelmann**

In a group structure, individual companies comprising a group are effectively managed as a single 'economic unit'. The economic unit concept refers to a group of companies that are collectively integrated on a financial, an organisational or an economic basis by virtue of common control, so that they are working towards a common purpose or goal. The South African income tax dispensation currently makes no provision for group taxation. Each legal entity within a group is taxed as a separate taxpayer. This study aims to evaluate whether there is a need for a system of group taxation in South Africa. In order to do so, the definition of a group was considered, the different tax treatment of divisions as opposed to a group structure were investigated, the current income tax dispensation for inter-group transactions was analysed, and anomalies arising from that were highlighted. The recommendations of the Katz Commission in its Third Interim Report, which addressed the issue of group taxation, were also examined to determine whether the report supports the implementation of a system of group taxation in South Africa. The impact of a system of group taxation for meeting the requirements of the canons of taxation, as well as the implications for the fiscus and the taxpayer, were also examined. The analyses and the conclusions clearly show that the status quo with regard to the inherent tax anomalies arising from the taxation of intra-group transactions is unsustainable, and that a system of group taxation should be implemented in South Africa.

Introduction and problem statement

According to the Katz Commission (South Africa 1995: 96), in a group structure, the individual companies comprising a group are effectively managed as a single 'economic unit'. The economic unit concept refers to a group of companies that are collectively integrated on a financial, an organisational or an economic basis by virtue of common control, so that they can be said to be working towards a common purpose or goal (Howitt 1992: 2). A group of companies is managed (including the strategic and financial decision-making) in the interests of the group as a whole (South Africa 1995: 96). Currently, the South African income tax dispensation makes no provision for a system of group taxation (Kannenberg 1999: 1). Each legal entity within a group is taxed as a separate taxpayer, in terms of section 5(1)(d) of the Income Tax Act (Act 58 of 1962).

The introduction of the Corporate Rules as Part III of the Income Tax Act by the Second Revenue Laws Amendment Act (Act 60 of 2001), promulgated on 12 December 2001, provided some relief in respect of transactions between group companies and between founding shareholders and their company (Huxham & Haupt 2004: 254). These measures are generally based on the view that where the group has retained a substantial interest in the assets that are transferred, it is appropriate to permit a tax-free transfer of assets to the entity in the group where they can be most efficiently used for business purposes (Department of Finance 2001: 6). These Corporate Rules provide some relief with regard to asset and share transactions for group companies, and they might be the first step in the direction of a group tax system, but no relief is provided for other important day-to-day transactions, such as timing mismatches in transactions between group companies.

When a single economic unit is treated as several separate tax entities, this encourages some manipulation of the taxable income of the various entities within the group, by means of profit shifting, in order to reflect the economic reality of the group's results. This also encourages groups to enter into elaborate tax schemes that either defer taxable income, or

* J.S. Wilcocks is a senior lecturer at the School of Applied Accountancy, University of South Africa. E-mail: wilcojs @unisa.ac.za. S.N. Middelmann is an independent consultant and a masters student in the Department of Taxation, University of Pretoria.

accelerate tax deductions, or change the tax base of the underlying asset in order for the group to obtain an overall tax benefit. A system of group taxation should act as a deterrent to such manipulation and schemes, as it disregards all intra-group transactions for tax purposes and eliminates tax anomalies that arise from intra-group transactions. Management is therefore less likely to be influenced by the tax effects of intra-group transactions, and more likely to concentrate on promoting growth in economic activity, thereby increasing the general tax base (South Africa 1995: 97; Middelmann 2003: 3).

It is clear that merely by changing the legal form of a business operation, without altering its economic substance, very different tax effects can be achieved. According to the Margo Commission (South Africa 1986: 199), divisionalisation is not always ideal because of issues such as the advantage of limited liability, compliance with certain regulated industry requirements, and certain rights and licences that are exclusive to a particular entity. Hence, separate legal entities will continue to exist, and group structures will continue to be present in our economy. A taxation system that taxes a group of companies as an economic unit is therefore essential for an equitable and sound system of taxation.

Group taxation is a system of taxation whereby a group of companies is effectively taxed as a single economic unit and the tax liability is met by a representative member. There are two principal systems of group taxation (South Africa 1995: 98):

- The loss transfer regime allows for a tax loss incurred by one company within a group to be set off against the income derived by another company, or other companies, within the same group. Each company in the group retains its own personality.
- The consolidation regime treats a group of companies as a single taxpayer. It effectively neutralises the tax effect of intra-group transactions, much like the consolidation process for accounting.

A system of group taxation for South Africa has been considered on various occasions. In 1986, the Margo Commission (South Africa 1986) investigated the issue and recommended that a system of group taxation should not be implemented in South Africa. The reasons advanced for not implementing such a system of taxation included loss of significant revenue to the state, the fact that minority shareholders could be prejudiced, as well as the complexity of the system and an increased administrative burden (South Africa 1986: 200–201). However, the majority of the bodies and corporations that made written representations to the Margo Commission argued strongly in favour of group taxation (South Africa 1986: 199).

The matter was subsequently investigated by the Katz Commission, which expressed the following opinion: "The Commission is mindful of the view amongst some that the issue of group taxation is not a priority. It disagrees with this view, and regards the current position as a structural defect in the system that cannot be passed over in any serious tax reform process" (South Africa 1995: 96).

The South African Chamber of Business (SACOB) fully supported this recommendation by the Katz Commission. SACOB acknowledged that the costs related to and the complexities of introducing a system of group taxation could not be ignored, but argued that these problems could be overcome. SACOB also stated that it believed that the benefits of adopting a system of group taxation significantly outweigh the disadvantages. The adoption of such a system would achieve greater fiscal control, minimise some of the economic distortions that currently exist at a corporate level, facilitate the corporate unbundling process and bring South Africa into step with the tax treatment of companies in industrialised countries (SACOB 1996: 4–5).

Although the proposal was accepted in principle by the South African Revenue Service, the decision to introduce a system of group taxation was held in abeyance until the new South African Revenue Service was fully operational (South Africa 1996: 2–25).

Research objective and methodology

The objective of this study is to evaluate the need to introduce a system of group taxation in South Africa. In order to achieve this objective, the following process was followed:

- The definition of a group was analysed.
- The anomalies arising from the different tax treatment of divisions as opposed to a group structure were investigated.
- The current income tax dispensation for inter-group transactions was analysed, and anomalies arising from that were highlighted.
- The recommendations made by the Katz Commission in its Third Interim Report, presented in

December 1995, which addressed the issue of group taxation, were examined to determine whether the Commission supports the implementation of a system of group taxation in South Africa.

- The effect of a system of group taxation in satisfying the requirements of the canons of taxation was examined.
- The impact of the implementation of a system of group taxation on the fiscus and the taxpayer was examined.

The study consisted of a review of relevant literature. The literature that was consulted included tax legislation, textbooks, studies undertaken by local and overseas research institutions and the respective commissions, articles published in legal and business journals and relevant court cases. The information was summarised, documented, evaluated and, where appropriate, examples were included to convey the issues clearly.

This study addresses group taxation only at a conceptual level. The focus is on broad principles and issues, rather than detailed design and implementation. The study does not include an in-depth analysis of the administrative issues that may also need to be considered in a further study. The study is limited to income tax in terms of the Income Tax Act (Act 58 of 1962), including legislation promulgated up to 31 December 2003 (Revenue Laws Amendment Act, Act 45 of 2003). A detailed discussion of the Corporate Rules (sections 41 to 47 of the Act), which were recently introduced into South African income tax legislation, falls outside the scope of this study. Donations tax (sections 54 to 64 of the Act) is also not referred to in this study, as donations made by a company to any other company that is a member of the same group of companies are exempt from donations tax (section 56(1)(r) of the Act). Secondary tax on companies (STC) (sections 64B and 64C) is also not covered in detail, as a company can elect that dividends declared by the company to a shareholder (a resident company) that forms part of the same group of companies be exempt from the payment of STC. The dividend should, moreover, be paid out of profits earned during the period when the shareholder formed part of the group (section 64B(f) of the Act).

The study deals only with groups of companies that are all registered, managed and controlled in South Africa. Aspects relating to transfer pricing and thin capitalisation (section 31 of the Act), as well as other aspects of international tax, have been ignored. The study does not include indirect taxes (such as value added tax, regional services levies, transfer duty and stamp duty).

Defining a group

Previously, a group of companies was defined in the Companies Act (Act 61 of 1973) as consisting of a holding company (which was not itself a wholly owned subsidiary) and its subsidiaries (paragraph 4(q) of Schedule 4 of the Companies Act of 1973). A company is a subsidiary company of another company in terms of the Companies Act (section 1(3)), if

- that company is a member of it and holds a majority of the voting rights; or
- it has the right to appoint or remove directors that hold the majority of the voting rights (at meetings of the board); or
- it has the sole control of a majority of the voting rights in it, in terms of an agreement with other members or otherwise.

With the introduction of the Corporate Rules (sections 41 to 47 of the Income Tax Act), the following definition of a 'group of companies' was also incorporated into section 1 of the Income Tax Act:

> '**group of companies**' means two or more companies in which one company (hereinafter referred to as the '**controlling group company**') directly or indirectly holds shares in at least one other company (hereinafter referred to as the '**controlled group company**'), to the extent that —
>
> (a) At least 75 per cent of the equity shares of each controlled group company are directly held by the controlling group company, one or more other controlled group companies or any combination thereof; and
>
> (b) The controlling group company directly holds 75 per cent or more of the equity shares in at least one controlled group company

The definition of a 'group of companies' in the Companies Act differs substantially from that in the Income Tax Act. According to the Companies Act (paragraph 4(q) of Schedule 4), a company is a member of a group if the holding company holds the majority (more than 50%) of voting rights, but in

terms of the Income Tax Act (section 1), a company is a member of a group if 75% of the equity share capital is held by the controlling group company. The Companies Act recognises control in terms of voting rights, while the Income Tax Act focuses on control in terms of shareholding. The definition set out in the Income Tax Act is the one preferred for the purposes of this study.

The Katz Commission (South Africa 1995: 101–102) stated in its report that, in order to qualify for group relief provisions, a group should comprise a holding company and all its wholly owned subsidiaries (limited to South African companies, but excluding close corporations, because the company law requirements are stricter for companies). The test for whether a company is wholly owned is determined with reference to interests, direct or indirect, in the equity share capital of the companies concerned. Equity shares (to a maximum of 10% of the total equity share capital of the company) held by full-time employees (including executive directors) in terms of a share incentive scheme should be taken into account for the wholly owned test. However, the Katz Commission (South Africa 1995) also stated that once a consolidation system has been successfully implemented, the ownership requirement could possibly be reduced to, say, 75% of the equity share capital. The main reason for the proposal that only wholly owned groups qualify for group tax relief is reduced cost and complexity (South Africa 1995: 100).

The essence of all these definitions of a holding company and its subsidiaries is therefore one of control (whether in terms of voting rights or shareholding). According to Cilliers, Benade, Henning, Du Plessis & Delport (1992: 432), the basic characteristic of a group is that the management of the various independent holding and subsidiary companies comprising the group is coordinated in such a way that management takes place on a central and unified basis in the interests of the group as a whole. This is due to the control implicit in the relationship between the holding and the subsidiary company or companies. This control makes it possible for the group to be managed as an economic unit. Although, in terms of company law, each company within a group is a separate legal entity, the courts have dealt with the group as a whole as an economic entity. This piercing of the corporate veil is indicated especially where a holding company has 100% holding and control of its subsidiary, and can therefore control every aspect of the subsidiary (Cilliers et al. 1992: 435).

If the South African sensitivity towards a concentration of economic power is taken into account, a system that requires 100% ownership within the group might not be acceptable. Minority shareholders should be accommodated within a group tax regime. However, to align the group with a divisionalised company, a minimum intra-group interest of 75% should be prescribed. This is in line with the current requirements of the Income Tax Act. It ensures that the holding company can still pass special resolutions (which will enable the holding company to manage the group in much the same way as a divisionalised company) and is substantial enough to qualify the holding company conceptually as the effective owner of the subsidiary's business (Kannenberg 1999: 3).

Divisionalisation versus a group structure

A divisional structure refers to a structure where separate businesses are housed in separate divisions within one company. From a legal point of view, this type of structure comprises a single legal entity. Transactions between the individual divisions are effectively ignored when reporting at the company level for accounting and tax purposes.

This differs from a group structure, as defined in the previous section. The group structure comprises separate legal entities within the group. Each entity houses one or more of the various businesses. Transactions between these separate legal entities have an effect for tax purposes that does not always correspond with the accounting treatment of these transactions. For accounting purposes, these intra-group transactions are effectively set off on consolidation, and give rise to similar results as transactions between divisions within the same company.

This can be clearly demonstrated where, for example, a business has generated a tax loss:

- If the business is structured as a division within a company, this tax loss is set off against the taxable income of the other divisions within the same company to derive the aggregate taxable income or loss for the company (where the legal entity is the taxpayer). As divisions within a single company do not each have a separate legal identity, the accounting and tax treatments will correspond for the legal entity.
- If the business is structured as a separate company within a group, the loss cannot be set off against the taxable income of the other group companies. For tax purposes, the group has to

pay tax in respect of each of the tax-paying entities within the group. The benefit of the tax loss can only be utilised in future when that specific entity has generated taxable income. The economic substance is that the group as a whole has made a net loss or smaller net profit before tax, but tax is levied on greater profits (which is illustrated when the results are consolidated for the compilation of the financial statements).

This anomaly often leads to 'financial engineering', that is, artificial manipulation of the affairs of companies in order to minimise the tax liability of the group as a whole. The principle of 'financial engineering' is usually altogether unproductive, and the efforts expended by executives in pursuit of these alleged benefits could be more profitably spent in more productive areas (*Taxpayer* 1985: 170).

According to the Margo Commission (South Africa 1986: 199), an argument that is often put forward against group taxation is that the tax effects of intra-group transactions can be neutralised by implementing a divisionalised structure within a single entity, as opposed to a group structure. Apart from the tax considerations, the compliance and administration costs of a divisional structure are claimed to be considerably less. In spite of these advantages, there may be sound commercial reasons why separate legal entities are required as opposed to a divisional structure.

Some of these considerations are cited in the Margo Commission report on group taxation (South Africa 1986: 199), including the following:

- The retention by companies in the group of valuable licences and rights that would lapse when these entities cease to exist
- The requirement by certain regulated industries that operations must be kept in separate entities
- The compliance of loan covenants and agreements that may restrict a transfer of assets
- Strategic business reasons that may include future listing, new risk ventures or foreign investment opportunities
- The protection of limited liability
- The need of new risk ventures for protection of limited liability
- The attraction for foreign investors who may wish to incorporate separate companies for specific operations.

It is therefore clear that in practice there may be legitimate commercial reasons for businesses to be housed in separate legal entities rather than in one divisionalised company. However, this does not detract from the fact that, from an ownership point of view, the entities within a group are managed as a single economic unit.

It is clear from the foregoing considerations that both group and divisional structures will continue to exist for reasons other than the tax implications of such structures. It is therefore important that the tax anomalies arising from intra-group transactions inherent in our current tax system be addressed.

Anomalies arising from the current tax treatment of intra-group transactions

This section analyses the tax effects of certain intra-group transactions and compares these to the tax effects that arise where the same transactions are carried out between the different divisions of the same company, in order to illustrate the need for the implementation of a group tax system.

A transaction between two parties generally gives rise to an expense or liability in the hands of the one party and income or an asset in the hands of the other party. For tax deduction purposes, the expense side of the transaction must meet the provisions of section 11(a) of the Income Tax Act (Act 52 of 1962, as amended, read together with section 23, particularly section 23(g), which prohibits a deduction to the extent that the deduction is not laid out or expended for trade purposes). In terms of section 11(a) of the Act, expenditure and losses may be deductible if they meet the following requirements:

- They have actually been incurred
- They have been incurred in the production of that entity's income
- They are not of a capital nature
- They are derived from carrying on a trade.

For the income to be recognised as gross income for tax purposes, the requirements of the gross income definition in terms of section 1 of the Act must be met. For income to be taxable, it must fall into the gross income definition in section 1. The requirements are that there should be:

- A total amount
- In cash or otherwise
- Received by or accrued to or in favour of
- During the year of assessment
- Excluding receipts of a capital nature.

However, as discussed in the following sections, income or gains from the disposal of qualifying assets (which do not constitute gross income as they are of a capital nature) are subject to the provisions of the Eighth Schedule of the Act. The taxable gains in respect of the disposal of these assets are included in taxable income in terms of section 26A of the Act.

For the purposes of this study, only the requirements pertaining to a resident of South Africa have been considered.

Anomalies could arise from the tax treatment of transactions between companies within the same group. For the purposes of this paper, the following categories have been identified:

- The timing mismatches of income and expenditure between group companies
- The taxability versus deductibility mismatches of certain transactions because the requirements of the deduction provisions are not met or do not correspond
- The taxability versus deductibility mismatches due to capital versus revenue differences of the same transaction
- A shift of income and expenditure between taxable and loss-making entities
- Capital gains tax and other tax consequences arising on certain intra-group transactions.

Timing mismatches of income and expenses between group companies

A timing mismatch of income and expenditure may result in income's being taxed in one year in the hands of one party, while the corresponding expenditure may only be deductible in the hands of the other party in a subsequent year. This is demonstrated by the following example:

Company A and Company B form part of the same group of companies. Company A makes an advance payment to Company B for administrative services to be rendered for the whole of the next year. This payment occurs at the year-end (Year 1), at which point, Company B has not rendered any services.

Company B is taxed on the full receipt in Year 1, as this payment constitutes gross income in its hands, since it is an amount actually received. Company B is not able to claim a deduction against this income, as it has not incurred any expenditure in producing this fee income by the year-end. Company A is not able to claim a deduction for the payment in Year 1, as the deduction is limited in terms of the provisions of section 23H of the Act (if the payment exceeds R50 000). Although the expense has been incurred by Company A, as the fee has been paid – based on the judgement in *Caltex Oil (SA) Ltd v SIR* (1975(1) SA 665 (A)) – section 23H of the Act limits the deduction in respect of any expenditure incurred where the related benefits are received over a period longer than six months. Effectively, where the related benefits are received over a period of more than six months, the expenditure is only deductible as and when the benefits are received. Company A can therefore only obtain the deduction in the subsequent year, when the services have actually been rendered by Company B. Although the inclusion of the receipt, in Year 1, in Company B's taxable income is followed by a deduction in the subsequent year for Company A, from the group's perspective, this has resulted in a negative cash outflow in the first year, as a result of the tax payable by Company B on the administration fee received.

The foregoing transaction would have no net accounting or tax effects between divisions within the same company.

Mismatch arising from not satisfying the requirements of sections 11(a) and 23(g)

The tax effects of intra-group transactions may be mismatched where the requirements of section 11(a), read together with section 23(g), are not met on the expenditure side of the transaction, although the gross income definition in terms of section 1 of the Act is applicable in respect of the income side of the transaction. Two main types of instance may arise.

The first scenario arises where the group company that incurs the expense or loss side of the transaction may not deduct the expense or loss in determining its taxable income, as the expense or loss does not meet the requirements of section 11(a), read with section 23(g), but the corresponding receipt or accrual is taxed in the hands of the group company receiving it. From a group perspective, this results in an inconsistency. This transaction's economic benefit has a zero-sum effect, whereas, from a tax perspective, it has created taxable income in the hands of one group entity, without the corresponding relief of a deduction in the hands of the group entity that has incurred the expense.

The requirements of sections 11(a) and 23(g) have already been set out. For the expense or loss to be deductible, it must be a 'non-capital expense' or

loss actually incurred in the production of income derived from trade. Over the years, the courts have established the meaning of this phrase. In *Port Elizabeth Electric Tramway Co Ltd v CIR* (1936 CPD 241), Judge Watermeyer AJP stated that, in the ordinary sense, this phrase does not refer to expenditure that produces income, but rather to business operations, and that expenditure is attendant upon these operations. He also stated that if a business operation is conducted *bona fide* for carrying on a trade that earns income, the expenditure is deductible, as it is incurred in the production of income derived from a trade. How closely linked the expenditure must be to the operations was formulated by Judge Watermeyer AJP in that case as follows:

> ... in my opinion, all expenses attached to the performance of a business operation *bona fide* performed for the purposes of earning income are deductible whether such expenses are necessary for its performance or attached to it by chance or are *bona fide* incurred for the more efficient performance of such operations provided they are so closely connected with it that they may be regarded as part of the costs of performing it.

From this case, it is clear that, as a general rule, ordinary business expenditure (as well as expenditure necessary for the performance of the business operation), because it is not of a capital nature, satisfies the production of income and trade requirements. It may be difficult for the remaining category of business expenses, namely expenses attached to business operations by chance, to satisfy the requirements in question.

In the case of group companies, this may happen where an incidental payment made by one group company to another group company is not deductible by the group company that incurs the expense, but the amount is taxed in the group company to which it accrues. An example of such expenditure is a payment for damages or negligence. In this instance, the transaction for the group is not tax neutral. The amount is taxed in the group company to which it accrues or by which it is received, but is not deductible by the group company incurring the expense, as it may not satisfy the production of income and trade tests.

The second instance, although it is similar to the first, in that the requirements of section 23(g) or section 11(a) are not met, is different in that the expense or loss is not deductible by the group company that incurred it, as the expenditure is not related to the production of income from its trade, even though from a group perspective it may have been deductible in relation to another group company, as the expenditure may have been related to the production of income from trade by another member of the same group. The corresponding receipt or accrual is taxed in the hands of the group company that receives the payment. This gives rise to the same tax effect as already discussed, namely, that the transaction is not tax neutral for the group as a whole (Kannenberg 1999: 23–25). This happens if, for example, Company A (a subsidiary in the group) does all the administration for the group, and charges all the group companies for administration services rendered. Company B (the holding company) may earn only exempt income (dividends) and not be able to deduct the administrative fee charged by Company A, as the expense is not related to the production of income, in terms of section 11(a), read together with section 23(f) of the Act. (It should be noted that the 'trade' requirement set out in section 11 of the Act is also not met.) If, however, this amount were allocated to other companies in the group that do earn taxable income, it would be deductible.

From the preceding, it is clear that, in some instances, intra-group transactions give rise to anomalous tax effects because each group company is treated as a separate taxpayer.

Mismatch of the capital versus the revenue nature of transactions between group companies

The requirements of the gross income definition and general deduction formula imply that an income or expense may not be of a capital nature in order for the amount to be taxable or deductible. However, receipts or accruals of a capital nature are dealt with under the provisions of the Eighth Schedule and, where applicable, 50% (for a company) of the capital gain may be subject to tax.

Over the years, the courts have established clear guidelines in determining what is regarded as an amount of a capital nature and what is regarded as an amount of a revenue nature. From a gross income perspective, a receipt or accrual is either capital or revenue; there is no half way house between the two (*Pyott Ltd v CIR* (1944 AD 610)). Generally, a revenue receipt is income that arises from a business enterprise or activity, personal exertion, or the employment of capital, either by using it or by letting it (Huxham & Haupt 2004: 22). In distinguishing between receipts of a capital nature and those of a revenue nature, an analogy

is often used – in *CIR v Visser* (1937 TPD 77), Maritz J stated: ''Income is what capital produces, or is something in the nature of interest or fruit as opposed to principal or tree.''

Over the years, the courts have laid down various tests to be applied in deciding whether a receipt is revenue or capital. The dominant test is that of the intention of the taxpayer (Arendse, Coetzee, Jordaan, Kolitz, Stein & Stiglingh 2004: 21; Huxham & Haupt 2004: 23), which was originally referred to in *CIR v Stott* (1928 AD 252). Various factors would influence the determination of the intention of the taxpayer, including whether there has been a change of intention between the time when an asset was acquired and the time when the asset was disposed of. This test was further expanded by the courts, to include the test that had its origins in the *California Copper Syndicate* case (1904), as to whether the intention of the taxpayer was that of a scheme of profit-making.

From the general deduction perspective, an amount is not deductible if it is of a capital nature. The main test for determining the capital or revenue nature of an expense or loss was established in *New State Areas Ltd v CIR* (1946 AD 610). If the expense or loss is incurred as the cost of performing the income-earning operations of the taxpayer, it is by nature revenue. If it is part of the cost of establishing, enhancing or adding to the taxpayer's income-earning structure, it is of a capital nature.

Although the broad principles used to distinguish between the capital and revenue nature of expenditure and receipts or accruals are similar, the application of these principles may result in the opposing sides of a transaction being treated inconsistently. This can be illustrated by means of a simple example:

A group of companies includes two subsidiaries. One is a property investor receiving rental income, and the other is a property developer. When the property developer develops and sells a property to the property investor, the developer is taxed on the receipt resulting from the sale of the property to its subsidiary, as the property is regarded as its stock in trade. The property investor in turn may not receive any tax relief on the acquisition cost of the property if the property does not qualify for any capital allowances, as is the case, for example, with office accommodation and shopping malls. The group has experienced a zero net cash flow, but has been taxed on the profit portion of the development.

It may therefore happen that the expenditure component of an intra-group transaction is viewed as a capital expenditure, while the receipt or accrual component is viewed as income. When parties to a transaction are viewed in isolation, it is possible that different facts and circumstances will be considered, or that these may be interpreted differently, resulting in anomalous results in respect of each component of the transaction.

Shifts of income between group companies with assessed losses

Section 20(1) states:

> For the purposes of determining the taxable income derived by any person from carrying on any trade, there shall, subject to section 20A, be set off against the income so derived by such person
>
> *(a)* any balance of assessed loss incurred by the taxpayer in any previous year

The definition of a 'person' in section 20(1) is restricted to a separate legal entity in terms of common law principles. A company is therefore prohibited from setting off an assessed loss arising from a fellow subsidiary company against its income. This gives rise to conflicting results compared to the situation where the same group entities are structured in the form of divisions within a single legal entity. In this case, the setting off of all the divisional income and losses is allowed.

For this reason, group companies often enter into elaborate schemes and transactions in an attempt to shift around income and losses between group entities in order to achieve the same tax effect as would have arisen if these separate legal entities had been set up as a divisional structure (Kannenberg 1999: 140). So, for example, expenditure may be channelled through a profitable entity within the group, creating income in the hands of the loss-making group entity. However, a deduction may not be allowed if the requirements of sections 11(a) and 23(g) of the Act are not met. Furthermore, section 103(2) of the Act specifically prohibits the utilisation of an assessed loss in a company by introducing income into that company.

The requirements of section 103(2) that apply to companies can be summarised as follows:

- Whenever the Commissioner is satisfied that any agreement affecting any company or any change in shareholding in a company has been effected,

- and has resulted directly or indirectly in income or any capital gain accruing to the company,
- solely or mainly for the purpose of utilising any assessed loss, any capital loss or any assessed capital loss,
- in order to avoid or reduce liability for any tax, duty or levy on income on the part of that company,
- then the setting-off of any such assessed loss or balance of assessed loss against such income (or any taxable capital gain) shall be disallowed, or
- the setting-off of such capital loss or assessed capital loss against such capital gain shall be disallowed.

What is envisaged is the situation where, as a result of an agreement or change in shareholding, income is injected into a company that has an assessed loss. However, section 103(2) only applies when the change of shareholding or agreement is carried out solely or mainly to utilise the assessed loss and thereby to reduce or avoid tax, since there is no abnormality requirement in section 103(2). If, for example, it can be proved that a company was acquired or an agreement was entered into for good commercial reasons and that the setting-off of income against the assessed loss was merely incidental to the main purpose, section 103(2) does not apply (Huxham & Haupt 2004: 357–358).

Should section 103(2) not apply, section 103(1) may be applied in cases where it can be shown that the transaction, operation or scheme was entered into solely or mainly to avoid, reduce or postpone any tax liability, as result of the abnormality of the transaction or scheme.

In practice, one of the most common ways to manipulate the tax liability of each group entity is often to make year-end adjustments to management or administrative fees charged between group entities (South Africa 1995: 97). There is often no basis for excessive charges, and should these be subject to enquiry from the revenue authorities, companies may in these instances find it difficult to demonstrate that a service has in fact been rendered in return for the management fee, or that the fee is not excessive for the nature of the service rendered (Middelmann 2003: 21).

Capital gains and other tax consequences of certain intra-group transactions

In terms of the Eighth Schedule of the Income Tax Act, the capital gains tax provisions apply to the disposal of assets on or after 1 October 2001. It should be noted that the introduction of the taxation of capital gains has not removed the necessity for determining the nature of the proceeds from the disposal of assets. If the asset is acquired with a revenue intention, proceeds on disposal are included in gross income and are taxable in full. If the intention is of a capital nature, the qualifying portion of the proceeds is subject to the provisions of the Eighth Schedule.

A detailed analysis of the Eighth Schedule falls outside the scope of this study, but some of the limiting provisions in respect of intra-group transactions are briefly analysed.

Paragraph 39 of the Eighth Schedule states that where an asset is disposed of between connected parties (generally this includes group companies), any capital loss must be disregarded. The loss may only be set off against capital gains arising from the disposal of assets between the same connected parties. This is the same effect as when the asset is transferred between two divisions within a single company. Any loss (as well as any gain) resulting from the transfer of assets is disregarded. Disposals to parties that do not form part of the same group of companies have normal capital gains tax consequences.

The limitation with regard to intra-group capital asset disposals gives rise to an inconsistency between the tax treatment and the economic reality in a group scenario. From a group perspective, where a company transfers an asset at a loss and it has other capital gains, these gains may not be reduced by this loss. Therefore a liability may arise in this entity, which results in an anomalous tax effect. This has been partially addressed by the introduction of the Corporate Rules and specifically section 45. Section 45 is only discussed briefly here, as a detailed discussion of these Corporate Rules falls outside the scope of this paper.

'Intra-group transactions' are discussed in section 45 of the Act. They are defined as any transaction in terms of which any asset is disposed of by one company (referred to as the transferor company) to another company which is a resident (referred to as the transferee company), and both companies form part of the same group of companies, as defined in section 1 of the Act (see the section in which the definition of 'group' is discussed), at the end of the transaction date. The transfer of assets from one company in the group to another may result in certain tax implications, such as recoupments, possible capital gains tax implications, transfer

duty, secondary tax on companies and donations tax. Section 45 provides for tax relief, if jointly elected by both companies, in respect of qualifying intra-group transactions. The transferor company is deemed to have disposed of the assets for proceeds equal to the base cost of the assets, which are transferred to the transferee company and deemed to be the base cost of the assets for the transferee company. Relief is therefore provided in respect of any possible normal tax implications (including capital gains tax), as it also applies to stock (section 45(2)).

Allowances previously claimed on fixed assets (defined in section 41 as 'allowance assets') are not recovered or recouped in the calculation of the taxable income of the transferor company. The two companies are deemed to be the same person, entitling the transferee company to the same qualifying allowances to which the transferor company would have been entitled. Future recoupments are for the account of the transferee company (section (45(3)).

The same principle applies in respect of qualifying section 24C allowances, if, for example, construction contracts are transferred as a going concern (section 45(3)(b)) (Arendse et al. 2004: 354–355). The relief measures provided for in section 45 are, in some instances, not available, for example, if the transferee company is exempt from income tax (section 45(6)(b) of the Act); or if the transferee company is not able to claim the same capital allowance or deduction as the transferor company (section 45(3)(a) of the Act), where, for example, Company A has used the asset for manufacturing purposes and claimed a section 12C allowance, but Company B, buying the asset, does not utilise the asset for manufacturing and can therefore only claim a section 11(e) allowance.

Section 45 therefore addresses most intra-group transfers of assets, but, as already mentioned, some types of transfers can still result in situations where a company transfers an asset at a loss and, although it has other capital gains, these gains may not be reduced by this loss, as the loss can only be utilised against capital gains made in transactions with a company within the same group (paragraph 39 of the Eighth Schedule).

Summary of anomalies arising within the present system of taxing group companies

It is clear from the foregoing discussion that the current tax treatment of intra-group transactions might result in certain anomalies. If these anomalies are favourable, this may promote an attempt to exploit transactions that may have no commercial substance but are tax beneficial and may result in a loss to the fiscus. If these anomalies are unfavourable, they may result in a situation where efficient business decisions are not made because of their potential negative tax effects on the group. Tax-induced economic activity can result in a misallocation of resources and have a detrimental effect on economic growth. Furthermore, such anomalies do not promote equity and neutrality within groups (Kannenberg 1999: 138). It is thus clear that the status quo is unsustainable.

Although the provisions of the Corporate Rules (sections 41 to 47 of the Act) offer some relief in respect of some intra-group transactions, these provisions do not cater for all circumstances. Only six types of transaction are catered for, namely:

- Company formations (section 42 of the Act)
- Share-for-share transactions (section 43 of the Act)
- Amalgamation transactions (section 44 of the Act)
- Intra-group transactions (section 45 of the Act)
- Unbundling transactions (section 46 of the Act)
- Liquidation transactions (section 47 of the Act).

The Corporate Rules were introduced as relief measures in respect of transactions between group companies or between founding shareholders and their company (Huxham & Haupt 2004: 254). These rules are based on the principle that the transfer of assets within a group structure should be tax neutral where effective ownership has remained the same (Middelmann 2003: 22). Unfortunately, these rules do not always provide for tax neutrality as intended. If one of the anti-avoidance provisions is triggered, taxable income may arise in the transferee company which cannot be set off against its assessed loss (if it has an assessed loss) or in determining its aggregate capital gain or loss (Middelmann 2003: 27).

The third interim report of the Katz Commission

The Third Interim Report of the Katz Commission (South Africa 1995) was presented in December 1995. The report addresses the issue of group taxation and makes a number of recommendations in this regard. A number of advantages and disadvantages of a system of group taxation were identified by the Katz Commission.

Advantages identified by the Katz Commission

The Katz Commission (South Africa 1995: 96–97) identified the following advantages of a system of group taxation:

- A closely held group of companies, although the group may consist of separate companies, can constitute a single economic unit for the purposes of strategic and financial planning. A tax system that ignores this reality can create economic and business distortions that can be addressed by a form of group taxation.
- One of the distortions created is the divisionalisation of companies into a single legal entity purely for tax reasons. This results in a loss of protection in respect of limited liability, and influences operational, management, compensation and competition policies. Such problems can be avoided with a system of group taxation.
- The alternative situation also results in certain distortions, for example, when a group of companies cannot be divisionalised for strategic reasons and hence the tax neutrality enjoyed under divisionalisation is not achieved. Once again, a group tax system addresses these anomalies.
- Management may invest large resources in establishing techniques that often have no commercial substance, mainly in order to avoid tax through the use of certain intra-group transactions, such as unsubstantiated management fees and transfer pricing. This undermines the integrity of the tax system, as these actions were induced by a desire to avoid tax rather than commercial considerations – the very influence that a tax system should avoid.
- Manipulation of intra-group transactions is not easy to police, and is often difficult to detect, even though complicated anti-avoidance measures may be in place. Furthermore, because there is no recognition in the tax law of the reality of a group's economic interest, tax avoidance and evasion do not end with merely trying to match profit or losses within a group. Although intra-group transactions usually have no real economic or commercial effect, they do have a tax effect as a result of common ownership or control. Further abuse is therefore possible by manipulating the cost bases to engineer timing, capital or revenue mismatches, or simply to 'lose' one end of a transaction.
- Under the current tax system, companies are assessed separately. This implies that the assessor often does not have access to all the information pertaining to all the companies within a group, because these companies may be registered in separate offices. A group tax system would ensure a full audit trail of all intra-group transactions, as well as the correct tax effects of transactions with outside parties. This would increase the power of the revenue authorities to police the system.
- Group taxation is sometimes regarded as being disadvantageous, as it encourages the formation of conglomerates. The Commission comments that in recent times this is no longer a valid argument, and that, in fact, the reverse is more common. Group taxation facilitates the unbundling of large organisations into more efficient multi-company structures. In the current tax system, this is discouraged, as it results in higher tax liabilities through higher profitability in each of the sub-units, as well as the loss of the benefit of large tax losses. The Commission is aware of the importance of facilitating the ownership and control of companies by emerging investors, both in the existing market and in the privatisation process.
- Because South Africa is now part of the international trade and investment community, it is important to align the current tax system with international practices, as foreign investors expect to find some form of group tax in South Africa.

Disadvantages identified by the Katz Commission

The Katz Commission (South Africa 1995: 98) identified the following disadvantages of a system of group taxation:

- A system of group taxation is complex.
- The cost to the fiscus is perceived to be high.
- There is a need for anti-avoidance measures.

Recommendations

The Commission proposed a gradual approach to the introduction of a system of group taxation, beginning with a simplified consolidation method. The suggested initial system of group taxation is not a fully-fledged consolidation system, but one that is able to progress towards a full consolidation system, once the impact of the shift to group taxation on the fiscus can be evaluated and administrative problems have been identified and addressed. This form of implementation should minimise the impact of the complexity of such a

system, as well as the cost. The consolidation system introduced should broadly follow international principles.

The Commission found that the claims that the fiscus would incur substantial losses were largely exaggerated and unfounded. Not all tax losses are available to group companies, and not all groups have profits that can be set off against such losses. The potential cost to the fiscus of setting off these losses could be largely countered by excluding losses prior to the first consolidation.

The Commission further found that a group tax system avoids the engineering of artificial transactions for the purposes of avoiding tax (which are difficult to control and police and undermine the entire corporate tax system). The Commission is of the opinion that the fiscus suffers more under the current tax system. The Commission recommended a compromise with the pure full consolidation system in respect of three areas (South Africa 1995: 100):

- There should be the requirement that only wholly owned groups qualify for consolidation in order to reduce cost and complexity. The fear that the 100% holding requirement will 'squeeze' out minorities is, in the Commission's view, a lesser problem.
- Any losses that arose prior to the consolidation of a group of companies should be excluded.
- A full consolidation method need not initially be implemented.

Canons of taxation

Fiscal policy has a critical impact on the political economy of any country, and many variables must be taken into account in the pursuit of an efficient, equitable and politically acceptable system of taxation. In this context, a number of so-called canons of taxation (which include equity, certainty, convenience, efficiency and neutrality) have been internationally accepted as representing the characteristics of a good tax system (Emslie, Davis, Hutton & Olivier 2001: 1).

The canons of taxation, first formulated by Adam Smith in 1776 (in a book titled *The Wealth of Nations*), are summarised by Huxham & Haupt (2004: 2) as follows:

(i) The subjects of every State ought to contribute towards the government, as nearly as possible, in proportion to their respective abilities; that is, in proportion to the revenue which they respectively enjoy under the protection of the State. The expense of government to the individuals of a great nation is like the expense of management to the joint tenants of a great estate, who are all obliged to contribute in proportion to their respective interests in the estate.

(ii) The tax which each individual is bound to pay ought to be certain, and not arbitrary. The time of payment, the manner of payment, the quantity to be paid, ought all be clear and plain to the contributor.

(iii) Every tax ought to be levied at the time, or in a manner which it is most likely to be convenient for the contributor to pay it.

(iv) Every tax ought to be so contrived as to both take out, and keep out, of the pockets of the people as little as possible over and above what it brings into the public treasury of the State.

These salient features of a good tax system are examined in the following sections, taking into account the comments made by the Katz Commission (South Africa 1995: 96–100), as discussed in the previous section, in order to establish whether a group tax system in South Africa would promote these features. Since it appears that a consolidation system would be the preferred choice if a group tax system were to be implemented in South Africa (see the recommendations made by the Katz Commission, as discussed in a previous section), the features of a consolidation system are also to be taken into account in the remainder of the article.

Equity and neutrality

The equity of a tax system is defined with regard to two related concepts. The first is the ability to pay, where one can distinguish between horizontal and vertical equity. Horizontal equity requires that similar individuals be treated similarly, or that a person in the same situation as another be treated equally. Vertical equity requires that taxpayers with a higher level of economic wellbeing should bear greater tax burdens. The second concept is the benefit principle, which states that those who benefit from the use of particular commodities or services should pay for them, according to the Margo Commission (South Africa 1986: 50–51).

Neutrality requires that taxpayers should not be influenced by the tax system in choosing one course of action over another, solely because the tax effects of one course of action are more beneficial under one of the options. A neutral tax

system is one that minimises the impact of the tax structure on economic behaviour, including business organisation, work effort and saving (South Africa 1986: 50).

A group tax regime addresses both these characteristics. This is clear from the manner in which the members of a group of companies are treated similarly and equally, as they are seen to form part of the same economic unit. With a system of group taxation, the same tax neutrality should therefore be enjoyed as that under divisionalisation (South Africa 1995: 96–97).

The current income tax system in South Africa clearly does not achieve this. Each company within a group is treated as a separate taxpayer, and all intra-group transactions result in tax effects. These tax effects often do not give rise to consistent results between the transacting group entities (refer to the anomalies examined in a previous section), and therefore do not emulate the group economic unit principle. However, intra-group transactions may be entered into solely for the purpose of exploiting favourable tax treatments, which does not necessarily result in sound economic decisions. Tax-induced business decisions may result in a misallocation of resources and are therefore disadvantageous to the economy.

Certainty and simplicity

Certainty and simplicity are also included among the characteristics of a good tax system and are interrelated. Certainty requires that taxpayers should be reasonably certain what their tax liabilities should be. A complex tax system results in uncertainty and increased costs because of the need for consultation with advisors. Simplicity requires that a tax should be easily assessed, collected and administered in order to minimise costs to both the taxpayer and the fiscus (South Africa 1986: 51). It refers to the ease of operation of a tax system from a technical point of view (Kannenberg 1999: 151).

A group tax regime is generally regarded as technically complex (South Africa 1995: 98), because intra-group transactions are subject to special treatment in order to neutralise their tax effects. However, this complexity depends on the manner and extent of implementation.

There are also some characteristics of a group tax system that could simplify tax treatment:

- There is no need for a separate exercise to identify intra-group transactions for the purposes of completing the group's tax return, because, in terms of generally accepted accounting practices, these intra-group transactions require specific detailed disclosure.
- In respect of income and expenditure in intra-group transactions, income and expenditure are mostly eliminated on aggregation of the group's results. So, for example, intra-group management fees received by one group company and included in its taxable income, will be set off against the management fees paid by the other group company, which claims it as a deduction, thereby eliminating this transaction.

Another important factor that should be considered is that income tax legislation could be simplified by the implementation of a group tax system, in that certain anti-avoidance provisions and the Corporate Rules (sections 41 to 47 of the Income Tax Act) would, to a large extent, become redundant. More consistent assessments would also be issued, as there are no separate assessments for each entity by various assessors under a group tax system.

It can therefore be concluded that a group tax system should achieve certainty and simplicity to a far greater extent than the current system. The Revenue authorities would have access to more information pertaining to the entire group, fewer returns would need to be submitted, and certainty regarding intra-group transactions would be greater, which should release management resources that can be concentrated on the economic activities of the group.

Cost and efficiency

The cost and efficiency of a tax system are directly linked to certainty and simplicity. A system that is more efficient administratively results in reduced costs (South Africa 1986: 51).

A system of group taxation appears to be more efficient and cost effective than the current tax regime, for a number of reasons:

- A group tax system is generally more efficient from an administrative point of view, since only one tax return needs to be completed for the group, and consolidated information provided by the financial statements can be incorporated, to a large extent, without any changes. Costs for the taxpayer should therefore also be lower.
- The initial cost to the fiscus to facilitate the implementation of a system of group taxation

(South Africa 1995: 98) should be set off against the benefits of the lower costs that result from fewer corporate tax returns being lodged.

- The submission of fewer corporate tax returns should also result in a decrease in administration and turnaround time for the fiscus.
- The assessor is privy to the information of the group as a whole, including the disclosure of all intra-group transactions, resulting in the provision of a more efficient service (South Africa 1995: 96–97).
- As intra-group transactions are tax irrelevant under a group tax system, the exploitation by taxpayers who try to manipulate such transactions to obtain a tax benefit is greatly reduced, thereby reducing the need for the revenue authorities to police such activities (South Africa 1995: 96–97).
- Management tends to invest resources in establishing techniques that often have no commercial substance, with the main purpose of avoiding tax through the use of certain intra-group transactions, such as unsubstantiated management fees (South Africa 1995: 96–97). A group tax system aligns the tax treatment of a group of companies for tax purposes with the economic unit principle, ensuring that companies utilise their resources effectively, not only taking into account their best interests, but also the best interests of the economy as a whole, thereby driving economic growth. This is critical in South Africa's developing economy.

Convenience

A group tax system appears to be more convenient for both the fiscus and the taxpayer, in the sense that only one tax return need be submitted and assessed for every company that elects to be treated as a group company for income tax purposes.

Summary

The foregoing analysis indicates that there appears to be a certain trade-off between the canons of taxation. A good tax system is one that achieves the best balance between equity, neutrality, certainty, simplicity, costs and efficiency. The only disadvantage that may be cited may be the complexity of the initial introduction of a group tax system. However, it is possible to design a system that is compatible with the current South African tax system and the administrative capacity of the South African revenue authorities.

Desirability for the fiscus and the taxpayer

From the preceding analysis, it is clear that a group tax regime would be beneficial to the South African tax system, and consequently to the South African economy. A number of aspects (some of which have already been mentioned) arising from the implementation of a group tax regime have been identified with regard to the desirability to both the fiscus and the taxpayer.

Desirability for the fiscus

The following are some of the reasons that a group tax regime would be desirable to the fiscus:

- A group tax system promotes efficient utilisation of management resources. Because it disregards all intra-group transactions for tax purposes, it eliminates tax anomalies that arise from these types of transactions. Management is therefore less likely to be influenced by the tax effects of these types of transactions, and is more likely to concentrate on promoting growth in economic activity, thereby increasing the general tax base (South Africa 1995: 71).
- A group tax system also promotes efficient utilisation of revenue authority resources. Revenue authorities are likely to spend less time policing intra-group transactions, and can therefore dedicate more time to other important areas, such as non-compliance.
- The information available to revenue authorities is substantially increased under a group tax (consolidation) system, as it provides information in respect of the entire group structure. All entities are effectively assessed by a single person, as opposed to the current situation where companies within the same group may be assessed by various assessors and even by different revenue offices (South Africa 1995: 71). This reduces the risk of exploitation of the system by the taxpayer and increases the consistency of assessments. It also improves turnaround time, and ultimately cash collection by the revenue authority.
- An argument often raised against the implementation of a group tax system is that the South African revenue authorities are ill-equipped to handle the complexities of such a system. However, over the last few years, the South African revenue authorities have shown that they are committed to transformation, and that in fact they are more than capable of competing with the fiscal systems of many

developed economies. This can be seen, firstly, from the way in which they have implemented new systems to improve the efficiency and effectiveness of assessments and administration of the South African tax system. Secondly, over the last few years, they have started to employ highly trained and skilled people and have accordingly increased salary packages to compete with the private sector.

- South Africa is now part of the international trade and investment community, and it is necessary to align the current tax system with international practices, as foreign investors expect to find some form of group tax in South Africa. This can ultimately contribute to an increase in the tax base, arising from new business derived from foreign investment.

Desirability for the taxpayer

The following are some of the reasons a group tax regime would be desirable to the taxpayer:

- A group tax (consolidation) system that ignores intra-group transactions for tax purposes promotes efficient utilisation of management resources. Management is able to concentrate on making business and operational decisions on commercial merit, and is not driven by the tax effects (South Africa 1995: 97).
- A group tax (consolidation) system supports the economic unity principle, as only transactions with outside parties give rise to tax consequences, and intra-group transactions are tax neutral.
- A group tax (consolidation) system reduces administration and compliance costs for income tax purposes by requiring only one return to be submitted for a group of companies.
- A group tax (consolidation) system promotes certainty, in that companies do not need to waste resources in consulting with special tax advisors regarding the intricate structures engineered to take advantage of intra-group transactions, as these are disregarded for tax purposes (South Africa 1986: 51).

Conclusion

The South African tax system has recently undergone radical transformation with the introduction of a residence-based tax system and capital gains tax, in order to achieve its globalisation policy. All the major developed economies, such as Australia, the United Kingdom and the United States of America, have some form of group taxation in their income tax dispensation. A foreign investor would thus expect to find a form of group taxation in South Africa. As South Africa makes progress as an emerging market, it is inevitable that a system of group tax will have to be introduced in order to integrate fully and compete globally. This will encourage foreign economic activity in the country and ultimately broaden the tax base.

On the basis of the analysis performed and the conclusions reached, it is evident that the status quo in respect of the inherent tax anomalies arising from intra-group transactions is not sustainable, and that a system of group taxation should be implemented in South Africa. A tax system that provides for a form of group taxation and recognises the economic unit principle will promote consistent results and will encourage sound business decisions, based on economic merit.

A tax system that includes a group tax regime would promote the canons of taxation more effectively than the current taxation system. It would be beneficial to both the fiscus and the taxpayer, and would ultimately support the growth of the South African economy. Although the introduction of a system of group taxation is not without its costs and complexities, it is both achievable and necessary for the advancement of the South African economy.

Acknowledgements

We thank Professor M.J. Nieuwoudt, our colleagues and three anonymous reviewers for their comments on earlier drafts.

References

Arendse, J.A., Coetzee, E.S.M., Jordaan, K., Kolitz, M.A., Stein, M.L. & Stiglingh, M. 2004. *Silke: South African Income Tax 2004.* Durban: LexisNexis Butterworths.

Caltex Oil (SA) Ltd v SIR. 1975(1) SA 665 (A).

Cilliers, H.S., Benade, M.L., Henning, J.J., Du Plessis, J.J. & Delport, P.A. 1992. *Corporate Law.* Durban: Butterworths.

CIR v Stott. 1928 AD 252.

CIR v Visser. 1937 TPD 77.

Department of Finance. 2001. Explanatory memorandum on the Second Revenue Laws Amendment Bill 2001. Pretoria: Government Printer.

Emslie, T.S., Davis, D.M., Hutton, S.J. & Olivier, L. 2001. *Income Tax Cases and Materials.* Cape Town. The Taxpayer.

Howitt, S.W. 1992. Group taxation in South Africa – an examination of the taxation implications and problems. Cape Town: Department of Accounting, University of Cape Town.

Huxham, K. & Haupt, P. 2004. *Notes on South African Income Tax*. Roggebaai: H & H Publications.

Kannenberg, E.A. 1999. The suitability of a system of group taxation for South Africa, with specific reference to the recommendations of the Katz Commission. Stellenbosch: University of Stellenbosch.

Middelmann, S.N. 2003. A study of the applicability of a system of group taxation in South Africa. Unpublished MCom mini dissertation, University of Pretoria, Pretoria.

New State Areas Ltd v CIR. 1946 AD 610.

Port Elizabeth Electric Tramway Co Ltd v CIR. 1936 CPD 241.

Pyott Ltd v CIR. 1944 AD 128.

South Africa. *Companies Act, Act 61 of 1973,* as amended. Pretoria: Government Printer.

South Africa. *Income Tax Act, 58 of 1962,* as amended. Pretoria: Government Printer.

South Africa. 1986. Report of the commission of inquiry into the tax structure of the Republic of South Africa (Chairman: CS Margo). Pretoria: Government Printer.

South Africa. 1995. Third interim report of the commission of inquiry into certain aspects of the tax structure of South Africa (Chairman: MM Katz). Pretoria: Government Printer.

South Africa. 1996. Budget review (Department of Finance). Cape Town: Department of Finance.

South African Chamber of Commerce. 1996. Comments on the third interim report of the commission of inquiry into certain aspects of the tax structure in South Africa. Johannesburg: SACOB.

Taxpayer. 1985. 'The Margo Commission report: some submissions by ASSOCOM', 34: 161–174.

Quality assurance in internal auditing: an essential tool in ensuring that the expectations of the users of internal auditing services are met

*M. Marais**

"Quality is not something you can buy, like a boat or automobile.
It is not a machine that can be installed on the third floor.
Nor does Japan have a monopoly on quality.
Quality is a way of doing business regardless of the continent, language or size of business." (Russell 1990: 3)

The managements of organisations, as well as oversight bodies, have certain needs that inevitably impact on their expectations of the services rendered by internal auditing functions. They further insist that the value that internal auditing functions add to their organisations should be in proportion to the financial expenditure attached to internal auditing functions. The question is raised whether or not internal audit functions are effectively supporting their organisations. Do they understand the needs of their clients and, if so, how can they ensure that they meet those needs?

Internal auditing functions will practise internal auditing optimally and meet their responsibilities in addressing clients' needs if they comply with the internal auditing standards and give effect to the definition of internal auditing. However, the compliance of internal auditing functions with the internal auditing standards cannot be taken as a given. For this reason, the internal auditing standards require that heads of internal auditing functions implement a quality programme, which entails a process that must ensure and improve the quality of the internal auditing function and the assessment of its efficiency and effectiveness by means of internal and external assessment.

The purpose of this article is to provide a theoretical overview of the importance of the provision of a quality service by internal audit functions in order to ensure that the expectations of the users of internal auditing services are met. Specific attention is given to the following:

- The meaning of quality and quality control in general
- The importance of the provision of a quality service by internal auditing functions
- Expectations of users of internal auditing services
- Conflict between the role of the internal auditing function and the needs of internal auditing clients
- The quality programme and its monitoring as methods of ensuring that internal auditing functions provide a quality service.

Introduction

The internal auditing profession defines internal auditing as follows (IIA 2001: 1):

> Internal auditing is an independent, objective assurance and consulting activity designed to add value and improve an organization's operations. It helps an organization accomplish its objectives by bringing a systematic, disciplined approach to evaluate and improve the effectiveness of risk management, control, and governance processes.

From this definition of internal auditing, it is clear that the internal auditing profession is directed towards the delivery of service. This service should add value to an organisation and improve the activities of the organisation.

* M. Marais is in the Department of Auditing, University of South Africa. E-mail: maraim@unisa.ac.za.

When the management of an organisation establishes an internal auditing function within an organisation, it is entitled to the service that, in terms of the internal auditing profession's definition of internal auditing, is the responsibility of an internal auditing function. The members of the top management of any organisation ought to set out, in a policy declaration, their specific expectations of the internal auditing function as part of the organisation' s policy document regarding control within the organisation. The charter of the internal auditing function emerges from this policy document and is, among other things, the written agreement between management and the internal auditing function, according to which the expectations that are spelt out in the policy document are carried out (Carolus & Nelson 1996: 8).

The Professional Practices Framework of the Institute of Internal Auditors (IIA) provides guidelines to internal auditing functions regarding the way in which internal auditing ought to be practised. Internal auditing standards and guidelines are regularly revised and updated to ensure that the needs of internal auditing clients may be continuously met. In adhering to the Standards for the Professional Practice of Internal Auditing (hereafter referred to as internal auditing standards), internal auditing functions carry out the responsibilities that appear in the definition of internal auditing. Conforming to the internal auditing standards is thus a method of ensuring that internal auditing functions deliver a service of quality and, in doing so, satisfy the needs of their most important clients.

The quality of internal auditing services can only be ensured if the service rendered is subjected to quality control. One of the internal auditing standards, Standard 1300, requires the head of an internal auditing function to develop and maintain a programme to ensure and improve the quality of the internal auditing function (hereafter referred to as the quality programme) which addresses all the aspects of the internal auditing function and which is continuously checked for effectiveness (IIA 2004: 11).

Objective

The objective of this study is to provide a theoretical overview of the importance of the provision of a quality service by internal auditing functions.

Research problem

The environments within which organisations operate are influenced by several factors, which could be either external factors (such as developments in information technology, global trading and the demand for transparency) or internal factors (such as strategic orientation, reengineering and risk management, to name but a few). These factors cause managements of organisations as well as oversight bodies to develop certain needs, which inevitably impact upon their expectations of the services rendered by internal auditing functions.

The question is raised whether or not internal audit functions effectively support their organisations. Do they understand the needs of their clients and, if so, how can they ensure that they meet those needs?

In addressing this issue, this article addresses the following:

- The meaning of quality and quality control in general
- The importance of the provision of a quality service by internal auditing functions
- Expectations of users of internal auditing services
- Conflict between the role of the internal auditing function and the needs of internal auditing clients
- The quality programme and its monitoring as methods of ensuring that internal auditing functions provide a quality service.

The article focuses on quality and quality control within internal auditing functions, and not on quality and quality control of the internal auditing profession as such. For instance, the requirements for admission to the Institute, disciplinary rules and other factors affecting quality control of the internal auditing profession in general are not discussed. Unless otherwise indicated, the information that appears in this article applies to both the private and the public sectors, but focuses specifically on internal auditing functions that operate as sections or departments within organisations.

Research strategy

In this study, the following research methodology was used:

- Both local and overseas literature on quality and quality control within internal auditing functions, in the form of published books, articles in

professional journals and other journals, as well as published guidelines in the IIA and the Institute of Internal Auditors of South Africa (IIA SA), were consulted.

- Appropriate dissertations, theses and surveys were studied.

The meaning of quality and quality control in general

Those who write about quality agree that it cannot be defined in absolute terms but that client satisfaction is a factor in determining quality. ''Any quality system should be based on meeting the customers' needs both internal and external'' (CIPFA 1993: 2).

The Chartered Institute of Public Finance and Accountancy (CIPFA) in London defines quality as the complete characteristics and traits of a product or service that satisfies implicit and explicit needs (CIPFA 1993: 53). This definition of quality is also used in the ISO 9000 series of standards and guidelines for the management of quality and quality assurance, as published by the International Organization for Standardization (ISO) (Ridley & Stephens 1996: 77).

According to Russell (1990: 7), from the client's point of view, quality is getting what you expected and, from the provider's point of view, getting whatever you deliver right at the first attempt. Stated differently by Russell (1990: 10): ''Quality is striving for 100 percent performance and achieving 100 percent customer satisfaction.''

Furthermore, it appears from the literature that the delivery of services or products of quality demands specific inputs from the provider of that service or product. These inputs entail the following:

- A grasp of and disposition towards quality. Quality should be defined according to the specific needs and aims of each organisation. Quality ought to accord with the organisations's or business's quality policy and *raison d' être* (Russell 1990: 14). If an organisation can succeed in combining the right attitude, perfect timing, best equipment, good facilities and a faultless product or service, client satisfaction ought to be a natural consequence, provided it is a product or service for which there is a market (Lefevre 1989: 20).
- Knowledge of the client. To be able to focus on clients' needs, an organisation must know who the clients are, identify their needs and keep in touch with them (Lefevre 1989: 23). As soon as an organisation knows who its clients are, it can focus on their needs. When an organisation knows what its clients' needs are, it is able to emphasise the right services (Lefevre 1989: 36).
- Knowledge of the enterprise. Managers should know what is going on in their enterprise. They should also be aware of what quality service (or the lack thereof) costs the enterprise (Lefevre 1989: 39). Russell (1990: 10), however, maintains that quality should not be viewed as costly but as remunerative for those who pursue it. According to him, an enterprise that does not pursue quality might improve its profitability in the short term but, eventually, it will lose its share in the market (Russell 1990: 23).
- A process for ensuring quality. Quality assurance entails planned, systematic steps to ensure that products or services comply with specific quality requirements (CIPFA 1993: 53; Ridley & Stephens 1996: 77). Russell (1990: 13) contends that a quality process should be built into the management structures and day-to-day activities of each organisation.

The latter implies quality control, which is described by Russell (1990: 6) as the processes followed by organisations in order to meet the expectations of a client in the provision of a specific product or service. Thus, although quality means different things to different people, the provision of quality products and services should always:

- Have the satisfaction of clients' needs as its main aim.
- Be properly researched, planned and managed within each organisation
- Demand purposive control and supervision on the part of the provider.

Importance to internal auditing functions of providing a service of quality

The quality of the services provided by internal auditing functions influences, in the first place, the suitability and viability of the internal auditing profession and, ultimately, the viability of internal auditing functions within organisations, or 'in-house' internal auditing functions.

Viability of the internal auditing profession

''Enhancing the [internal audit] profession's future viability requires ensuring the quality of internal

auditing services'' (IIA 1999: 3). This is not unique to the internal auditing profession. Other professions in South Africa, such as public accounting, engineering and law, even though regulated by South African law regulates them, still have to ensure that they meet the expectations of their major clients in order to enhance their future viability.

According to Miller (2000: 27–35), it would seem that the perceptions of internal auditing by those outside the profession have changed over the years. Miller conducted interviews with individuals in various positions in the American business world. For example, to David McIntosh, senior manager at an organisation, internal auditing was initially a combination of process checking and random statistical sampling. Now he predicts that internal auditing will be the first profession to be tasked with the appraisal of non-financial assets such as intellectual assets and business relationships (Miller 2000: 28). Initially, Chris Carney, the chief executive officer of an organisation, believed that the main aim of internal auditing was to catch management out for making mistakes (Miller 2000: 29). Today he utilises his organisation's internal auditing function to add value to the business (Miller 2000: 30).

According to Patrick Lencioni, speaker, consultant and author, internal auditing ought to work at rectifying false perceptions of internal auditing on the part of those who make use of the service. Internal auditors should market a different image of themselves so that executive management will welcome the valuable role that internal auditing could play and teach their employees to do the same (Miller 2000: 29).

However, in the United Kingdom in 1999, a CIPFA survey entitled 'Perceptions of Audit Quality' indicated that at the time of the survey, a misconception persisted among internal auditing clients regarding the role that internal auditing functions are supposed to play. The nature of the problem was conveyed as follows in the report (IIA 1999: 35):

> Auditors recognise that the historical probity-centred role of auditing has been superseded by a positive, forward-looking emphasis on auditing systems and operations. They see their role as adding value to an organisation through offering appraisal and advice on internal systems, backed up by wider business skills. However, the survey shows that auditees, particularly non-financial managers, still primarily associate internal auditing with finding errors and preventing and detecting fraud.

Thus, although internal auditors are aware of the role they ought to play within organisations, in practice they do not always succeed in fulfilling that role in such a way as to change their clients' perceptions of internal auditing. In order to address these false perceptions, the Guidance Task Force[1] has suggested that an image-building programme be implemented by the IIA, with a view to improving the knowledge and worldwide acceptance of internal auditing standards and enhancing the prominence of the internal auditing profession (IIA 1999: 35). The aim of this recommendation is, firstly, to improve the quality of internal auditing by improving internal auditors' knowledge of standards and, secondly, marketing compliance with the standards to internal auditing clients and other interested parties as a sign of quality. In response to this recommendation, the internal auditing standards have been modified to specify that internal auditing functions can declare that they have conducted an audit based on the internal auditing standards only if they prove their compliance with the internal auditing standards by means of an external quality assessment of the internal auditing function every five years (IIA 1999: 35).

Viability of the 'in-house' internal auditing function

For the past decade, the total or partial outsourcing of internal auditing functions has developed into a serious threat, which the internal auditing profession needs to take into account. For instance, Burns (2000: 80) writes: ''Every time an internal

1 During 1997, the Guidance Task Force was commissioned by the Institute of Internal Auditors to investigate the following matters:

- The possibility that there was a gap between developing internal auditing practice and the existing internal auditing standards
- The possible improvement of existing processes according to which internal auditing standards develop and are laid down and the ways in which improvement could be assisted.

The Guidance Task Force consisted of 16 leaders in the field of internal auditing and was representative of all interested parties throughout the world. They met six times, with the first meeting being held in December 1997 and the last meeting in September 1998. The details and results of their investigation are contained in a publication of the Institute of Internal Auditor's Research Foundation entitled 'A Vision for the Future: Professional Practices Framework for Internal Auditing' (IIA 1999: 1).

audit department falls victim to complete outsourcing, it's a blow not only to the individuals involved but also to the credibility of the profession.''

In an investigation into the outsourcing of internal auditing services (Rittenberg & Covaleski 1997: 71), the researchers came to the conclusion that most of the internal auditing functions that were eventually outsourced did not focus on adding value and improving the management of the enterprise. Nor, very often, were the sections proactive: they failed to initiate change, with the result that their services became less important to the enterprise. This research did not, however, show that only unsuccessful internal auditing functions are outsourced. Successful internal auditing functions are also outsourced. The challenge to each organisation is thus to obtain the right internal auditing service for the right price, and the challenge to existing internal auditing functions is to prove that they provide the right services, which add the most value to the organisation.

Michael Corbett, co-founder and director of the Outsourcing Institute, suggests the following as the motive for the outsourcing of services in general: ''Few organizations have the ability to create and maintain world-class capabilities in every aspect of their business, and yet none can afford to be less than world class in any'' (Rittenberg & Covaleski 1997: 1).

In the competitive market in which internal auditing functions find themselves today, their pursuit of excellent service is probably the element that will ensure their positions within organisations. ''The essential ingredient for exceptional performance by a professional is a passion for excellence'' (Fabrizius 1997: 23–24).

Anthony Walz (1997: 51) views as naïve and shortsighted internal auditing functions that think of themselves as essential functions with a guaranteed exclusive right to existence and immunity to critical investigation, downscaling and outsourcing. He maintains that, without clear structures through which the adding of value may be demonstrated, internal auditors run the risk of being labelled by management as consumers of resources rather than adders of value.

''Unfortunately, few audit departments base their project list on what the organisation needs most. Instead, internal audit plans are often determined according to what the internal audit department does well – which in many cases means little more than assessing control activities'' (Leithhead 2000a: 68). This means that, because internal auditors do not address the most critical issues of many organisations, they endanger their own positions within organisations. Leithhead (2000a: 68) believes that there is a very good chance that an internal auditing function that does not make a meaningful attempt to address the needs of the organisation will be outsourced.

Strategies to manage the risk of becoming irrelevant

Based on the Competency Framework for Internal Auditors (CFIA), of which he was a co-author, Leithhead (2000a: 69) identifies the following strategies that internal auditing functions could use to address the organisation's needs and to manage the risk of becoming irrelevant:

- Consultation with senior management and line management to ascertain which services are expected of the internal auditing department and to ensure that risk management is properly understood and managed in changing circumstances
- Determining the skills needed by the internal auditing function, in the form of staff, equipment and infrastructure, to meet the needs of the organisation
- Recognising the dynamics of the internal and external environments within which the internal auditing function operates and identifying the risks these hold for the internal auditing function
- Developing a suitable plan of risk management for the internal auditing function
- Creating opportunities for the internal auditing function and optimally utilising such opportunities.

Drent (2002: 53–55) agrees with this and identifies the following four techniques that may be used to prove that the internal auditing function is of importance to an organisation and ought to enjoy recognition:

- An external assessment of the internal auditing function, during which the effectiveness of the function is measured in terms of the expectations of the most important clients
- Discussion with interested parties of the focus and planning of the internal auditing function, bearing in mind the definition of internal auditing
- Continual feedback from internal auditing clients about the quality of internal auditing processes and client satisfaction (either by means of

anonymous questionnaires or direct questions) and the presentation of such feedback to top management and the auditing committee
- Identification of key performance areas that are important to the client and appropriate criteria that may be used to monitor the performance of the internal auditing function.

According to the IIA, the practice that auditors ought to adopt in order to produce the best quality auditing services and to remain effective over the long term entails the following:

- Complying with the guidelines set out in the ethical code and internal auditing standards
- Aspiring to individual certification as internal auditors
- Executing internal assessment
- Submitting to independent external quality assessment to ensure compliance with the internal auditing standards (IIA 2003: 2).

The foregoing points to the importance of the internal auditing function's providing proof of the quality of internal auditing services and of their pursuit of the satisfaction of their most important clients' needs, in order to justify their viability. As discussed earlier, client satisfaction is the chief criterion of quality (CIPFA 1993: 53). When internal auditing functions ensure that they provide a quality service in accordance with Standard 1300, they are satisfying the needs of their chief clients and they can justify their viability (Fabrizius & Serafini 2004: 43).

Expectations of users of internal auditing services: Who are the users of internal auditing services?

"Quality is defined as meeting the expectations of customers. Any organization that implements quality standards must make sure that the customers' (both external and internal) requirements are understood and that the organization is capable of fulfilling those requirements" (Ridley & Stephens 1996: 7).

The internal audit function should thus firstly identify who its clients are, after which the requirements for a quality service in the circumstances can be planned. When the needs of a client change, the internal auditing function should adapt, and, as the pressure on a client increases, the auditing function should seize the opportunity to better assist the client (Bishop, Hermanson, Lapides & Rittenberg 2000: 51).

During their research into a new practice framework for internal auditing, the Guidance Task Force (IIA 1999: 79) investigated the issue of who could be viewed as the internal auditor's client. The most distinguished internal audit clients singled out by this research are the organisation/department/activity that is audited, and various role-players within such an organisation/department/activity, the external auditors, audit committee, and providers of products and services to the organisation.

Another avenue of thought is that it is better to focus on one client, namely the organisation as a whole. The definition of internal auditing also refers to support of the organisation as a whole, in preference to distinct role-players within the organisation. Ewert (1997: 55) believes that, in practice, this approach means that all internal auditing activities are focused on service to the organisation. It puts the organisation's interests above any others and brings the aims of the internal auditing function into line with those of the entire management team, by adding value to the organisation, improving the organisation's activities and achieving the organisation's objectives.

Whether the focus is on the organisation as a whole, or on the various role-players within the organisation as clients, it is well nigh impossible to compile a complete list of the needs of internal auditing clients. Each organisation is unique as a result of factors such the nature of the organisation, the composition of the organisation, and the market served by the organisation.

Needs of users of internal auditing functions

In the course of the literature review for this study, certain needs common to users of internal auditing services became apparent. These common needs will be discussed under eight main points.

Internal auditing functions must add value to organisations

"Non value-adding functions are in the spotlight, and internal audit departments that don't address 'the things that really matter' run the risk of becoming dispensable" (Leithhead 2000a: 68).

In the prevailing business environments, each department and employee should add value that is directed to achieving business objectives. As in the case of other business functions, internal audit departments are also expected to show that they are able to produce a reasonable return on the investment that has been made in them (IIA SA 1997: 3). If senior officials within an organisation do not consider the internal audit function as value adding, the function is seen as an expense rather than a resource. It is the responsibility of every internal auditor to make sure that the latter does not happen (Burns 2000: 80). Thus it is crucial that internal auditing functions should seek opportunities to illustrate their value to organisations. They should build relationships based on mutual respect, trust and teamwork with managements, organisations and boards (IIA SA 1997: 3). According to an article by Miller (2000: 27–35), in which the requirements of various users of auditing functions are expressed, internal auditing functions will be respected if both senior management and the board are convinced that the chief aim of the internal auditing function is to improve the organisation (Miller 2000: 33). Internal auditors should do the following to instil the latter conviction in management and the board of directors:

- Focus on ways of adding value to the organisation instead of simply reporting on procedures and policy that have not been followed (Miller 2000: 32; Leithhead 2000a: 68)
- Be proactive by reporting mistakes when they are identified so that processes can be rectified immediately, rather than after the formal internal audit report is submitted (Miller 2000: 29)
- Be involved in identifying risks as well as identifying and suggesting methods of dealing with those risks (Miller 2000: 32; Leithhead 2000a: 68)
- Adopt auditing approaches that are flexible and adaptable in order to reflect changed business environments (IIA SA 1997: 3)
- Not rely only on their own skills, but obtain assistance from specialist consultants when necessary (IIA SA 1997: 3)
- Anticipate the organisation's future needs and develop internal audit expertise in those fields (Burns 2000: 80)
- Be capable of rapidly summing up new situations and making valid recommendations (IIA SA 1997: 3)
- Inform decision-makers about the most probable outcomes of possible decisions and provide meaningful feedback on the results of decisions taken (Steinberg & Pojunis 2000: 36)
- Accomplish changes that improve processes and promote the effectiveness of the organisation (Miller 2000: 34).

Internal auditing functions should perform a consultative function

As internal auditors become more involved in corporate control, opportunities arise for them to give guidance and act as management consultants. Because they know so much about the inner workings of their organisations, internal auditors ought to be the first to be approached – in their professional capacities – by top management to assist with strategic decision-making issues (Miller 2000: 28). For instance, internal auditors should advise any team that decides on changes in the composition of an organisation, and should assist in identifying risks that need to be taken into account in the consideration of new business opportunities (Miller 2000: 30). Internal auditors should support management. They must identify areas in which management might have an interest, provide support where necessary and give management the information it requires to do its work (IIA SA 1997: 2–3).

In addition, internal auditors should foster best practice within organisations, thereby improving processes. Internal auditors should act as more than consultants, by assisting staff members to achieve their objectives. They should not only point out areas for improvement but also make recommendations that will solve problems (Miller 2000: 34).

Internal auditing functions should contribute to good corporate governance and exercise a positive influence on corporate culture

"An effective internal audit function should provide:

- Assurance that the management processes are adequate to identify and monitor significant risks;
- Confirmation of the effective operation of the established internal control systems;
- Credible processes for feedback on risk management and assurance .." (Institute of Directors in Southern Africa 2002: paragraph 4.2.2).

Internal auditors need to understand the pressure exerted by shareholders and interested parties on the management of an organisation and should assist management with proper corporate govern-

ance. This implies that internal auditors must identify all factors that influence corporate governance in general and the organisation's specific circumstances in particular, and at all times keep abreast of all the factors that might influence good corporate control. They need to process this knowledge in order to suggest actions to management and the audit committee that will help management safeguard the organisation against familiar as well as hypothetical risks (IIA SA 1997: 2–3).

An important contribution that the internal auditing function can make to corporate governance is to advise boards of directors about the adequacy of internal control systems. To this end, the internal audit function should have direct access to and good communication with the board or top management of the organisation (Miller 2000: 30).

Internal auditing functions may also help to establish effective audit committees within organisations. By cooperating in a consultative capacity with senior management, internal auditors can explain the skills that an audit committee ought to have in order to adapt to the organisation's unique requirements. They could make available to the board the latest guidelines published in the commercial sector relating to the requirements with which audit committees are expected to comply, and give their input as to whether those who are nominated to serve on audit committees have the appropriate knowledge of financial and business matters (Miller 2000: 32; Marks 2003: 42).

"A Mc Kinsey worldwide study showed investors are willing to pay up to 27% more for companies with good corporate governance" (Binneman 2002: 12). Investors thus place a premium on good corporate governance and a sound corporate culture within an organisation. Internal audit contributes to security, which elevates the organisation's value in the eyes of its clients, owners and credit providers (Miller 2000: 31). Internal and external auditing is necessary to ensure honesty, predictability, continuity, accuracy and reliability (Miller 2000: 35).

In order to contribute to good corporate culture, internal auditors need to have a thorough knowledge of the organisation's culture (Miller 2000: 32) and advise management on the exercise of best practices, pursue standards and comply with legislation (Miller 2000: 34).

For instance, internal auditing functions may assist in establishing an ethical code within organisations and introducing this to the staff. They could then analyse any transgressions of the ethical code and point out repetitive matters and trends. They could evaluate and monitor the existence, utilisation and effectiveness of communication channels within the organisation (Steinberg & Pojunis 2000: 38).

Compliance with the audit committee's requirements means that the internal auditing function also needs to ensure that there are proper control measures in place, even when these apply to senior management (Wagner 2000: 1; Reding, Barber & Digirolamo 2000: 45). Internal auditors ought to assess the corporate structure to identify opportunities that could give rise to fraud, such as irresponsible conduct on the part of management and boards, inefficient and ineffective boards and inadequate control systems. They should also identify and report on circumstances that could influence senior management to make unethical decisions or behave unethically, such as transactions that could lead to conflicts of interest, excessive remuneration of directors and top management, and transactions that are not made public (Rezaee 2002: 61).

Internal auditing functions should be supplied with well-trained staff whose knowledge is appropriate to the organisation

Internal auditing functions have a complicated mandate that makes great demands on the knowledge and skills of internal auditors (IIA SA 1997: 2). Managements of internal auditing functions should thus anticipate the future needs of organisations and analyse the internal auditing department's capacity to meet those needs. With this information, the internal auditors are able to prepare themselves by undergoing the necessary training and acquiring skills that might be needed to provide for the organisation's future requirements (Burns 2000: 80). Equipped with the appropriate knowledge, internal auditors can appropriately advise management on appropriate steps to discern, understand and deal with critical vulnerabilities (Leithhead 2000b: 59).

Besides having good academic qualifications, whether these are oriented towards the financial, computer or business fields, as well as the qualification of Certified Internal Auditor (CIA), internal auditors should also have personality traits such as the following:

- Creativity, ability to make accurate assessments of situations and a natural tendency towards correct behaviour
- Good judgement, strong character and a balanced view of things (Wagner 2000: 5; Leithhead 2000c: 69).

Internal auditing functions should learn to understand organisations from a business perspective and to grasp the aspects that are important to management and business functions

"Directors want to sleep comfortably knowing major surprises won't hit their companies" (Steinberg & Pojunis 2000: 39). From this need follows the basic role of internal auditing functions, namely, to help organisations to achieve their goals. The goals of management and the organisation should form the focal point of internal auditing practice. The closer the internal auditing function can get to knowing what and how managers think, the greater the potential contribution the function will be able to make and the greater the chances that it will contribute to achieving the goals of the organisation (IIA SA 1997: 2–3; Steinberg & Pojunis 2000: 38–39; Van Wyk 2002: 14).

The management of the organisation is responsible for dealing with risks. Internal auditing is supposed to have an in-depth knowledge of risk and ought to be involved in the processes whereby risks are ascertained, measured and managed. Internal auditors are in the best position to identify and investigate risk at the highest level (Steinberg & Pojunis 2000: 36).

Internal auditing functions ought to report objectively to the managements of organisations about the effectiveness of the organisations' risk management processes. By contributing to an organisation's risk management process, internal auditors may help to ensure that projects are effectively run and that project objectives are accomplished. If they understand the organisation, internal auditors are also able to identify what does not necessarily relate to a specific audit. Internal auditing functions ought to ensure that the organisation's top management receives the necessary information regarding the extent to which strategies are brought to fruition. Moreover, they can assist management by providing more future-oriented information, as well as criteria and comparisons for the measurement of organisations' performance in relation to the market (Leithhead 2000b: 59; Steinberg & Pojunis 2000: 38; Miller 2000: 30; Wessels 2002: 10).

Internal auditing functions should provide appropriate information

"An effective internal audit function should provide: ... objective confirmation that the board receives the right quality of assurance and information from management and that this information is reliable" (Institute of Directors in Southern Africa 2002: paragraph 4.2.2).

Because boards draft policy, they need a constant, available and reliable source of information (Steinberg & Pojunis 2000: 36). Internal auditors are able to collect information and to provide assurance that all information is accurate and complete (Steinberg & Pojunis 2000: 37).

To be able to function effectively, audit committees also need timely and reliable information about the risks and control systems of organisations with which they are involved. Internal auditors could take the lead in developing effective information systems that provide audit committees with objective, detailed and consumable information. They should also assess the effectiveness, reliability and usability of information on a continuous basis (Rittenberg 2002: 32).

Clikeman (1999: 32–33) maintains that, traditionally, internal auditors assessed the reliability and accuracy of information, which is important, but that they add more value by expanding their scope to all the aspects of quality information, namely appropriateness, promptness and costs. He illustrates these aspects as follows:

- Appropriateness. Before internal auditors confirm the accuracy of a report or a database, they should ascertain whether the information is appropriate, in the sense of being informative and suitable for making decisions. They should also consider whether the collection and reporting of information is not being duplicated elsewhere in the organisation.
- Promptness. Internal auditors should analyse the information needs of decision-makers and make sure that the information is received regularly enough and in sufficient time to support optimal decision-making.
- Costs. Internal auditors can assist their organisations to make significant cost savings by assessing the efficiency of their organisations' information systems and ensuring that they are economically run. In addition, internal auditors should consider whether their organisations are making adequate use of new innovations in information technology.

Internal auditing functions should be a partner in business – not simply criticise

Chapman (1998: 40) predicts that the viability of internal auditing functions that continue to cling to

the traditional role of internal auditing are in jeopardy, but that internal auditing functions that adapt to new avenues of thought can play a challenging and exciting role within organisations and add a great deal more value to their organisations. She maintains that internal auditors should sometimes stand aside from the rules and look at the auditing process as a whole. It is important for auditors to take some time to understand the activities they examine and to think about the aim of an auditing project.

Internal auditors should understand their position within the organisation. They should always take into account their role as partners and consultants (Miller 2000: 34). Their aim should be to help. Instead of simply reporting problems, they should make recommendations to solve the problems. Internal auditors ought to be encouraged to make meaningful suggestions. They should enhance their role as partners in business by behaving reasonably and objectively (Miller 2000: 30).

Instead of rapping others over the knuckles, internal auditors should make a positive contribution by cooperating with policy-makers and systems developers to realise a joint goal (Miller 2000: 31).

Control should be driven by processes and systems rather than by people. If people apply control, it arouses resistance. For this reason, internal auditors should not point a finger at anyone or tell people what to do. They should treat members of staff whose work they scrutinise as reliable adults who are capable of doing the work to the best of their ability. They ought to take into account the fact that their clients are human beings and focus on auditing processes that foster respect, confidence and adult relationships (Miller 2000: 35).

Internal auditing functions ought to market themselves

"There is a close correlation between the amount of communication between internal audit and senior management, and top management's perception of the function" (Van Wyk 2002: 14).

Managements of internal auditing functions cannot expect senior officials within organisations to discern and appreciate the advantages of internal auditing themselves. The abilities of internal auditing functions and what they have achieved need to be deliberately brought to the notice of top managements and auditing committees. Where possible, internal auditors should be assigned to tasks where their visibility is promoted. They should be encouraged to play a public role within the organisation. Marketing the talents and contributions of the internal auditing staff draws attention to the value of internal auditing functions to their enterprises and reduces the probability of their being outsourced (Burns 2000: 80).

Conflict between the role of the internal auditing function and the needs of internal auditing clients

One of the objectives set when the internal auditing definition was revised in 1999 was that the new definition should indeed be used to provide guidance to the internal auditing profession. For this endeavour to succeed, the definition had to explain the true nature of internal auditing as practised today, and convey the vision of the internal auditing profession to all who practise internal auditing or have another interest in it (IIA 1999: 5).

In the light of this, the definition of internal auditing that has been acknowledged by the IIA should be used as the point of departure in any discussion of the role of the internal auditing function within an organisation.

The internal auditing definition and internal auditing standards provide sufficient latitude to internal auditing functions to meet the needs of their clients, provided that their independence and objectivity is retained. The extent to which an internal audit function gives effect to the image portrayed by the internal auditing definition and complies with the internal auditing standards depends on each internal auditing function. However, the boards of organisations are largely responsible for the efficiency and effectiveness of their internal auditing functions in ensuring that such functions have proper mandates according to which they can fulfil their responsibilities. "The board will get the type of internal audit function it deserves. In their own interests, all directors should act to ensure an effective internal audit, providing assurance in relation to risk management, control and governance processes" (Payne 2002: 7).

Nonetheless, the needs of various internal auditing clients can still cause a degree of conflict between internal auditing clients and internal auditing functions. Drent (2002: 50) describes this conflict as follows: "Daily, internal auditors face the challenge of reconciling their need to make a meaningful

contribution to the business with meeting the needs of their key customers – the audit committee, executive management and line management."

Internal auditing clients have various needs, which should be met by the internal auditing function. If internal auditing functions do not play their part properly with their prominent clients, it could give rise to conflict, with the result that organisations would not receive optimal service from internal auditing functions (Van Wyk 2002: 14).

Internal auditing functions should at all times consider whether or not they are meeting the needs of both the audit committee and the board. For instance, the board would be interested to know whether the control measures are adequate, properly implemented and receiving the attention of management. The most valuable contribution on the part of the audit committee is to conduct independent investigations of business processes that ensure the proper safeguarding of the activities of an organisation and its preparedness for future events (Miller 2000: 33).

Drent (2002: 51) and Van Wyk (2002: 14–5) point out the following underlying conflicts in the requirements of internal auditing clients:

- Executive and line management do not always understand the fact that internal auditors must be independent. Many executives and line managers view internal auditors as employees of the enterprise, and their reporting to the audit committee is seen as no more than a formality to provide reassurance for the purpose of corporate governance. If internal auditors insist on their independence, management tends not to view them as part of the team. In contrast to this, the auditing committee attaches great value to the independence of the internal audit function.
- Sometimes executive management is inconsistent about defining the role of internal auditing functions, expecting them to provide independent opinions in some areas and requesting them to fulfil operational responsibilities in other areas. The result of this is that internal auditing functions are obliged to play the partial role of advisors while being used partially to keep the activities of the business under control. It is difficult for internal auditors to market their advisory services as long as management uses the internal auditing function as a weapon for exercising control.
- By their very nature, advisory services are more time-consuming than insurance services. With limited resources and the priority given by management to regulatory requirements, internal auditors do not have sufficient time to act in an advisory capacity. This means that employees might experience the auditing function as a policeman rather than as a business advisor.

It is the chief audit executive's responsibility to manage such conflicts. Communicating openly with the most important clients and being prepared to ask for feedback, even if this is not pleasant, could influence internal auditing clients to form a better understanding of the role of the internal auditing function and could contribute to the ability of the auditing function to satisfy the needs of their clients (Van Wyk 2002: 14).

The quality programme and its monitoring as methods of ensuring that internal auditing provides a quality service

"The internal audit profession is based on the concept that impartial assessments, feedback and recommendations help to render overall enhancements. In fact, there's often no better performance indicator than the observations of a knowledgeable, objective, third party. Reality is sometimes hardest to see – and improvements most difficult to make – by those closest to a subject matter" (Miller 2000: 2).

One of the objectives set for the internal auditing profession by Bookal (2002: 47), the 2002/2003 international chairperson of the IIA, was to promote professionalism and the competence of internal auditors. He believed that one way of achieving this would be for the IIA to encourage internal auditors to provide evidence of quality within internal auditing functions by regularly having the quality of their services assessed.

Within the internal auditing function, quality control has the satisfaction of the client's needs as its chief aim. According to the Guidance Task Force (IIA 1999: 56), the existence of internal auditing functions is justified by the service they deliver to organisations.

The definition of internal auditing, which defines this service to the organisation, indicates that internal auditing is designed to add value, to improve the operation of organisations, and to help them to achieve their objectives (IIA 2004: xxvii).

The Professional Practices Framework, in which the definition of internal auditing, the internal

auditing standards, the ethical code and the practice guidelines for internal auditing are contained, is the recognised guideline of the internal auditors' profession, according to which internal auditing practice must be implemented. Internal auditing functions will practise internal auditing optimally and meet their responsibilities in addressing client's needs if they comply with the internal auditing standards and give effect to the definition of internal auditing.

However, the compliance of internal auditing functions with the internal auditing standards cannot be taken as a given. For this reason, the internal auditing standards determine that heads of internal auditing functions must implement a quality programme (IIA 2004: 11). Standard 1300 of the internal auditing standards stipulates that a quality programme must be put into operation within internal auditing functions, developed with the purpose of helping the internal auditing function to add value and to improve the activities of the organisation. In addition, the programme should provide assurance that the internal auditing function will act in accordance with the internal auditing standards and the ethical code (IIA 2004: 11).

The quality programme entails a process that must ensure and improve the quality of the internal auditing function and the assessment of its efficiency and effectiveness by means of internal and external assessment.

The internal auditing standards require that the quality programme has to:

- Cover all aspects of the internal auditing function, including consulting services
- Continually monitor the effectiveness of the internal auditing function and develop performance measures
- Assure compliance with the internal auditing standards and the Code of Ethics
- Help the internal auditing function add value and improve organisational operations through surveys and dialogue with management
- Include both periodic and ongoing internal assessments
- Include an external assessment at least once every five years, the results of which are communicated to management and the audit committee or appropriate oversight board (IIA 2003: 56).

The omission of any of these elements from the quality programme represents non-compliance with the internal auditing standards. Standard 1330, which encourages internal auditors to report: ''that their activities are conducted in accordance with the International Standards for the Professional Practice of Internal Auditing'', also includes a provision that internal auditors may use the statement only if assessments of the quality programme demonstrate that the internal audit activity is in compliance with the internal auditing standards (IIA 2004: 12).

The purposes of the quality programme are threefold: firstly, to assess the effectiveness of the internal auditing function in providing assurance and consulting services to the board, senior executives and other interested parties; secondly, to assess conformance to the internal auditing standards and provide an opinion on whether the internal auditing activity generally conforms to all of them; and thirdly, to identify opportunities, offer recommendations for improvement and provide counsel to the chief audit executive and internal auditing staff for improving their performance and services and promoting the image and credibility of the internal auditing function (IIA 2003: 7).

The quality programme, and the regular assessment thereof, should thus assist the internal auditing function in adding value to and improving the activities of the enterprise, and ought also to ensure that the internal auditing function gives effect to the definition of internal auditing, meets the internal auditing standards and, ultimately, fulfils the needs of its clients.

Conclusion

Internal auditing functions should endorse the fact that the quality of their service is directly dependent on the following two factors:

- Knowledge of the requirements of their clients
- Provision of a superlative service that meets the needs of their clients.

If internal auditing functions provide a quality service that complies with the requirements of their users and adds value to their organisations, they will be able to ensure the viability of internal auditing functions. Producing products and services of quality demands purposeful action on the part of the provider (in other words, thorough research, planning and managing within each organisation). Marketing the services of the internal auditing function, as well as communication, feature prominently here. The services of the internal auditing function need to be made known within the

organisation, and good communication with all levels of management within the organisation is necessary in order to reach agreement on the role to be played by the internal auditing function within the organisation. Where there is conflict between the needs of clients and the services that the internal auditing function is able to deliver in terms of its ethical code and the internal auditing definition, negotiation with the client will be necessary until both parties have reached an agreement with which they are satisfied. Proper quality control within internal auditing functions is indispensable in order to ensure that internal auditing functions meet the requirements of their users on a continuous basis. Quality control also serves as an opportunity for open communication between management and the chief audit executive and as a marketing opportunity for keeping management abreast of the services that the internal auditing function can offer.

References

Binneman, D. 2002. 'Protecting organisational reputation: your role', *IIA IA Adviser.* May: 12–13.

Bishop III, W.G., Hermanson D.R., Lapides, P.D. & Rittenberg L.E. 2000. 'The year of the audit committee', *Internal Auditor,* 57(2) April: 46–51.

Bookal, L.E. 2002. 'Internal auditors: integral to good corporate governance', *Internal Auditor,* 59(4) August: 44–49.

Burns, S. 2000. 'Reinforcing internal audit value', *Internal Auditor,* 57(5) October: 80.

Carolus, R.N. & Nelson, D.J. 1996. *Business-focussed Quality Assurance Review Manual* (3rd edition). Altamonte Springs Florida: Institute of Internal Auditors.

Chapman, C. 1998. 'Just do it', *Internal Auditor,* 55(3): 38–41.

Chartered Institute of Public Finance and Accountancy. 1993. *Quality and the Internal Auditor. Part of the Shaping the Future Initiative.* London: Chartered Institute of Public Finance and Accountancy.

Clikeman, P.M. 1999. 'Improving information quality', *Internal Audito*r, 56(3): 32–33.

Drent, D. 2002. 'The quest for increased relevance', *Internal Auditor*, 59(1): 49–55.

Ewert, G.A. 1997. 'How to sell internal auditing', *Internal Auditor,* 54(5): 54–57.

Fabrizius, M.P. 1997. 'A passion for excellence', *Internal Auditor*, 54(5): 22–27.

Fabrizius, M.P. & Serafini, R.M. 2004. 'Learning to love the scrutiny', *Internal Auditor*, 61(1): 38–43.

Hodson, P. 2001. 'Quality the path to excellence. The internal auditor's ultimate goal', *IIA IA Adviser,* February: 16–17.

Institute of Directors in Southern Africa. 2002. *King Report on Corporate Governance for South Africa 2002.* Parkland: Institute of Directors in Southern Africa.

Institute of Internal Auditors. 1999. A Vision for the Future: Professional Practices Framework for Internal Auditing. Report of the Guidance Task Force to the IIA's board of directors. Altamonte Springs, Florida: Institute of Internal Auditors.

Institute of Internal Auditors. 2003. *Quality Assessment Manual* (4th edition). Altamonte Springs, Florida: Institute of Internal Auditors.

Institute of Internal Auditors. 2004. *The Professional Practice Framework.* Altamonte Springs, Florida: Institute of Internal Auditors.

Institute of Internal Auditors South Africa. 1997. 'Corporate governance. Doing the right things right', *IIA IA Adviser,* February: 2–3.

Lefevre, H.L. 1989. *Quality Service Pays. Six Keys to Success!* Milwaukee, WI: American Society for Quality Control.

Leithhead, B.S. 2000a. 'Risk watch: ensuring audit relevance', *Internal Auditor*, 57(2): 68–69.

Leithhead, B.S. 2000b. 'Risk watch: product development risks', *Internal Auditor*, 57(5): 64–69.

Leithhead, B.S. 2000c. 'Risk watch: in touch with the top', *Internal Auditor*, 57(6): 67–69.

Marks, N. 2003. 'Training the audit committee', *Internal Auditor*, 60(6): 41–45.

Miller, L. 2000. 'The looking glass', *Internal Auditor*, 57(5): 27–35.

Payne, N. 2002. 'Preventing another Worldcom', *IIA IA Adviser,* September: 7.

Reding, K.F., Barber, C.H. & Digirolamo, K.K. 2000. 'Benchmarking against CFIA', *Internal Auditor*, 57(4): 41–46.

Rezaee, Z. 2002. 'The three C's of fraudulent financial reporting', *Internal Auditor*, 59(5): 56–61.

Ridley, J. & Stephens, K. 1996. *International Quality Standards: Implications for Internal Auditing.* Altamonte Springs, Florida: Institute of Internal Auditors Research Foundation.

Rittenberg, L.E. 2002. 'Lessons for internal auditors', *Internal Auditor,* 59(2): 32.

Rittenberg, L.E. & Covaleski, M.A. 1997. *The Outsourcing Dilemma: What's Best for Internal Auditing*. Altamonte Springs, Florida: Institute of Internal Auditors Research Foundation.

Russell, J.P. 1990. *The Quality Master Plan. A Quality Strategy for Business Leadership.* Wisconsin: American Society for Quality Control.

Steinberg, R.M. & Pojunis, D. 2000. 'Corporate governance: The new frontier', *Internal Auditor*, 57(6): 34–39.

Van Wyk, A. 2000. 'Business unusual – internal audit's changing role'. *IIA IA Adviser.* April: 8–9

Van Wyk A. 2002. 'Internal audit transformation', *IIA IA Adviser*, July: 14–18.

Wagner, J.K. 2000. 'Evaluating your internal auditor', *Directors and Boards.* 24(4): 1–5.

Walz, A. 1997. 'Adding value: creating value has become a matter of survival', *Internal Auditor,* 54(1): 51–54.

Wessels, R. 2002. Risk management issues facing directors. *IIA IA Adviser*, July: 8–11.

The tax deductibility of donations, with specific reference to donations of property made in kind to public benefit organisations

R. Oberholzer[*]

Section 18A of the Income Tax Act (Act 58 of 1962), as amended by the Revenue Laws Amendment Act (Act 45 of 2003) entitles a taxpayer (an individual or a legal entity, including a trust) to deduct annual donations to certain public benefit organisations, not exceeding 5% of taxable income. The question arises whether the same tax deduction applies when a taxpayer decides to make a donation in cash as opposed to a donation of property in kind. The aim of this paper is to compare the effect on a taxpayer's taxable income of making a cash donation compared with a donation of property in kind. It appears that the tax deduction in respect of donations is greater when a taxpayer decides to donate an amount in cash rather than a donation of property in kind. The paper shows that, under current legislation, a donation of property made in kind can be structured in such a way that it will provide a taxpayer with an identical tax deduction. It is hoped that current legislation pertaining to the deduction of a donation of property in kind will be amended soon, as the provisions are clearly inequitable in relation to a taxpayer who wishes to donate property in kind rather than a cash amount to a public benefit organisation.

Deduction of donations – an introduction

According to the Tax Exemption Guide for Public Benefit Organisations in South Africa, issued by the South African Revenue Service (SARS) (2004: 9), it is widely accepted that the tax deductibility of donations influences donor behaviour. Government has recognised this, and donations to a limited number of categories of public benefit organisations may be deducted from the taxable income of the donating taxpayer.

The question arises whether the same tax deduction applies when a taxpayer decides to make a donation in cash as opposed to a donation of property in kind. The aim of this paper is to compare the effect on a taxpayer's taxable income of making a cash donation compared with a donation of property in kind.

Research objective and research problem

The research objective of this study is to determine whether a taxpayer qualifies for the same tax deduction when he or she decides to make a donation in cash as opposed to a donation of property in kind.

The paper concentrates on the sections of the South African Income Tax Act (Act 58 of 1962) (the Act) relevant to a taxpayer who wishes to donate either an amount in cash or property in kind to a public benefit organisation.

The paper is limited to a consideration of income tax implications. No value-added tax or other indirect tax implications are taken into account for the purposes of this paper.

Research strategy

The research method followed included the following:

- A review of the Act was carried out in order to analyse the applicable sections in the Act relating to a taxpayer who makes a donation.
- SARS officials at the Law Interpreters division (Brooklyn, Pretoria) were interviewed in October 2003 in order to establish the view of SARS on issues identified in the paper.
- Examples were used to illustrate the effect on a taxpayer's taxable income of making a donation in cash compared with a donation of property in kind.

* R. Oberholzer is in the Department of Taxation, University of Pretoria. E-mail: ruanda.oberholzer@up.ac.za.

The paper is based on current income tax legislation. The most recent law amendments considered in this paper are the Second Revenue Laws Amendment Act (Act 45 of 2003), which was promulgated on 22 December 2003.

The research strategy involved the following steps:

- All the relevant tax provisions that need to be taken into account when making a donation were analysed.
- Examples were set out to illustrate the effect on a taxpayer's taxable income of making a donation in cash or of property in kind.
- The comments received from SARS on the issues identified were set out.
- The paper recommends an alternative approach for taxpayers in order to overcome the inequalities that have been identified in the paper.

Tax provisions relating to donations

All the relevant tax provisions that need to be taken into account when making a donation were analysed as a basis for the rest of the study.

Deduction of the donation

Section 18A of the Act, as amended by the Revenue Laws Amendment Act (Act 45 of 2003) (Department of Finance 2003a), entitles a taxpayer (an individual or a legal entity, including a trust) to deduct annual donations to certain public benefit organisations, not exceeding 5% of his or her taxable income. Before the Revenue Laws Amendment Act was enacted, the allowable deduction was limited to the greater of R1 000 or 5% of the taxpayer's taxable income. The limit of R1 000 was repealed because 5% of taxable income at a tax threshold of R30 000 equals R1 500, and granting a R1 000 deduction serves no purpose, as was explained in the Explanatory Memorandum on the Revenue Laws Amendment Bill (Department of Finance 2003b: 59). The provision for a minimum donation of R1 000 ensured that a taxpayer who incurred an assessed loss would not be denied the benefit of a deduction (Stiglingh, Venter & Hamel 2003a: 95). The scenario of a taxpayer in a loss situation is not addressed at all in the Explanatory Memorandum on the Revenue Laws Amendment Bill (Department of Finance 2003b: 59). It is submitted that a taxpayer in a loss situation will forfeit the deduction of a qualifying donation. The Act contains no provision in respect of the unclaimed portion of donations being carried forward to a subsequent year of assessment.

No claim for a deduction for any donation is allowed unless the claim is supported by a receipt issued by the public benefit organisation concerned in accordance with section 18A(2) of the Act. Annexure I of the Tax Exemption Guide for Public Benefit Organisations in South Africa (SARS 2004: 39) provides an example of such a receipt.

The 'taxable income' contemplated in the section is the donor's taxable income as calculated before allowing a deduction for medical expenses and the deduction of the donation (see section 18A(1)(*c*) of the Act).

According to section 18A(1) of the Act, the allowable deduction is in respect of *bona fide* donations in cash or of property made in kind, actually paid or transferred during the year of assessment, to any

- public benefit organisation approved by the Commissioner under section 30; or
- institution, board or body contemplated in section 10(1)(*c*A)(i)

  which –
- carries on in the Republic any public benefit activity contemplated in Part II of the Ninth Schedule.

This approval of a public benefit organisation is subject to the condition that the section 18A public benefit activities are ring-fenced, and to certification by an auditor that the donations for which tax-deductible receipts have been issued have indeed been utilised solely in carrying out such public benefit activities. The public benefit activities that qualify for section 18A approval must be carried out in South Africa (SARS 2004: 10).

Conduit funds providing funds or assets to public benefit organisations listed in Parts I or II of the Ninth Schedule also qualify for section 18A approval. The conduit fund will qualify if, during the year of assessment, the fund distributes or incurs the obligation to so distribute at least 75% of the funds received during its preceding year of assessment by way of donations for which tax deductible receipts were issued. The Commissioner may, on good cause shown, and subject to conditions, either generally or in a particular instance, waive, defer or reduce the organisation's obligation to make this distribution, having regard to the public interest and the purpose for which the organisation wishes to accumulate the funds (section 18A(1)(*b*)

of the Act). Income generated by donations received is not required to be distributed (Meyerowitz 2003: 12–50).

Value of the donation

When a taxpayer makes a donation in cash, the amount of the qualifying deduction can easily be determined (being the actual amount donated). Where property in kind is donated, however, section 18A(3) of the Act determines the amount that qualifies for a possible deduction. This section of the Act has four subsections dealing with different types of donation of property in kind.

The amount of the deduction allowed depends on the type of property:

- If it is trading stock (including a financial instrument which is trading stock and livestock or fresh produce of farmers), the deduction is the lower of
  - the market value (fair market value for financial instruments) on the date of the donation; or
  - the amount taken into account in respect of the value of the trading stock for the purposes of section 22(8) or paragraph 11 of the First Schedule of the Act (section 18A(3)(*a*)).
- If it is an asset used by the taxpayer for the purpose of his trade, the deduction is the lower of
  - the fair market value on the date of the donation; or
  - the cost to the taxpayer of such property, less any allowance deducted from the income of the taxpayer under the provisions of the Act in respect of that asset (section 18A(3)(*b*) of the Act).

  According to Huxham & Haupt (2004: 168), the amount of the donation where a taxpayer donates trading assets is deemed to be the lower of the cost of the asset less any tax allowances (the tax value), or the fair market value of the asset on the date of the donation.
- If it is another asset (not trading stock or a business asset), the deduction is the lower of
  - the fair market value on the date of the donation; or
  - the cost to the taxpayer of such property, less any reasonable depreciation in the case of movable property which has deteriorated in condition (section 18A(3)(*c*) of the Act).

  Section 18A(3)(*c*) of the Act specifies that the depreciation allowance must be calculated in the manner contemplated in section 8(5)(*b*B)(i), which refers to depreciation using the 20% reducing balance method.
- If it is property purchased, manufactured, erected, assembled, installed or constructed by or on behalf of the taxpayer, the deduction is the lower of
  - the fair market value on the date of the donation; or
  - the cost to the taxpayer of such property (section 18A(3)(*d*) of the Act).

The amount qualifying for the possible deduction is still subject to the 5% limitation of a taxpayer's taxable income in terms of section 18A (see section on deduction of the donation).

Other income tax provisions to consider when making a donation

Section 22(8) – 'non-trade' disposal of trading stock

In terms of section 22(8) of the Act, where a taxpayer has applied trading stock for the purposes of making any donation, and the cost price of such trading stock has been taken into account in the determination of the taxable income of the taxpayer for any year of assessment, the taxpayer is deemed to have recovered or recouped an amount equal to the amount that has been taken into account for that year of assessment in respect of the value of that trading stock (the lower of the cost or the market value of such stock).

Section 8(4)(k) of the Act – taxing allowances previously claimed

Section 8(4)(*k*) comes into operation when a person donates an asset. Section 8(4)(*k*) applies when the asset is one on which a tax deduction or an allowance has been granted to the taxpayer in terms of sections 11 to 20 and sections 24D, 24F, 24G and 27(2)(*b*) and (*d*) of the Act. According to section 8(4)(*k*), a taxpayer who donates an asset, for which an allowance has been granted as already discussed, is deemed to have recovered or recouped an amount equal to the market value of such asset on the date of such donation.

Arendse, Coetzee, Jordaan, Kolitz, Stein & Stiglingh (2003: 224) state that section 8(4)(*k*) is

ambiguous. However, they add that the Explanatory Memorandum on the Income Tax Bill, 1993 makes the point that, although the market value of the asset is deemed to have been recouped, the actual inclusion in income is still governed by section 8(4)(*a*) (the general recoupment section). The result is that the amount actually subjected to tax is limited to the extent of the deductions or allowances previously granted.

Paragraph 11(1)(a) of the Eighth Schedule donation equals a capital gains tax event

In addition, with the introduction of capital gains tax, a donation is also specifically included in the definition of a disposal (paragraph 11(1)(*a*) of the Eighth Schedule to the Act). A taxpayer must therefore include the market value of the donation, being the proceeds of the disposal, where the disposal took the form of a donation (paragraph 38(2)(*a*) of the Eight Schedule to the Act). Any capital gain or loss is disregarded, however, if the donation is made to a public benefit organisation (paragraph 62 of the Eighth Schedule to the Act).

Donations tax implications of a donation

In terms of section 54 of the Act, donations tax is payable on the value of any property disposed of as a donation by a resident. The current rate of donations tax is 20% (section 64 of the Act). In terms of section 56(*h*) of the Act, certain donations made by a resident are specifically exempt from the payment of donations tax, including donations to any institution, council or body referred to in section 10(1)(*c*A). No donations tax is therefore payable on a donation to any public benefit organisation approved in terms of section 30(3) of the Act.

Summary

The treatment of donations of property made in kind can generally be divided into two groups, which are summarised as follows:

- Donations of trading stock and property purchased, manufactured, erected assembled, installed or constructed – it is clear that a taxpayer making a donation of property in kind (trading stock or property purchased, manufactured, erected assembled, installed or constructed) is taxed on the lower of cost or market value in terms of section 22(8), and that the qualifying amount in terms of section 18A of the Act is exactly the same amount, namely the lower of the market value or the cost of the stock item (the taxpayer will generally be in the same position if he or she donates trading stock or an amount of cash or she same value).
- Donations of trade and non-trade assets – a taxpayer who wishes to donate trade or non-trade assets on which a tax deduction or an allowance has been granted (in terms of sections 11 to 20 and sections 24D, 24F, 24G and 27(2)(*b*) and (*d*) of the Act) is taxed in terms of section 8(4)(*k*). The taxpayer is deemed to have recovered or recouped an amount equal to the market value of such an asset. The qualifying amount in terms of section 18A is, however, only the tax value of the particular asset.

The effect on a taxpayer's taxable income of making a donation in cash or of property in kind consisting of an asset (both trade and non-trade) will be illustrated in the next section by means of examples. The examples deal only with trade assets, but apply equally to non-trade assets.

Practical examples

The following examples illustrate the effect on a taxpayer's taxable income of making a donation in cash or of property in kind:

Example 1

A taxpayer owns a vehicle, which he bought three years ago at the beginning of his financial year for R120 000. He brought the vehicle into use immediately in his business. The market value of the vehicle at the end of the third year equals R80 000. The taxpayer claimed wear and tear on this vehicle for the three years in terms of section 11(*e*) (Practice Note 19) at 20% annually.

	Rand
Cost price	120 000
Wear and tear allowed (a section 11(*e*) deduction): Year 1	(24 000)
Wear and tear allowed (a section 11(*e*) deduction): Year 2	(24 000)
Wear and tear allowed (a section 11(*e*) deduction): Year 3	(24 000)
Tax value	**48 000**

The taxpayer decides to make a donation to an approved public benefit organisation. He can either donate R80 000 cash (Scenario 1) or donate this particular vehicle valued at R80 000 (Scenario 2) to the approved public benefit organisation. The taxpayer obtains the necessary receipt from the public benefit organisation (as required by section 18A of the Act) and seeks to deduct the donation from his taxable income.

Assuming that the taxable income of the taxpayer before this donation equals R2 500 000, the effect of the donation is as follows:

	Scenario 1 Donate R80 000 in cash Rand	Scenario 2 Donate vehicle valued at R80 000 Rand
Taxable income before donation	2 500 000	2 500 000
Plus section 8(4)(*k*) recoupment Market value (limited to original cost price) less tax value: (R80 000 – R48 000)	n/a	32 000
	2 500 000	2 532 000
Less section 18A deduction in respect of donation made to a public benefit organisation:		
Qualifying deduction: Scenario 1 R80 000 cash, limited to 5% of R2 500 000 (taxable income before this deduction) = R125 000	80 000	
Qualifying deduction: Scenario 2 The lower of: • the tax value of the asset (R48 000); or • the fair market value (R80 000).		
R48 000, limited to 5% of R2 532 000 (taxable income before this deduction) = R126 600		48 000
Final taxable income	**2 420 000**	**2 484 000**

The difference of R64 000 (R2 484 000 – R2 420 000) consists of two amounts:

- R32 000 recoupment in terms of section 8(4)(*k*); and
- R32 000 (R80 000 – R48 000) in terms of section 18A in respect of donations made, if the taxpayer decides to donate an amount in cash rather than of property in kind (the vehicle) to the public benefit organisation.

Taxpayers who donate cash instead of property in kind (by means of a trading asset) with the same market value therefore receive a much greater tax advantage.

Example 2

Where a taxpayer identifies a particular asset as redundant in a business, that particular asset might be identified as a suitable donation to a public benefit organisation. One can compare this to a situation in which the taxpayer decides to sell the asset and then donates the cash generated by the sale, instead of donating the asset itself. One can examine these possibilities assuming the same details as in example 1, but with Scenario 1 representing the situation in which the taxpayer sells the vehicle for R80 000 (the market value) and Scenario 2 representing the situation in which the taxpayer donates the vehicle itself.

Assuming once again that the taxable income of the taxpayer before this donation equals R2 500 000, the effect of the donation is as follows:

	Scenario 1 Sell the vehicle and donate R80 000 in cash Rand	Scenario 2 Donate vehicle valued at R80 000 Rand
Taxable income before donation	2 500 000	2 500 000
Plus section 8(4)(*a*) recoupment Selling price (limited to original cost price) Less tax value: (R80 000 – R48 000)	32 000	
Plus section 8(4)(*k*) recoupment Market value (limited to original cost price) Less tax value: (R80 000 – R48 000)		32 000
	2 532 000	2 532 000
Less a section 18A deduction in respect of a donation made to a public benefit organisation:		
Qualifying deduction: Scenario 1 R80 000 cash limited to 5% of R2 532 000 (taxable income before this deduction) = R126 600	80 000	
Qualifying deduction: Scenario 2 The lower of:		48 000
• the tax value of the asset (R48 000); or • the fair market value (R80 000). R48 000 limited to 5% of R2 532 000 (taxable income before this deduction) = R126 600		
	2 452 000	2 484 000
Capital gains tax		n/a
Proceeds (R80 000 – R32 000) less Base cost (R120 000 – R72 000) = Rnil	nil	
Final taxable income	**2 452 000**	**2 484 000**

The difference of R32 000 (R2 484 000 – R2 452 000) consists of one amount, namely, R32 000 (R80 000 – R48 000), in terms of section 18A, for donations made when a taxpayer decides to donate cash rather than property in kind (the vehicle) to the public benefit organisation.

Taxpayers who donate cash realised from the sale of a redundant asset instead of property in kind (by means of a trading asset) with the same market value, therefore still receive a greater tax advantage. This clearly does not make sense. In an effort to clarify this inconsistency, SARS was approached for comment.

Comment from SARS

The anomaly of receiving a greater tax deduction when a taxpayer decides to donate an amount in cash, compared with a donation of property in kind, was pointed out to SARS. An official from the SARS Law Administration division, in an e-mail dated 31 October 2003, commented as follows:

> Section 8(4)(*k*) was designed to achieve the same result that a person would have if the asset was sold for cash and the cash was donated. In other words it is designed to prevent people escaping recoupment by donating assets instead of cash.
>
> I suspect that tax value was allowed in terms of section 18A to avoid giving a double deduction for the same asset. For example, if I bought an asset for R100, depreciated it fully and then donated it when its market value was R100, had market value been allowed I would have had R200 as a deduction (R100 as depreciation, and R100 under section 18A). However, the use of tax value now seems wrong since section 8(4)(*k*) was introduced. It's probably something that needs to be amended.

The response received from SARS indicates that the current provisions, relating to the deduction of a donation of property made in kind in section 18A, are outdated and need to be amended. In the next section, an alternative approach to overcome these inequalities for taxpayers wanting to make a donation of property in kind is proposed.

What can be done in the interim?

When a taxpayer wants to make a donation to a public benefit organisation in the form of property in kind and obtain the same tax deduction as he or she would be allowed when making a donation in cash, he may prefer to sell the specific asset to the public benefit organisation at market value. The public benefit organisation then pays the selling price in cash to the taxpayer, who donates the exact amount to the public benefit organisation (for which he receives a valid section 18A receipt).

The only problem with this suggestion is that SARS may enforce the provisions of section 103(1) in relation to this transaction.

Section 103(1) of the Act reads as follows (author's emphasis):

> Whenever the Commissioner is satisfied that any **transaction**, operation or scheme (whether entered into or carried out before or after the commencement of this Act, and including a transaction, operation or scheme involving the alienation of property) –
>
> *(a)* has been entered into or carried out which has the effect of avoiding **tax**, duty or levy imposed by this Act or any previous Income Tax Act, or **reducing the amount thereof**; and
>
> *(b)* having regard to the circumstances under which the transaction, operation or scheme was entered into or carried out
>
> (i) was entered into or carried out
>
> *(aa)* in the case of a transaction, operation or scheme in the context of business, in a manner which would **not normally be employed for *bona fide* business purposes, other than the obtaining of a tax benefit**; and
>
> *(bb)* in the case of any other transaction, operation or scheme, being a transaction, operation or scheme not falling within the provisions of item *(aa)*, by means or in a manner which would not normally be employed in the entering into or carrying out of a transaction, operation or scheme of the nature of the transaction, operation or scheme in question; or
>
> (ii) has created rights or obligations which would not normally be created between persons dealing at arm's length

under a transaction, operation or scheme of the nature of the transaction, operation or scheme in question; and

(c) was entered into or carried out solely or mainly for the purposes of obtaining a tax benefit,

the Commissioner shall determine the liability for any tax, duty or levy imposed by this Act, and the amount thereof, as if the transaction, operation or scheme had not been entered into or carried out, or in such manner as in the circumstances of the case he deems appropriate for the prevention or diminution of such avoidance, postponement or reduction.

Before the Commissioner can apply the provisions of section 103(1) to set aside a transaction or act in such a manner as to prevent the avoidance of tax, all of the preceding requirements must be present:

- There must be a transaction, operation or scheme.
- It must have the effect of avoiding or postponing liability for tax.
- The manner in which a business transaction is entered into must be one that would not normally be employed for *bona fide* business purposes (the business purpose test or the abnormality test and an abnormal rights test) – according to Stiglingh, Venter & Hamel (2003b: 345), the manner in which a business transaction is entered into must be one that would not normally be employed for *bona fide* business purposes (according to the business purpose test or the abnormality test and an abnormal rights test, which requires an element of abnormality in either the transaction, operation or scheme or in the rights or obligations arising from it).
- Its sole or main purpose must have been to obtain a tax benefit (this includes any avoidance, postponement or reduction of liability for payment of any tax, duty or levy imposed under the Act or any other law administered by the Commissioner).

The onus of proof that the sole or main purpose was not tax avoidance, however, rests upon the taxpayer himself. If the Commissioner is satisfied that the conditions are met, as described, he determines a taxpayer's tax liability as if the transaction had not been entered into or in such manner as in the circumstances the Commissioner deems appropriate for the prevention or diminution resulting from the reduction of tax.

When a taxpayer sells a specific asset at market value to a public benefit organisation for cash and in turn donates exactly the same amount to that public benefit organisation (for which he receives a valid section 18A receipt) as already suggested, there is a **transaction** with the effect of **reducing the amount of tax** payable. The third requirement of being **abnormal** in either the transaction, or the rights or obligations arising from it, might be debatable. There have not been many cases dealing with the application of section 103(1), but all these clearly show that each case must be decided on its own merits and according to its own particular circumstances (Stiglingh et al. 2003b: 346).

In conclusion, the four canons of taxation (first formulated by Adam Smith in 1776 in *The Wealth of Nations*) are summarised as follows by Huxham & Haupt (2004: 2):

(i) The subjects of every State ought to contribute towards the support of the government, as nearly as possible in proportion to their abilities; that is, in proportion to the revenue which they respectively enjoy under the protection of the State. The expense of government to the individuals of a great nation is like the expense of management to the joint tenants of a great estate, who are all obliged to contribute in proportion to their respective interests in the estate.

(ii) The tax which each individual is bound to pay ought to be certain, and not arbitrary. The time of payment, the manner of payment, the quantity to be paid, ought all to be clear and plain to the contributor.

(iii) Every tax ought to be levied at the time, or in the manner in which, it is most likely to be convenient for the contributor to pay it.

(iv) Every tax ought to be so contrived as to both take out, and keep out, of the pockets of the people as little as possible over and above what it brings into the public treasury of the State.

The Katz Commission (1987: 55) also distinguished between horizontal and vertical equitability. Horizontal equitability determines that similar individuals or persons in the same situation must be treated equally. Vertical equitability, however,

means that persons in different situations must carry different tax liabilities (individuals with a higher level of economic prosperity must carry a higher tax burden).

In the modern context, these principles must also include the broader principles of social justice (Stiglingh et al. 2003a: 3).

From the discussion in this paper, it is clear that the provisions relating to a taxpayer who wishes to donate a depreciable asset instead of a cash amount to a public benefit organisation are in conflict with the basic principles of taxation.

Conclusion

This paper shows that the tax deduction in respect of donations is greater when a taxpayer decides to donate an amount in cash rather than of property in kind.

SARS responded in this regard by agreeing that the current provisions relating to the deduction of a donation of property made in kind in section 18A are outdated and need to be amended.

The paper shows that, under current legislation, a donation of property in kind can be structured in such a way that it will provide a taxpayer with an identical tax deduction as when donating cash.

The paper concludes with a statement that the provisions relating to a taxpayer who wishes to donate a depreciable asset instead of a cash amount to a public benefit organisation are in conflict with the basic principles of taxation.

It is hoped that the current legislation pertaining to the deduction of a donation of property made in kind will be amended soon, as the provisions are clearly inequitable towards a taxpayer who wishes to donate a depreciable asset rather than cash to a public benefit organisation.

References

Arendse, J.A., Coetzee, E., Jordaan, K., Kolitz, M.A., Stein, M.L. & Stiglingh, M. 2003. *Silke: South African Income Tax 2004.* Durban: LexisNexis Butterworths.

Huxham, K. & Haupt P. 2004. *Notes on South African Income Tax 2004.* Roggebaai: Hedron Tax Consulting and Publishing.

Katz Kommissie. 1987. Verslag van die Kommissie van ondersoek na die Republiek van Suid-Afrika se belastingstruktuur. Pretoria: Staatsdrukker.

Meyerowitz, D. 2003. *Meyerowitz on Income Tax 2003–2004 The Taxpayer,* Cape Town.

South African Revenue Service (SARS). 2004. *Tax Exemption Guide for Public Benefit Organisations in South Africa.* [Online]. Available from: www.sars.gov.za/it/pbo. Accessed: 28 May 2004.

South Africa. Department of Finance. 2003a. Revenue Laws Amendment Act (Act 45 of 2003). *Government Gazette No 25864*. Pretoria: Government Printer.

South Africa. Department of Finance. 2003b. *Explanatory Memorandum on the Revenue Laws Amendment Bill 2003*. [Online]. Available from: www.sars.gov.za/legislation/bills_memos/memos/2003/Revenue. Accessed: 11 May 2004.

Stiglingh, M., Venter, J.M.P., & Hamel, E.H. 2003a. *A Student's Approach to Income Tax Natural Persons*. Durban: LexisNexis Butterworths.

Stiglingh, M., Venter, J.M.P, & Hamel, E.H. 2003b. *A Student's Approach to Income Tax Business Activities*. Durban: LexisNexis Butterworths.