CHAPTER 4
RESEARCH METHODOLOGY

4.1 INTRODUCTION

Oberholzer (2008:ii) recently found that non-compliance by taxpayers is one of the main causes for the significant gap between the amount of tax theoretically collectable from economically active persons and that actually collected in South Africa. The development of a service quality model for the assessment of the services provided by SARS is therefore justified, because it is an essential means to improving the services of SARS and therefore also to improving voluntary compliance.

The objective of the present research is to develop a service quality model that can be used to establish the perceptions of tax practitioners with regard to the service quality of SARS. In the previous chapters, the relevant theoretical constructs with regard to services, quality, service quality and perceptions of service quality have been identified, described and analysed. In order to achieve the objective of the present research, it is important to build on these theoretical constructs to develop a service quality model that can be used as a framework for a quantitative survey instrument to measure the service quality of SARS. In this chapter, the research design used in the present research to achieve this objective is described.

4.2 RESEARCH ORIENTATION

The research can be categorised as falling within the qualitative paradigm and, more specifically, an interpretive orientation, which is an approach that seeks to understand phenomena and to develop theory or build models or frameworks that can be tested empirically in later research (Cooper & Schindler 2001; Leedy & Ormrod 2005; Welman *et al*. 2005). For this reason, the research problem was not stated in the form of null hypotheses which the research could then attempt to reject using statistical techniques, but was instead framed as a broad research objective.

Although the research was mainly qualitative in its approach, it also had a positivist underpinning, as it was based on the broad premise that there is an ideal norm or standard against which the service delivery levels of SARS can be tested. The research did not

seek merely to understand the service quality construct, but also to develop a model based on an ideal standard or norm.

4.3 DEFINITIONS

In the process of developing a model based on an ideal standard or norm (in other words, the service quality model proposed in the present research), the research focused on a number of key concepts. The theoretical underpinnings of the definitions of **service**, **quality**, **service quality**, **perceived service quality, service dimensions, service determinants and service aspects** have already been discussed in Chapters 2 and 3. The relevant definitions of the abovementioned concepts as adapted for the present research are summarised below.

Services were defined as differing from goods. They were analysed with reference to their characteristics, namely the intangibility, relative inseparability, interdependence and heterogeneity of services (Boshoff 1990; Eiglier & Langeard 1977; Grönroos 1978; Schneider & White 2004; Upah & Fulton 1985). For the purposes of the present research, all the actions taken by SARS in collecting taxes were therefore regarded as services.

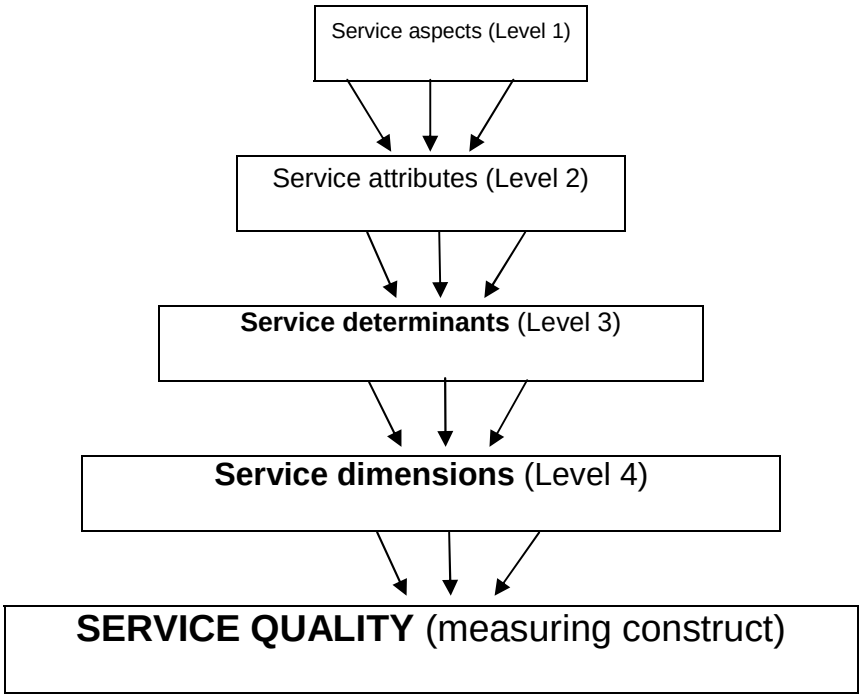
In the present research, a user-based approach to defining **quality** was predominantly used, but the principles of the manufacturing approach were also incorporated to ensure that the full spectrum of quality, as perceived by tax practitioners, was measured. In a user-based approach to defining quality, the definition of quality as formulated by the customer (in this case, the tax practitioner) is relevant. It is therefore subjective, hinging on the perceptions of individual tax practitioners – the services that best satisfy their preferences are those they regard as having the highest quality (Berry *et al*. 1985:45; Boothe 1990:65; Garvin 1984:27; Gummesson 1992:184; Philip & Stewart 1999:2; Schneider & White 2004:10). By contrast, **manufacturing quality** refers to a precise and measurable variable (Garvin 1984:25). In terms of a measurement of manufacturing quality, services are regarded as special types of goods. High quality then implies that goods are compliant with standards and free of deficiencies (Juran 1988:5; Klaus 1985:19).

For the purposes of the present research, it is acknowledged that the combined term **service quality** is a multidimensional, hierarchical construct. This means that customers

form their service quality perceptions on the basis of an evaluation of performance at multiple levels (Grönroos 1984, 1988; Gummesson 1992; Kang & James 2004; Parasuraman *et al*. 1985, 1986, 1988; Parasuraman *et al*. 1991a). Figure 4.1 illustrates the different levels in the multidimensional service quality construct. In the present research, the first level on which a customer forms his or her service quality perception is called **service aspects** (Level 1). This refers to the detailed service aspects that contribute to service factors (service attributes) for each specific service. The term **service attributes** (Level 2) refers to service factors relevant to service quality evaluations (Dabholkar *et al*. 2000:169). The **service determinants** (Level 3) represent a conceptual framework for summarising the service attributes tax practitioners use in assessing service quality (Parasuraman *et al.* 1991a:440). The **service dimensions** (Level 4) are similar in meaning to service quality components (Dabholkar 2000 *et al*. :169).

An example is helpful in assisting in an understanding of the multiple levels in the service quality construct. One service attribute (factor) that the participating tax practitioners identified as relevant to the service quality of SARS's traditional services is the communication service attribute (Level 2). In defining the communication service attribute, it was found that this service attribute could be divided into different sub-service attributes, which are referred to as service aspects (Level 1) in the present research. The understandability of contact personnel, the understandability of documentation and the communication skills of employees were, for example, identified as three different service aspects that are all antecedents to the service attribute of communication. In turn, the communication service attribute (Level 2) was found to be an antecedent of the empathy service determinant (Level 3) in the functional service quality dimension (Level 4). The different definitions for each service aspect, service attribute, service determinant and service dimension are presented in Chapters 5 and 6, together with the development of the proposed service quality model.

Figure 4.1: Multiple levels in service quality construct



The proposed service quality model is to be used in future as a framework for a quantitative survey instrument to measure the service quality of the services SARS renders as perceived by tax practitioners. The model would therefore not measure service quality as such, but would measure **perceived service quality**. For the purposes of the present research, the perceived service quality of the services SARS provides would tend, firstly, to be viewed subjectively by the tax practitioner; secondly, to be predominantly a cognitive and, to a lesser extent, an affective judgement; thirdly, to be represented by the difference between the perceived performance and expectations, and fourthly, to be related to, but not equivalent to, satisfaction (refer to Section 2.5).

With the rapid expansion of information technology, it has now also become necessary to distinguish between **traditional services** and **e-services**. **Traditional services** refer to all non-internet or non-electronic customer interactions and experiences with suppliers (Parasuraman *et al*. 2005:214). For the purposes of the present research, SARS's **e-services** are regarded as all services provided through the internet or through SARS's e-filing (refer to Section 3.5). E-mail was therefore still classified as a traditional service, because, with modern technology, faxes are often delivered in the e-mail inbox of the recipient. Hence, it was difficult to distinguish between these two service channels. Consequently, e-mail was regarded as only one of the service channels of traditional services.

4.4 THE UNIT OF ANALYSIS AND THE POPULATION

The unit of analysis and population consisted of all the tax practitioners registered with SARS in terms of section 67A of the Income Tax Act at the time when the questionnaires were distributed. From SARS's perspective, the term "tax practitioner" refers to and includes any person giving advice to any other person in respect of an Act administered by the Commissioner, or who, for reward, completes or assists with the completion of any document to be submitted to the Commissioner.

Tax practitioners were chosen as the target population because such practitioners play a crucial role in enhancing the efficiency of tax collection. It is they who communicate SARS's intentions to their clients (Friedman 2003:13). They fulfil an important mediating role and serve as a conduit for SARS by passing information to the most affluent taxpayers, which may inevitably influence decisions taken by these taxpayers, as Smith (2003:11) points out. Even if they try to think of ways of "beating the system", they remain essential cogs in its workings (Friedman 2003:13).

Between 60% and 80% of South African businesses use tax practitioners to help them to comply with the burdens of tax compliance (SBP 2005:49; UNISA 2005:114; Upstart Business Strategies CC 2004:36). During extensive consultations with the small business sector, it became clear that an even higher percentage of approximately 95% of small and medium enterprises have to outsource some of their compliance issues to tax practitioners (Arendse, Karlinsky, Killian & Payne 2006:17).

However, in South Africa, it is not only businesses that rely on the assistance of tax practitioners. Turner, Smith and Gurd (1998:99) argue that limited leisure time is available to individuals, so that when they have large enough disposable incomes, individuals are also prepared to pay tax practitioners to complete their returns, because it is faster and more convenient to do so than to try to learn about the requirements themselves. Furthermore, it is estimated that the tax practitioners represent approximately 4 million of the 6.3 million taxpayers in South Africa (SARS 2007:19; Snyckers 2006:4). SARS (2007:39) also indicated that it wants to form a strategic alliance with advisors and tax practitioners to ensure that they are provided with a differentiated service.

Previously, the views of the tax practitioners might have been biased, as SARS can be the subject of severe hostility from the tax practitioners (Friedman 2003:6). The employment of

tax specialists from the private sector by SARS, has, however, given the tax authority a better understanding of business culture, and a more effective basis for engaging with tax consultants working for taxpayers in the private sector (Centre for the future state 2005:16). During 2006, SARS launched a specialised tax practitioners' unit that specifically focuses on increased service delivery to practitioners. Various initiatives undertaken by SARS have also already favourably influenced the relationship between SARS and tax practitioners. SARS lists four different categories of taxpayers on its website, one of which is tax practitioners.

It is therefore submitted that tax practitioners not only represent the majority of taxpayers, but that the frequency of their interaction with SARS is probably much higher than that of an individual taxpayer. Therefore, they are probably the individuals best able to identify service excellence and deficiencies with regard to the services SARS delivers.

The entire population of registered tax practitioners was used for the present research, as the aim of the present research was not to use inferential statistics to predict outcomes. The total population of tax practitioners is approximately 17 000 (Snyckers 2006). Access to SARS's tax practitioners' database was granted by the then head of the tax practitioners' unit, Ms Telita Snyckers.

The client must experience the provision of the service in order to begin to form a mental representation of it (Eiglier & Langeard 1977: 42) and therefore only tax practitioners – who already engage with SARS – were included in the population.

4.5 THE DATA AND ITS COLLECTION

The first step in the research was a detailed literature review, which was carried out to establish the definitions of service, quality, service quality and perceived service quality to be used in the research. The outcome of the literature review served as a theoretical underpinning for the development of the proposed service quality model.

The literature review suggested that a user-based approach to quality was the most relevant to the present research – as Johnson and Gustafsson (2000:47) put it, it is important to build the "lens of the customer". Doing so will assist SARS to fulfil its aim of using a measuring instrument that captures the tax practitioners' perceptions of the quality of the services rendered by SARS. The "lens of the customer" in other service contexts

(typologies of service determinants and models) that have been presented so far is intended to serve only as a framework for an instrument to measure service quality, as suggested by Schneider and White (2004:38). In other words, such typologies are designed to be modified and changed to fit the needs of specific contexts (Schneider & White 2004:38,40). In order to develop the specific "lens of the customer" for evaluating the services of SARS, an in-depth, qualitative approach was required to identify a comprehensive range of determinants that potentially drive service quality in the revenue service industry and setting, as suggested by Johnson and Gustafsson (2000:47). One such qualitative method is the critical incident technique (CIT).

4.5.1 What is the critical incident technique?

The critical incident technique was introduced and originally defined by Flanagan (1954:1) as

> a set of procedures for collecting direct observations of human behaviour in such a way as to facilitate their potential usefulness in solving practical problems and developing broad psychological principles. The critical incident technique outlines procedures for collecting observed incidents having special significance and meeting systematically defined criteria.

Bitner, Booms and Tetreault (1990:73) argue that the critical incident technique is essentially a classification technique employing a content analysis of stories or "critical incidents" as data.

The critical incident technique relies on a set of procedures to collect comments on service experiences, to perform a content analysis and to classify the observations of service experiences. The specific descriptions of events are identified as critical incidents. Several authors have tried to define the critical incident technique – they are generally of the opinion that this method does not consist of a single rigid set of rules governing data collection. They appear to agree that the method should rather be thought of as a flexible set of principles which should be modified and adapted to meet the specific situation at hand (Flanagan 1954:9; Urquhart *et al*. 2003:63).

4.5.2 Applicability of the critical incident technique

Gremler (2004:67) claims that, generally speaking, the critical incident technique method

has been demonstrated to be a sound method since Flanagan first presented it in 1954. One of the advantages of the critical incident technique is that the context is developed entirely from the respondent's perspective and in his or her own words, and that the observations are not restricted to a limited set of variables or activities (Bitner *et al*. 1990:73; Gremler 2004:66-67; Odekerken-Schröder, Van Birgelen, Lemmink, De Ruyter & Wetzels 2000:110). An additional advantage is that critical incidents provide concrete areas for improvement from a customer's point of view (Odekerken-Schröder *et al.* 2000:110). Finally, it is contended that the critical incident technique is not particularly culturally-bound and that there is no prior determination of what will be important (Gremler 2004:67). The classification of critical incidents also allows for the identification of customer-defined service determinants, allowing more freedom in measuring service quality and preventing researchers' "blind spots" (Odekerken-Schröder *et al.* 2000:109). The critical incident technique is thus exactly what is required for building the "lens of the customer". Johnson and Gustafsson (2000:52) regard this technique as particularly well suited for this purpose.

The critical incident technique provides a valuable means for service researchers to study a phenomenon rigorously and to identify issues that have not previously been considered. It has been used successfully in a variety of service contexts in the last three decades: more than 140 critical incident technique studies have appeared in marketing research in the service context. More than 125 of these studies have been published since 1990 (Gremler 2004:65, 68, 69, 78). The most frequently researched issue using the critical incident technique is customer evaluations of service (31% of its use), including issues related to service quality (Gremler 2004:71). As is the case in the present research, the critical incident technique has primarily been used in a business-to-customer context (Gremler 2004:77).

The method itself appears to be a credible approach for service researchers to use, as virtually none of 168 studies investigated by Gremler (2004) have identified any substantial problems with the method itself. Odekerken-Schröder *et al.* (2000:109) found it useful to incorporate critical incidents in a relationship-oriented assessment of service quality. In the focus groups that Parasuraman *et al.* (1985) conducted for SERVQUAL, they also employed the critical incident technique to elicit examples of when customers were satisfied with a service and when they were not (Schneider & White 2004:54). The study

by Bitner *et al*. (1990:83) also supports the appropriateness of the critical incident technique for studying marketing questions and for assessing customer perceptions.

4.5.3 Conclusion on the use of the critical incident technique

The critical incident technique was chosen as the method to be used for building the "lens of the customer" for the evaluation of the tax practitioner's (customer) evaluation of the service quality of SARS, because the evaluation of a tax practitioner's perceptions of the service quality of SARS

- is a relationship-oriented assessment of service quality (Odekerken-Schröder *et al.* 2000);
- is done by the customers (tax practitioners) (Bitner *et al*. 1990; Odekerken-Schröder *et al.* 2000);
- is done in the business-to-customer context (Gremler 2004);
- seeks to provide the answer to a question in the service research environment (Gremler 2004); and
- is measured where the user-based approach of quality has been identified as the most suitable approach to apply (Johnson & Gustafsson 2000; Parasuraman *et al.* 1985).

4.5.4 Purpose of using the critical incident technique in the present research

Although an attribute-based measurement of service quality (the evaluation of identified service determinants) differs substantially from an incident-based assessment of service quality (the evaluation of specific service incidents) with regard to its methodology and perspective, there are some advantages to using both approaches in a complementary manner (Stauss & Weinlich 1997:34). Thus far, the call for an integration of these two methods has not been answered in any substantial manner, but generally researchers do not rely solely on the critical incident technique data as a single method in their attempts to understand the phenomenon of interest – the critical incident technique is usually used as a companion research method (Gremler 2004:76). Various researchers, such as Gremler (2004:67), Johnson and Gustafsson (2000) and Schneider and White (2004:54), recommend the critical incident technique as a way to generate qualitative data, which is a good starting point for developing quantitative models or measures of service quality. The critical incident technique method has also been used in the development of quantitative survey instruments by Martin (1996) and Miller, Craighead and Karwan (2000).

The studies that have examined service quality in service industries to date have generally been concerned with the dimensions of service quality and the identification of service quality determinants (Bitner *et al*. 1990:73; Urquhart *et al*. 2003:64). The purpose of using the critical incident technique in the present research was to assist in the development of a service quality framework to be used to develop a quantitative survey instrument to measure the quality of the service SARS provides. The critical incidents that were gathered were classified into categories of different service determinants (using content analysis), so that the important service determinants that are relevant and need to be incorporated into the service quality model could be identified. The critical incident technique was thus used in the present research both to confirm the service determinants identified in the literature review and to assist in the development of new service determinants.

4.6 SPECIFIC DATA COLLECTION METHOD

Data can be collected for the purposes of the critical incident technique from respondents at the time of observation or from observations made earlier and reported from memory, as suggested by Flanagan (1954:14). Because the services SARS provides involve more than only one-on-one interactions (for example, an assessment received in the mail) and because each respondent in the present research was required to list both positive and negative service experiences, the critical incidents that respondents were asked to report on were reported from memory. Gremler (2004:67) regards the need to report from memory as a possible disadvantage of the critical incident technique, as it relies on the chance that respondents will indeed remember incidents and it requires the “accurate and truthful reporting” of these incidents. Other researchers (Flanagan 1954:14; Johnson & Gustafsson 2000:56; Odekerken-Schröder *et al*. 2000:109) are not very concerned about this possible disadvantage, as they argue that it seems reasonable to assume that the incidents respondents recall can be relied on to provide adequate data, as critical incidents from customers’ memories are salient and relatively easy for customers to retrieve, as these incidents are at the top of their minds even in the long term. Flanagan (1954:14) also found that evidence regarding the accuracy of reporting is usually contained in the incidents themselves: if full and precise details are given, it can usually be assumed that this information is accurate, but vague reports suggest that the incident is not well remembered and that some of the data may be incorrect. In an attempt to avoid

generalisation, respondents in the present research were asked to be as specific as possible. (The words “Please be as specific as possible” were included in all the questions.)

Data reported from memory can be collected by means of personal interviews, group interviews or questionnaires (Flanagan 1954:15-18). The group interview technique has been well developed. It has the same advantages as an individual interview with regard to the amount of personal contact, explanation, and the availability of the interviewer to answer questions (Flanagan 1954:17). In this method, an interviewer presents introductory remarks to a group (very much as he or she would do in an individual interview). There is then an opportunity for questions and clarification (Flanagan 1954:17). Each person in the group is requested to write down incidents in response to specific questions contained in a specially prepared form (Flanagan 1954:17). Although such interviews were not the primary source of data in the present research, one group interview was conducted for the purposes of validating the final data collection instrument, as well as to assist in the development of the classification scheme used for the content analysis.

Apart from the interview method, critical incident data can also be collected by means of questionnaires (Gremler 2004:80; Flanagan 1954). Flanagan (1954:18) found that questionnaires tend to elicit results which are not essentially different from those obtained by means of the interview method. Except for the addition of introductory remarks, the forms used in collecting critical incidents by means of questionnaires are virtually the same as those used in group interviews. The questionnaire method was successfully used by Odekerken-Schröder *et al.* (2000) and Stauss and Weinlich (1997).

In addition to the group interview, the questionnaire option was chosen as the primary instrument for data collection for building the “lens of the customer” in the present research. The reason for choosing this method is that the results of the present research should be representative for SARS as a whole and not only for a specific region in which SARS operates, because SARS provides services throughout South Africa. By using a questionnaire, all the tax practitioners in the whole of South Africa could be reached easily, at the lowest possible cost. It was also anticipated that using a questionnaire would provide a greater number of responses, which would in turn result in a bigger pool of raw data.

To ensure that the tax practitioners who were included in the group interview did not also complete the questionnaire, in the introductory remarks made in the group interview, the respondents were told that a questionnaire was to be circulated and that part of it would duplicate what they would already have done in the group interview. They were requested not to answer those specific questions in the second questionnaire again. This is not an infallible method of ensuring that a particular tax practitioner would not complete the questionnaire twice, but completing a critical incident technique questionnaire is probably sufficiently time-consuming to deter most of the group interview participants from completing the form a second time. There is no guarantee, however, that practitioners have not submitted more than one questionnaire (this risk is discussed in greater detail in Section 4.8).

4.7 DESIGN OF THE DATA COLLECTION INSTRUMENT

Flanagan (1954:12) suggests that the first necessary specification for data collection is a delimitation of the situations that are to be observed. For the purposes of the present research, the respondents were asked to evaluate the **service quality of SARS as perceived by the tax practitioners in all interactions with SARS.**

Practitioners interact with SARS in respect of different types of taxes and by means of various service channels. With regard to the different types of taxes, observations on the service quality in connection with all taxes (excluding Customs and Excise) that are administered by SARS were requested.

Incidents relating to Customs and Excise duties were excluded from the data that was collected, because Customs and Excise duties are not paid by a significant proportion of the taxpaying population. A second reason for the exclusion is that most of the services rendered in administering the Customs and Excise duties are not part of the normal service channels. SARS itself acknowledges that Customs and Excise should be treated separately, as SARS circulated a customer satisfaction survey to all its Customs clients in January 2007, requesting information only on client satisfaction with regard to selected Customs services. Snyckers (2007a:pers. comm.), then head of the tax practitioners' unit at SARS, also agrees that Customs and Excise services are not part of the mainstream service channels offered by SARS.

The interactions through the SARS service channels include interactions with SARS at a local branch office or through a call centre, e-mail, e-filing, SARS's website, mail or fax. In order to ensure that respondents considered all the possible service channels, Questions 1 and 2 included a list of all the possible service channels. The respondents were not required to comment separately on each of the service channels. The positive responses on all the service channels were grouped into Question 1, while the negative responses on all the service channels were grouped into Question 2. In the web-based questionnaire, text messaging was added as a service channel.

It was considered appropriate to incorporate a business process approach in the service quality measurement instrument, as suggested by Rust *et al*. (1995). In order to evaluate the services SARS provides, it was decided to measure the different services separately, because services are regarded as intangible (Boshoff 1990:37; Eiglier & Langeard 1977:36; Grönroos 1978:591; Schneider & White 2004:6; Upah & Fulton 1985:255) and because each of SARS's services can lie at a different point on the separability-inseparability continuum. This choice was confirmed by referring to Haywood-Farmer's (1988) conceptual model of service quality, which suggests that it may be important to rate the services provided more specifically in terms of individual departments within SARS, to ensure that the results from the survey can be used in practice to improve the design of the processes of the service offerings. The evaluation of the service channels would possibly only have resulted in an evaluation of the functional quality (the "how") of the services of SARS.

In an attempt to prompt the respondents to evaluate the technical quality (the "what") of the service as well, or at least aspects of the technical quality of the service, the various tax processes that are dealt with by different departments in SARS were listed separately and were added as a second list of observations that respondents were asked for. These processes relate to tax registrations, the submission of tax returns, tax payments, tax refunds, account queries, updating details, assessments and dispute resolution issues (Questions 3 and 4 addressed the respondents' positive and negative experiences with regard to these aspects). For the web-based questionnaire, status queries (that is inquiries with regard to the status of any service aspect, for example, the status of a request for a change of address, or the status of a tax return that is in the tax assessment business process) were added to the list.

Gremler (2004:73) found that most critical incident technique studies include positive and negative critical incidents. In order to ensure that the full service offering was evaluated, the critical incident technique instrument used in the present research therefore included a specific request to report both positive and negative experiences. When a service quality model is used to measure service quality (ideally not only once but frequently, at regular intervals), if only negative critical incidents are asked for, such a model may only indicate the service quality problem areas current at the time when the model is developed – it would not reflect the full spectrum of important services offered by the service provider.

In the four questions used in the present research, the tax practitioners responding to the survey were requested to list the things they "appreciate" (Questions 1 and 3) and then the things they "dislike" (Questions 2 and 4) about their interactions with SARS. Johnson and Gustafsson (2000:158) found that the number of things respondents "dislike" usually exceeds the number of things respondents "appreciate", so the tax practitioners were guided to first list the things they "appreciate" (Questions 1 and 3).

In the group interview, the respondents were also provided with a range for the number of responses that should be provided. Johnson and Gustafsson (2000:158) recommend a range of between five and ten responses for each category. Because there are different service channels likely to elicit different perceptions of service quality levels, it was thought that a range of five responses would perhaps limit the number of responses, as there are already approximately six different service channels. For the purposes of the present research, the respondents were encouraged to list as many experiences as possible, but a guideline range of ten experiences (for both the positive and the negative answers) was provided for the group interview respondents. For the web-based questionnaire, an open block without any range was provided, but respondents were still encouraged to list as many experiences as possible.

4.8 DATA COLLECTION PROCESS

The group interview was conducted at the University of Pretoria on Monday, 12 November 2007, at 16:00, with members of the Tax Committee of the Northern Region of the South African Institute of Chartered Accountants (SAICA). All the SAICA members in the Northern Region who practise as tax practitioners were invited to attend this meeting. It was decided to use this group for the group interview because it was the most convenient

option for both the researcher and the respondents, given that all the parties live and work in Pretoria, Gauteng.

The sample for a qualitative study does not need to be truly random, as no statistical validity is required, and most studies do not report using a probability sample (Gremler 2004:72; Johnson & Gustafsson 2000:52). A convenience sample of tax practitioners is thus acceptable for the purposes of a group interview.

In this case, a total of 22 members of SAICA actually attended the meeting. The researcher addressed the group with an introductory presentation. This presentation included all the information that is usually provided in a letter of consent, for example, the scope and purpose of the study, the importance of the study and the confidentiality of individual responses (see Annexure A for a copy of the letter of consent that was also given to the group interview members).

After the introduction, a hard copy of a questionnaire containing the four questions was handed out to all the attendees. This questionnaire was available in both Afrikaans and English and each attendee therefore received it in his or her language of choice. The translation of the questionnaire was done by the researcher. The utmost care was taken to ensure that both languages communicated the same message to the respondent (see Annexure B for a copy of the English questionnaire and Annexure C for a copy of the Afrikaans questionnaire).

Initially, the intention was that the attendees at the meeting were to complete the questionnaire at the meeting, but after discussing the matter, some of these tax practitioners felt that they would prefer to consult with their personnel before completing the questionnaire. There were two reasons for this: firstly, some practitioners did not have to deal directly with SARS officials themselves and they therefore felt they had to consult with the relevant personnel at their offices who deal with SARS on a day-to-day basis; secondly, although some of these tax practitioners did deal with SARS on a daily basis, they wanted the results to incorporate all the positive and negative experiences of their employees and therefore wanted to consult with them, to add, where possible, to their own responses. In order to assist them to complete the questionnaire electronically, the group interview questionnaire was distributed to all 22 attendees by e-mail on 16 November 2007. Six completed questionnaires were returned, which represents a

response rate of 27.3%. The data generated by this group were kept separate, as they could be classified as a questionnaire option with the benefit of a formal introduction, as well as initial contact with the researcher. (In further references to the responses to this questionnaire, this option is referred to as the "distributed questionnaire").

The collection of the data by means of the distributed questionnaire was followed by a web-based questionnaire included as part of a larger survey administered by SARS. For the purposes of the present research, this questionnaire is referred to as the "web-based questionnaire". The distributed questionnaire was also used to validate the content of the questions relating to critical incidents to be included in the final web-based questionnaire. The distributed questionnaire was found to be suitable for the purposes of the present research, but the respondents to the distributed questionnaire referred to a business process (status queries) that was not originally added as a business process. The status query business process was therefore added to the web-based questionnaire administered by SARS. The group interview respondents also referred to text messaging as a service channel at SARS. The text messaging service channel was therefore also added as a service channel for the web-based questionnaire administered by SARS. The four open-ended critical incident questions were included as part of a bigger questionnaire administered by SARS. The purpose of the SARS questionnaire was to gather information so that SARS could enhance its service offerings and the four-open ended questions formed part of it. Notice of the questionnaire, as well as the link to the website, was e-mailed to the full data base of tax practitioners registered with SARS. The questionnaire was open for completion from 21 November 2007, with a closing date of 26 November 2007.

The total population of approximately 17 000 tax practitioners returned 811 completed questionnaires, which represents a response rate of approximately 5%. This response rate may be considered satisfactory, given that the average response rate for questionnaires in marketing-related studies is often as low as 5% (McDaniel & Gates 1996 in Odekerken-Schröder *et al.* 2000:110). The response rate is also not a matter for concern, because the purpose of open-ended questions is to assist in an exploratory study. It should also be taken into account that a large, unknown number of questionnaires may not have reached

the relevant tax practitioners. Snyckers (2007b) suggested the following reasons:

- Because the questionnaire was mailed as a bulk e-mail, some e-mail system operators may have identified it as “spam” and it was therefore either not sent to the recipient or it was marked as “spam” on the mail delivered to the recipient, which may have resulted in immediate deletion by the recipient.
- Some e-mail boxes may have been full and the questionnaire was therefore not delivered.
- The e-mail addresses for some of the tax practitioners on the database may have been incorrect.
- Some tax practitioners did not provide an e-mail address when they registered. This may be either because they do not have an e-mail address, or because they did not want to make their e-mail address available.

Apart from the response rate (which may be considered satisfactory), the number of responses should also be considered, because, in 115 of the critical incident technique studies evaluated by Gremler (2004:73), the response rates were not reported, and it is therefore common for critical incident technique studies to report only on the number of responses and the number of critical incidents. Gremler (2004:73) found that the average number of responses in the 115 critical incident technique studies investigated was 341. In the present research, the 811 responses received may therefore be considered to be adequate.

4.9 DATA ANALYSIS

The data analysis involved three processes. The first was the identification of usable critical incidents (see Section 4.9.1), the second was the development of a classification scheme for the content analysis (see Section 4.9.2) and the third was a content analysis of the identified critical incidents (see Section 4.9.3).

4.9.1 Defining and identifying a critical incident

According to Hays (in Johnson & Gustafsson 2000:52), an expert in the critical incident approach,

> [a] critical incident is a specific example of the service … that describes either positive or negative performance. A positive example is a characteristic of the service … that the customer would like to see every time he or she receives that service. A negative example is a characteristic of the service that would make the customer question the quality of the company.

Elements in comments by Bitner *et al*. (1990:73) and Flanagan (1954:12) can be added to the above definition. They believe that an incident can only be critical if it makes a "significant" contribution, either positively or negatively, to the general aim of the activity.

For the purposes of the present research, critical incidents are defined as positive and negative service experiences by tax practitioners of the services provided by SARS. To try to ensure that only experiences relating to "critical incidents" were gathered, the words "really appreciate" and "really dislike" were used in the wording of the questions put to the tax practitioners.

The respondents to both the group interview (distributed questionnaire) and the web-based questionnaires were asked to provide a list of the things that they appreciated and disliked about SARS services. The analysis procedure advocated by Flanagan (1954) indicates that the critical incident itself is the basic unit of analysis. Hence, for the purposes of the present research, the basic unit of analysis (the critical incident) is defined in such a manner as to include statements about SARS service delivery. These statements had to be as specific as possible. A critical incident therefore did not include any comment relating to

- the specific tax legislation applicable;
- the fairness of the tax system;
- the interpretation of a particular piece of legislation;
- services rendered by the SSMO;
- the value received as a *quid pro quo* for taxes paid;
- any comments on Revenue Stamps; or
- Customs and Excise duties.

Any statement that was too vague was also excluded from the list of comments on the critical incidents identified.

To ensure that the right type of data was used in the analysis, it was important that each response was measured against the general definition of a critical incident for the purposes of the present research. Only critical incidents as defined were then used in the data analysis. From the six completed distributed questionnaires, a total of 164 critical incidents were identified and analysed. From the 811 responses to the web-based questionnaire, 5 252 critical incidents were identified and analysed.

In evaluating the adequacy of the number of critical incidents, it should be remembered that services are heterogeneous. Because it is impossible to measure them against exact, uniform standards, even when each customer receives exactly the same quality of service, each customer could evaluate these services differently (Anthony & Govindarajan 2000:621; Czepiel *et al.* 1985:3; Gaster & Squires 2003:7; Haywood-Farmer 1988:20). In order to ensure that the results of the present research truly reflected the perceptions of tax practitioners (and indirectly the perceptions of the taxpayers they represent), the number of critical incidents had to be high enough. The question of whether the actual number of critical incidents was in fact high enough does not appear to have a simple answer, but, according to Flanagan (1954:18), for most purposes, the number of critical incidents could be considered to be adequate when the addition of 100 critical incidents to the sample adds only two or three additional determinants. Gremler (2004:73) reported that the average number of usable critical incidents across 115 critical incident technique studies was 443.

In the present research, many more critical incidents than the average number of 443 usable critical incidents were identified. It was found that fewer than three additional determinants were added with the addition of 100 critical incidents. In fact, this result was already achieved when more or less 50% of the critical incidents were analysed. It can therefore be concluded that the number of critical incidents identified was indeed high enough to draw a relevant conclusion.

4.9.2 Development of a classification scheme for the purposes of the data analysis

After the data has been collected and the relevant critical incidents have been identified, the next step in the critical incident technique is to analyse the data. The first step in the

data analysis in the present research was to develop a classification scheme for the purposes of the content analysis. The aim of a content analysis is to summarize and describe data in an efficient manner, so that it can be used effectively for many practical purposes (Flanagan 1954:19). The main categories of classification (the classification scheme) can either be deduced from theoretical models or established on the basis of inductive interpretation (Gremler 2004:66).

As a starting point, the previous studies that specifically focused on the quality of the services SARS provides (Department of Taxation 2005; De Clercq *et al*. 2006; SARS 2005a:35; SARS Practitioners Unit 2007; Smulders 2006; Tustin *et al*. 2006) were investigated. Most of these studies used closed-ended questions chosen by the researcher concerned or the researcher in conjunction with SARS and other bodies. These questions could therefore not be used to indicate what should be included in a model of evaluating service quality as viewed through the “lens of the customer”.

The next possible option was the theoretical models derived from the literature study. It was found that the service quality models (for example, SERVQUAL) tend to be fairly generic. This implies that it is important to identify the service quality determinants for each particular type of service or service industry, because standard instruments are unsuitable to measure the service quality in different industries (Babakus & Boller 1992:253,264; Badri, Abdulla & Al-Madani 2005:842; Barnes & Glynn 1993:51; Carman 1990; Haywood-Farmer 1988; Phillip & Hazlett 1997:272; Schneider & White 2004:33). The measurement of service quality should therefore be conducted using instruments which have been developed by identifying the determinants of service quality that are important to the customers of the company whose service quality is being measured (Barnes & Glynn 1993:51). The existing models could serve as a framework or basis to be modified and changed to fit the needs of specific contexts (Schneider & White 2004:38). In the present research, the existing service quality models were used as a basis to develop a classification scheme to assist in identifying the determinants that are important in evaluating the service quality of services provided by SARS.

As a starting point, the original ten service quality determinants from the study by Parasuraman *et al.* (1985) were used. All ten determinants taken from Parasuraman *et al.*’s (1985) study were then listed in a classification scheme. Next, this classification

scheme was expanded using the other service quality instruments investigated in the literature review.

Because Kang and James (2004) and Philip and Stewart (1999) found that the SERVQUAL dimensions do not measure the technical quality of a service, but only its functional quality, all the different business processes were also added to the classification scheme.

Santos (2003), Zeithaml *et al.* (2002) and Zhu *et al.* (2002) found that e-service quality is influenced by determinants that differ from traditional service quality, so the service channels through the website, as well as e-filing, were listed separately in the classification scheme. The determinants identified in the models analysed in the literature review were also summarised separately according to these service channels. For the purposes of the present research, e-services were regarded as all services provided via the internet or SARS's e-filing.

The literature study indicated that models of service quality are equally applicable to both the private and the public sectors, but to check whether this was really the case, specific aspects were included in the classification scheme that might only be relevant to SARS as part of the public sector. Market forces in the private sector should, for example, ensure that the accessibility of the service provider can be optimised. Although the accessibility determinant was excluded from Parasuraman *et al*.'s (1991a) study, it was included in the classification scheme in the current study, as it might be important for the public sector. The same may be true for communication – in the public sector there are no market forces to automatically provide the optimum situation. In the private sector, contact personnel are, for example, mainly appointed because of their communication skills, or they receive training in these skills if these skills are relevant to their job descriptions. SARS has had to reorganise, and many of the employees who previously only captured information have now been redeployed (moved to a different department) in SARS. This could imply that many people may have been moved into departments for which they do not necessarily possess the required skills. Communication was therefore also included in the classification scheme.

The first draft of the classification scheme distinguished between traditional services and e-services. Hence, all the different determinants already identified in existing models were included. Based on the experience of the researcher, additional determinants were added.

Flanagan (1954:20) argues that the classification scheme is usually developed or confirmed by starting with a relatively small sample of critical incidents and sorting them into clusters or “piles” related to the frame of reference that has been selected. Hayes (in Johnson & Gustafsson 2000:56) recommends a sample of between ten and twenty as a starting point. The number may increase, depending on the rest of the analysis.

In the present research, the 164 responses from the distributed questionnaire were analysed to refine the classification scheme. The classification scheme was adjusted to incorporate all of these responses. In applying the classification scheme to the bulk of the data (the critical incidents from the web-based questionnaire), the classification scheme was amended in a constant process which resulted either in the expansion of the definitions of current categories or in the addition of new categories.

Flanagan (1954:20) argues that the development of the classification scheme should be a continuous process, but does not highlight any risk with regard to the classification of incidents before changes to the classification model are made. Although the initial classification scheme in the present research was not materially adjusted throughout the process of the content analysis, two risks still remained: that the data analysed in the beginning were not measured against the same definitions and categories as the ones analysed at the end of the process and that because of a lack of richer definitions or additional specific categories, the category closest to the response was chosen by the data analysts (refer to Section 4.9.3 below for a discussion of the involvement of analysts).

One way to empirically test (or pre-test) a classification scheme is to use a holdout sample (Gremler 2004:82; Johnson & Gustafsson 2000:60). Such a practice entails setting aside a portion of the incidents and using only the other incidents to develop the categories (Gremler 2004:82; Johnson & Gustafsson 2000:60). In the current study, all the critical incidents were grouped into 35 different groups. Group 35 represented the critical incidents obtained from the group interview respondents. All the critical incidents from the web-based questionnaire were evenly distributed into Groups 1 to 34. Groups 33 and 34 were treated as a hold-out sample, but, because they added nothing new to the

classification scheme, it was concluded that the categories of the classification scheme were comprehensive.

4.9.3 Description of the content analysis process

After a classification scheme has been developed, content analysis of the data is done. Generally the goal of a content analysis is to develop a classification system to provide insights regarding the frequency and patterns of factors that affect the phenomenon of interest (Flanagan 1954:29; Gremler 2004:66). The aim is to increase the usefulness of the data, while sacrificing as little as possible of their comprehensiveness, specificity and validity (Flanagan 1954:19). It is usually not possible to obtain 100% objectivity in this coding procedure.

The first step in the coding procedure is to have independent judges sort the critical incidents into their underlying categories (Johnson & Gustafsson 2000:57). Generally, two to three judges (coders) are used to analyse and ultimately categorise the critical incidents (Gremler 2004:73). Usually the judges categorise incidents without prior knowledge of other judges' coding (Gremler 2004:73).

In the present research, a total of nine judges (the researcher and eight research assistants) performed the content analysis of the critical incidents. Six of the research assistants had recently completed their Honours degrees in Marketing Management. This degree includes a research project and a research methodology course (the content of the course was evaluated by the researcher and it was found to be sufficiently extensive to equip these assistants with the necessary background to be able to assist in analysing the critical incidents). One research assistant had recently completed an Honours degree in Human Resources Management. This course also includes a research methodology course, the content of which was also found to be extensive enough to equip this assistant with the necessary background to be able to assist in analysing the critical incidents. Another research assistant was busy with her academic traineeship toward becoming a chartered accountant. She had completed the first year of her course-based master's degree in Taxation and Part 1 of the qualifying examination for admission as a chartered accountant. Although this assistant had not as yet completed a course on research methodology, she assisted the researcher with some language editing on the first three chapters of the present research. Taxation was one of her major subjects from her second

year onwards. She was able to bring a theoretical background on the research topic to the research process. This academic candidate did not initially form part of the group of research assistants, but was only added later, as described below.

To ensure that sufficient training was provided to the analysts, they were all given the first three chapters of the present research to read. They were also provided with Chapter 3, "Building the lens of the customer", in the book *Improving customer satisfaction, loyalty and profit – An integrated measurement and management system* by Johnson and Gustafsson (2000). After they had worked through these documents, the researcher met with them to clear up any questions or uncertainties. The researcher also explained the purpose of the study and the methods used to gather the data on the critical incidents. The research assistants each received a copy of the classification scheme (refer to Annexure D for the classification scheme for the traditional services and Annexure E for the classification scheme for the e-services). The researcher worked through all the definitions with them. Without using any of the responses actually received, *ad hoc* examples were provided and the analysts were each given an opportunity to classify the examples into the classification scheme.

After their training had been completed, the research assistants (excluding the academic trainee candidate) received the 164 responses from the distributed questionnaire (Group 35) and had to perform the content analysis of these responses independently. After they had completed their own work, the whole group of assistants compared their analyses. As a group they agreed on one final classification. The researcher then met with the group to compare her own content analysis with the analyses performed by the group. It was found that there was only a 52% agreement between the researcher and the group. This was considered to be too low. When the analyses were discussed, it was found that the research assistants did not fully understand either the tax principles or the technical terminology used. They also did not understand how SARS operates. The researcher explained to them how the service channels and business processes at SARS work. They then also familiarised themselves with the terminology used. Each assistant was provided with two chapters from different tax textbooks to assist them with additional background knowledge on taxation and SARS. It was also decided to add the academic trainee candidate to the group. Although she later assisted with the content analysis, her initial

function was only to answer questions on tax aspects not fully understood by the research assistants.

After the additional training, each research assistant had to reclassify the responses from the distributed questionnaire. Again the group compared their analyses. The final version was then compared with the researcher's results. A level of agreement of more than 90% was achieved in the second round of analysis, with the added advantage that the differences were no longer one-sided. In relation to the remaining 10% differences between the results of the group and the researcher, nearly half were resolved by accepting the analysis of the group, as their individual academic backgrounds brought different understandings to the various critical incidents. After their understandings were explained to the researcher, the researcher adopted the classification of the group. The remaining differences were mainly the result of an incomplete understanding of the context or technicalities. However, because that resulted in a difference of only 5%, the research assistants were deemed to be sufficiently equipped to classify the responses to the web-based questionnaires.

The responses to the web-based questionnaires were organised into 34 groups (with no criterion other than the order in which the responses were received). Each critical incident in each group was then independently classified by two different research assistants. As they proceeded, interaction with the researcher resulted in an expansion of current definitions and the addition of new categories to the classification scheme. Any changes to the classification scheme were always communicated to the whole group. After the independent analysts had completed their separate evaluations, they compared their analyses and provided a final analysis agreed upon by both. If they could not reach agreement, both their versions were provided for on the final list. Any critical incidents which they did not understand or where they were not completely sure of whether the classification was correct were also identified. Care was taken to ensure that each research assistant always had to work with a different person – a schedule was drawn up for this purpose. To ensure that they worked independently, the research assistants working on the same groups were not allowed to sit next to each other.

After the content analyses of all the groups had been completed, the classification scheme was deemed to be final. A third research assistant (different from the ones who had already worked on each specific group) was then allocated to each group and that

assistant then had to reclassify the whole group independently. The three research assistants that worked on the group then discussed the differences between the final list of Research Assistants One and Two and the results obtained by Research Assistant Three. The changes that were then effected were mostly because of the changes to the classification scheme or because of the additional experience all the candidates had gained by then.

To ensure that the researcher, who is in an authoritative position compared to the other analysts, did not dominate any changes to the classifications, the other three analysts first reached an agreement on all their differences. Only after their differences had been resolved did the three analysts and the researcher meet to resolve the remaining differences. All the items on which they still did not agree, or where they were uncertain, were dealt with first, and then agreement was reached for all the critical incidents between the four parties (the researcher and the three research assistants).

The researcher carried out spot checks on the analysis of the group results, depending on the amount of agreement between the researcher and the research assistants. A minimum of 80% of the critical incidents of each group were reviewed by the researcher. For some groups, all were reviewed. The changes made by the researcher (in agreement with the group) were, however, never more than 20% in total for any particular group. This implies an interjudge reliability of at least 80% for all the groups. Although it is acknowledged that the position and age of the researcher might have intimidated the research assistants, the position of the researcher had the added benefit that it also resulted in a situation in which the research assistants wanted to be regarded as “as successful” as possible. Because it was important for them to be “successful” in their classification, they tried to identify all aspects which they felt even remotely unsure about. This resulted in a situation where they would not agree with each other unless they were totally convinced. It also meant that they communicated their views strongly to the researcher in an attempt to convince the researcher of those views. They were frequently successful in convincing the researcher of their views (something they enjoyed greatly).

The critical incidents from the distributed questionnaire (which was originally used to develop the classification scheme) were identified as Group 35. After the content analyses of Groups 1 to 34 had been finalised, two research assistants were identified to reclassify Group 35. The researcher then reviewed all these results.

4.10 RELIABILITY OF THE DATA ANALYSIS

Gremler (2004:75) indicates that reliability is a key component in content analysis methods. A variety of interjudge reliability indices is available in evaluating the reliability of critical incident technique data analysis. Clearly the most common reliability index used is interjudge agreement (the total number of agreements divided by the total number of coding decisions) (Gremler 2004:74; Johnson & Gustafsson 2000:59; Perreault & Leigh 1989:137).

The utility of an estimate of interjudge reliability is not necessarily limited to an *ex post facto* evaluation of coded data (Perreault & Leigh 1989:137). Instead, researchers often need a diagnostic application (in pre-tests or on a subset of data early in the coding process) to determine whether the classification scheme, definitions, directions and training (that of the coders) can be improved (Perreault & Leigh 1989:137). Once an adequate level of reliability is achieved, attention can be given to the general implementation of the coding process (which may or may not involve multiple judges to code every observation) (Perreault & Leigh 1989:137).

According to Johnson and Gustafsson (2000:59) and Perreault and Leigh (1989), an agreement index of 80% is a reasonable cut-off level to determine whether content analyses are reliable. This percentage agreement has been shown to underestimate interjudge reliability when there are a large number of categories (Perreault & Leigh 1989).

Although no formal indices are available for the reliability of the interjudge classifications, it is reasonable to assume that the thoroughness of the process, as well as the individual agreement ratios in excess of 80% for all the groups, should indicate that the results of the content analysis were reliable. The initial training of the research assistants and the pre-tests on the subset of data (Group 35) that were done early in the coding process also contributed to the reliability of the results. It can also be assumed that the result of 80% is underestimated, as there were a very large number of categories in the classification scheme.

The raw data were organised per respondent. As the critical incidents were numbered per incident and not per respondent, the same points mentioned by the same person carry the same weight as the same point mentioned by more than one person. To try to eliminate any double inclusion, all research assistants were requested to ensure that the same

respondent did not have two or more critical incidents for the same category. Duplications of the same principle by respondents did, in fact, occur very often and this then meant that the duplicated principle was not counted as a critical incident. The researcher also ensured that all data classifications per respondent were reviewed, and not only for specific critical incidents marked. It was found in nearly all of the cases that the research assistant had ignored the duplicated items, and that these items had therefore not been counted more than once.

Although it is acknowledged that a tax practitioner might wish to emphasize a particular issue by mentioning it more than once, it was felt that double inclusion may not reflect the most accurate picture, as other respondents may feel even more strongly about a specific item, but may only have mentioned it once. It should also be noted that the questionnaires were completed electronically, which also allowed respondents to use the “cut” and “paste” options to repeat themselves without much effort. It was assumed that the researcher’s review of between 80% and 100% of the classifications of the critical incidents was sufficient to conclude that the risk of double inclusion of the same critical incident by the same respondent was eliminated as far as possible.

It is acknowledged that, although the participants in the group interview had been requested not to complete the open-ended questions of the questionnaire administered by SARS, there is no way of determining whether any duplication did in fact take place, but no repetition was detected by the researcher.

After a careful evaluation of the process followed for the content analysis, the results were considered to be reliable and the results could therefore be reported as they stand.

4.11 INTERPRETATION OF THE DATA

After the preparation of the summaries of the frequencies of the responses in terms of the relevant classification scheme, a process of **natural language argument** was used to convert the data analysis results and the relevant elements of the theory from the literature survey into the two parts (traditional and the e-service quality) of the proposed model as developed in Chapters 5 and 6 of the present research. Chapter 5 discusses the process for the development of the part of the proposed service quality model for the traditional services SARS renders. Chapter 6 discusses the process for the development of the part

of the proposed service quality model for the e-services SARS provides. Both the traditional and the e-service quality parts of the model are presented in the final chapter of the present research (Chapter 7).

Some general issues were also relevant to the interpretation of the data and these issues are set out below.

4.11.1 SARS Service Charter

In developing the model that is to be used to assess service quality, SARS's Service Charter (SARS 2005b:1), officially released to the public on 19 October 2005, was relevant to the extent that it provides SARS's perspective of the "lens of the customer". The Service Charter commits SARS to clearly defined deliverables that were to be implemented by 2007 (SARS 2005b:1). However, it was not very clear from the Service Charter exactly when in 2007 the Charter was to be fully operational. Croome (2005) and Kieswetter (2006b) were under the impression that the terms of the Charter only applied from the 2008 tax year (that is, from 1 March 2007). There was also no official announcement of a commencement date for the SARS Service Charter. Hence, it was assumed that the service standards set out in the SARS Service Charter were already applicable for the purposes of the service quality model to be developed to measure the quality of SARS's services. These service standards were therefore also incorporated into the results of the study when applicable.

It is important to note that currently the service standards indicated in the SARS Service Charter are applied, but it is not clear how these standards were established. Although a comparison with the standards may be helpful in evaluating whether SARS adheres to its promises, meeting these standards could not be used as an indicator of service quality in the present research, as the "customer" (in this case, the participating tax practitioners) was not consulted in the drafting of the service standards in the Charter. It was therefore still very important that the expectations of the tax practitioners should also be measured.

4.11.2 Importance of responses

It should be remembered that the analysis set out in the present research represents the results of a qualitative study and that the respondents were asked to comment on aspects that they either highly appreciated or really disliked with regard to SARS's services. All the

aspects mentioned by the respondents therefore formed a starting point that was relevant to the development of the service quality model, except where the researcher specified why they should be excluded.

Moreover, the frequencies of the same aspect as mentioned by more than one respondent were used to interpret the relative importance of that particular aspect in the context of the present research. If a particular aspect was mentioned by 100 or more different respondents, the aspect was regarded as more important than an aspect that was listed only once or twice. Apart from the fact that the relative importance of the various service determinants can assist SARS to focus its service strategies on the most relevant service aspects, the importance ratings can also assist in the refinement of a service quality model, especially if the length of the model becomes a problem. The benefits of having specific information versus the risk of not having any information at all meant that a balance had to be found, and in this regard, frequencies were relevant.

4.11.3 Service quality versus building the lens of the customer

Although the results of the content analyses (as presented in Chapters 5 and 6) also reflected the perceptions of the participating tax practitioners with regard to the quality of SARS's services, the purpose of the results of the content analysis was primarily to assist in identifying the determinants that are most important to the tax practitioners in order to draft the "lens of the customer". Chapters 5 and 6 therefore focus on presenting the results for this purpose and should not be interpreted as any reflection on the quality of the services SARS delivers.

4.11.4 Validation of the proposed service quality model

The validity of the proposed model was explored by comparing it to two existing service quality models. The traditional service quality part of the model proposed was compared with the SERVQUAL model, while the e-service quality part of the model proposed was compared with E-S-Qual. The outcome of the comparison is presented in the final chapter of the present research (Chapter 7).

4.12 CONCLUSION

This chapter discussed the research design and the process followed in analysing the

responses for the qualitative research performed. The critical incident technique was chosen as the qualitative method used to build the “lens of the customer” for the evaluation of the service quality of SARS by the practitioners. The purpose of using the critical incident technique in the present research was to present a service quality framework to assist in the development of a quantitative survey instrument to measure the service quality of SARS.

The chapter described the data collection method and process, the design of the data collection instrument, the definition of a critical incident for the purposes of the present research, the processes followed in developing the classification scheme and the content analysis process that followed. The reliability of the research process followed was also evaluated and it was considered that the data analysis method used was reliable and that the results could be reported as they stand. Finally, the chapter discussed the interpretation of the data. In the next chapter, the outcome of the data interpretation (to the extent that it relates to the traditional services) is presented.