THE UNIVERSITY OF CALGARY

AN APPLICATION OF GENERALIZED SERVICE QUALITY MEASUREMENT SCALES

TO SELECTED SEGMENTS OF THE TOURISM INDUSTRY

BY

GAVIN ROSS FICK

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF

MASTER OF BUSINESS ADMINISTRATION
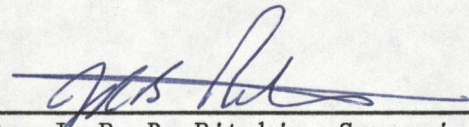

FACULTY OF MANAGEMENT
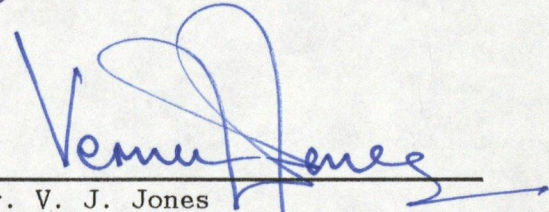

CALGARY, ALBERTA

SEPTEMBER, 1989

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "An Application of Generalized Service Quality Measurement Scales to Selected Segments of the Tourism Industry" submitted by Gavin Ross Fick in partial fulfillment of the requirements for the degree of Master of Business Administration.
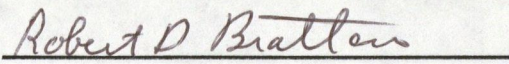
Dr. J. R. B. Ritchie, Supervisor
Faculty of Management

Dr. V. J. Jones
Faculty of Management

Dr. R. D. Bratton
Faculty of Physical Education

September, 1989

ii

ABSTRACT

This study represents an attempt to evaluate the effectiveness of a generalized service quality measurement scale within segments of the tourism industry. It involves an application of the scale to one of the service segments used in its development and four tourism service segments. Additional examination of the scale after the effectiveness evaluation attempt to examine the functioning of the scale to a greater extent.

The results of the study indicate the scale does not have the generalization abilities that have been professed, and that it is not as effective as it might be for measuring service quality within the tourism industry. Further examination of the functioning of the scale indicated the scale has significant problems with using negatively worded statement sets, its calculation of the quality score, and the relationship between its quality score and actual perceived quality. While it is not without its flaws, however, the scale is a good initial attempt at measuring perceived service quality and further recommended research should build on its foundation.

ACKNOWLEDGEMENTS

A thesis is the culmination of a single person's determination and judgement, but is not independent of the assistance of others. This thesis is no exception, and I will try to acknowledge some of the main contributors to its completion.

To several of the instructors at the University of Calgary, Southern Alberta Institute of Technology, Mount Royal College, and Alberta Vocational Centre, I express deep gratitude for allowing me access to their classes for sampling.

I would like to express my sincere appreciation to Dr. J. R. Brent Ritchie, for his instruction and guidance through the course of this research study and my tenure in the Graduate Program. I have great respect for Brent and hope that our academic relationship and personal friendship continues into the future.

Any potential success of an undertaking the size of this project is greatly enhanced by the existence of a strong support group. My

father and mother, Ross and Myra, my brother Colin, and my wife Patti have provided such a group. They have provided an environment conducive to work when appropriate, and to recreation when necessary. It is to my family that this document is dedicated.

TABLE OF CONTENTS

Chapter 6:   CONCLUDING DISCUSSION

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.0 SERVICE QUALITY

There are many elements that affect the quality of travel and tourism services. The physical facilities, service personnel, and delivery time are all factors that can influence consumers' perceptions of the quality of service they have received. These perceptions can play a vital role in the word of mouth communication that has been identified as being an important factor in the success of service industries (Eiglier et al., 1977). Studies have also shown that upon finding service organizations that provide a satisfactory level of performance, consumers often restrict their search activities for that type of service organization in the future and limit their choice to the organization with which they have had a positive experience (Booms and Bitner, 1981).

The concept of quality is an elusive construct; one that is difficult to delimit and measure (Parasuraman et al., 1984). When asked, consumers often have a difficult time defining quality and its requirements, yet they are consistently demanding increasing levels of quality in products (Takeuchi and Quelch, 1983). From an organization's point of view, quality levels have been demonstrated to have a positive impact on market share (Buzzell and Weirsema, 1981), on return on investment (Phillips et al., 1983), and serve to increase productivity and lower manufacturing costs (Garvin, 1983; Phillips et al., 1983). The emphasis on quality was identified as an emerging marketing trend in the 1980s (Rabin, 1983).

A primary goal of any study of service quality should be to provide assistance to participants in service industries in their

management of service quality. In order to effectively use the concept of service quality to improve product offerings, techniques must be available to facilitate the measurement of such quality. Measurement is an important step in devising action plans to meet a service organization's goals (Lewis and Booms, 1982).

Parasuraman et al. (1986) have developed a multiple item scale (SERVQUAL) for measuring consumers' perceptions of service quality. The researchers identified five dimensions of service quality which attempt to capture the criteria used by consumers in assessing the quality of services. One of the main purposes of the research was to produce a reasonably universal scale that would be applicable to a wide range of services. The categories of services that were sampled in developing the scale included:

1) Appliance Repair and Maintenance

2) Retail Banking

3) Long Distance Telephone

4) Securities Brokerage

5) Credit Card Organization

These categories were chosen to represent a broad cross-section of service types that vary along key dimensions identified by Lovelock (1980; 1983). For those service categories tested, the scale developers contend that the SERVQUAL scale is a concise instrument with good reliability and validity. They further believe that it can be used to assess the quality of service in a wide range of service categories. If this is accurate, service organizations now have at

their disposal a powerful tool in their endeavors to manage service quality.

1.1 TOURISM SERVICE QUALITY

Research into the attributes employed by consumers when evaluating the quality of service offered by components of the tourism industry is lacking. Several studies attempted to extract specific attributes that consumers use in choosing and evaluating service organizations within the tourism industry (O'Brien et al., 1977; Ritchie et al., 1980; Pizam et al., 1978; Etherington and Var, 1984). However, these attributes are very specific to industry segments and no attempt has been made to relate them to other components within the tourism industry or to other types of service industries.

Other attempts have been made to evaluate the importance of different factors in service. Two of these studies employed researcher-selected attributes in analyzing the consumers' perceptions (Perry and Friedman, 1973; Makens and Marquardt, 1977). However, O'Brien et al. (1977) illustrate the importance of having the consumers of the service determine the attributes used in evaluating satisfaction and attach an importance factor to each. This method of using consumer generated attributes was applied by Parasuraman et al. (1986) in the SERVQUAL research.

O'Brien et al. (1977) identified eight factors considered by consumers in choosing airline companies. A discussion of the specific

factors of this study, and others that will be mentioned, will follow in the literature review section. Makens and Marquardt (1977) evaluated the importance of fifteen different factors in airline travel, and illustrate differences in the evaluation of the importance of these factors between first class and coach travellers. Similarly, Ritchie et al. (1980) discuss eight dimensions along which consumers base their decisions on airline preference.

Pizam et al. (1978) and Etherington and Var (1984) identify factors that are considered by consumers in choosing or evaluating a member in a specific segment in the tourism industry. While similarities exist between factors or dimensions identified by Parasuraman et al. (1986) and those discussed by O'Brien et al. (1977), Makens and Marquardt (1977), Ritchie et al. (1980), Pizam et al. (1978), and Etherington and Var (1984), there are also many significant differences and inconsistencies. An important result of the evaluation of these differences could be a determination that they are the result of some of the unique attributes of tourism services.

As Burkart (1981) states: "Of course tourism is a service industry. Its output is a particular kind of service rather than a tangible physical good, which characterizes the manufacturing industries." (p. 2). The tourism industry shares the special characteristics of service industries that have been identified in the literature and will be discussed in the literature review. What must be determined is the extent to which the tourism industry shares the quality evaluation characteristics of those industries.

1.2 BASIS FOR THE RESEARCH

The proposed research is based on the necessity of determining if generalized service quality measurement scales may be effectively applied to the tourism industry. Testing existing scales within segments of the industry will help assess the degree to which the quality of tourism service may be measured by the generalized scale. Examination of individual dimensions may suggest additions, deletions, and modifications of items within the scales that serve to aid the instrument in accurately representing the perceptions of service quality within the tourism industry.

While attempts have been made to develop items and scales that may be used within segments of the tourism industry (Perry and Friedman, 1973; Makens and Marquardt, 1977; Ritchie et al., 1980; O'Brien et al., 1977; Pizam et al., 1978; Etherington and Var, 1984), these attempts had some weaknesses in their approaches. They have either been limited in the range of service categories investigated, employed researcher-defined criteria which may not correspond to the perceptual structure of consumers or have been developed without the rigor, attention to detail, examination of reliability and validity, or generalizability that are characteristic of good scale construction procedures.

Proper construction procedures and evidence of the scale's 'goodness' have been identified as being required for good scale construction (Churchill, 1979). This research will attempt to test Parasuraman et al's (1986) SERVQUAL scale in the tourism industry, and determine if it may be used to measure and evaluate the quality of

service offered in four different tourism service categories. Where the scale is found to be inadequate, recommendations will be made to enhance its application to the tourism industry.

CHAPTER 2

LITERATURE REVIEW

2.0 HOW SERVICES ARE DIFFERENT THAN PRODUCTS

Service industries have only recently begun to attract attention as a separate entity from traditional goods producing industries. Much of the work that has been done in the area of product research has been focused on the production and marketing of goods. With the growing importance of service industries, this same type of research is becoming necessary to support those organizations providing services to the marketplace. The lack of concentration of research in the area of goods production and marketing would cause little concern if the problems faced by those in service production and marketing were the same or similar. Unfortunately, due to some unique characteristics of services, they are not.

The rationale for the separate treatment of services relies on the importance placed on a number of characteristics of services consistently cited in the literature. These characteristics are:

1)   Intangibility

2)   Inseparability of production and consumption

3)   Heterogeneity

4)   Perishability

The fundamental difference between goods and services, recognized by many researchers (Rathmell, 1979; Shostack, 1977; Berry, 1980; Lovelock, 1981), is intangibility. Berry (1980) states: "A good is an object, a device, a thing; a service is a deed, a performance, an effort."(p. 24). When a consumer purchases a good, he has something

tangible to show for the transaction. When a service is purchased, there is generally nothing tangible in nature to show for it.

Shostack (1977) points out, however, that goods and services are actually combinations of tangible and intangible elements, and the essence of the purchase determines its classification as a good or service. For example, airline transportation involves many tangible elements: the check-in and baggage areas, the airplane, the interior decor, and the food and beverages served on the flight. The essence of the product, however, is transportation. This transportation is itself intangible; hence the classification of airlines as belonging in the service industry. The company that builds the airplanes and sells them to the airline has the physical airplane as the essence of the transaction, and is classified as belonging in the goods or manufacturing industry.

Other unique attributes of services are actually functions of the intangibility characteristic. While they are discussed separately and individually, it is important to note that their origin is the result of intangibility as the essence of services.

Simultaneous production and consumption is a characteristic of most services. Goods are first produced, then sold, then consumed; services are sold first, then produced and consumed simultaneously (Regan, 1963). This simultaneous production and consumption means that the consumer is present when the production takes place, and the service provider is present when consumption takes place. This has several important implications. The buyer of the service may become intimately involved with the production process (Carmen and Langeard, 1980), and increasing importance is placed on the delivery aspects of

the transaction · (Berry, 1980). In most service situations, the producer and the seller of the services are the same entity (Upah, 1980), and there is no middleman. This duality of roles usually forces production and marketing to become highly interrelated (Gronroos, 1978).

Heterogeneity refers to the high potential for variability within a service situation. As Zeithaml et al. (1985) state: "The quality and essence of a service can vary from producer to producer, from customer to customer, and from day to day."(p. 34). Service industries differ on the extent to which they are 'people-based' or 'equipment-based' (Thomas, 1978), and the outcomes of service situations involving 'people-based' service firms tend to be less standardized and uniform than outcomes from 'equipment-based' service and goods-producing industries (Berry, 1980). Service situations may also involve many different employees coming in contact with the customer, and problems may arise with regard to the consistency of their behavior (Langeard et al., 1981). Individual employee performance levels may fluctuate as well (Knisely, 1979), and consistency from day-to-day and even moment-to-moment is never assured.

Generally, services cannot be saved or inventoried. The perishability of services means that factors such as capacity, production scheduling, and peak load periods become extremely important to service firms (Rathmell, 1974). It is also difficult for service providers to synchronize supply and demand (Zeithaml et al., 1985), which may result in either too much or too little of the service being available at any given time. With services, excess supply is wasted, and excess demand is lost sales.

2.1 IMPACTS OF THE UNIQUE CHARACTERISTICS OF SERVICES

An examination of some of the effects the unique attributes of services have on the delivery can further serve to distinguish between goods and services.

Quality has traditionally been defined in terms of goods production. The traditional 'Japanese philosophy' defines quality based on the number of defects detected, with 'zero defects' in the product being the goal (Ebrahimpour, 1985). Quality may also be determined by the degree to which goods conform to requirements (Crosby, 1979), with these requirements being defined or influenced by management, government, and consumers. Counting the number of failures of a good is another proposed way of measuring quality (Garvin, 1983). In discussing this, Garvin (1983) also distinguishes between internal failures, which are detected before the good leaves the factory, and external failures, which are detected after the sale is complete.

Unfortunately, knowledge regarding quality in goods production is generally not transferable to the area of service production. This is a result of the unique service dimensions already mentioned: intangibility, inseparability of production and consumption, heterogeneity, and perishability. It becomes impossible, in most situations, to set precise manufacturing standards to control the production of services. Whereas goods are physically tangible products, services are performances, which are intangible and very time dependant. They generally cannot be stored, tested, or verified in advance of sale to assure quality standards (Parasuraman et al., 1984).

Due to the high labor content and interaction requirements of many types of services, it is more difficult to control the quality of service a consumer receives than the quality of a good. A point made in the earlier discussion serves to illustrate: "... a service can vary from producer to producer, customer to customer, and from day to day."(p. 34) (Zeithaml et al., 1985). So much of the content of a service transaction is personal in nature that it becomes difficult to maintain levels and consistencies of behavior. This high degree of interaction introduces all the ambiguities of human behavior and interpersonal communication, and permits further complications to the service delivery process (Booms and Nyquist, 1981). Also, as a result of these factors, it is often not easy to determine if the service is being delivered in the fashion it was originally planned and promoted.

The inseparability of production and consumption may also lead to some of the same problems. Lehtinen and Lehtinen (1982) point out that in many service situations, the quality of the process is determined during a personal interaction between service personnel and the consumer. Again, it can be difficult to maintain levels of quality service during these transactions. Another level of unpredictability can also be introduced when participation by the consumer in the transaction becomes an important input into the final product, and by association, into the evaluation of service quality. In service situations, the quality often occurs during the actual delivery of the service, which generally coincides with the production of the service.

Another important point regarding service quality is illustrated when considering the fact that services are perishable. If services cannot be stored, then testing them for defects and substandard

quality becomes impossible; the consumer receives what is produced in the first attempt at the service procedure. This further stresses the importance of producing the service properly and to satisfaction the first time.

2.2 THE COMPONENTS OF SERVICE QUALITY

The level of perceived service quality has been equated with a consumer's degree of satisfaction. Howard and Sheth (1969) defined satisfaction as a cognitive state relating to rewards obtained for sacrifices undergone. Oliver (1981) stated that satisfaction is a psychological state resulting from a combination of unconfirmed expectations and prior feelings about the situation. Other researchers specify attitude as the major influence on a consumer's perceptions of service quality (Parasuraman et al., 1986). An important issue then becomes the degree to which satisfaction can be equated with attitudes regarding perceived quality.

Several researchers have noted the important distinction between satisfaction and attitudes. Oliver (1981) stated that satisfaction is very transaction specific and is generally an emotional reaction to a specific situation. Attitude, on the other hand, is an enduring affective orientation toward a process or category. It is usually the result of repeated incidents of satisfaction over time (Oliver, 1981).

Olshavsky (1985) explains quality perception as the result of an evaluation process a consumer undertakes after consumption, the final

result being similar to attitude. Holbrook and Corfman (1985) defined quality as being a personal value judgment about the attributes of a product. This value judgment can also be viewed as a form of attitude.

Parasuraman et al. (1986) relate the concepts of satisfaction and attitude to service quality as follows:

> ...service quality is a global judgment or attitude relating to the superiority or excellence or the service whereas satisfaction is related to a specific transaction. The two constructs are related in that incidents of satisfaction over time result in perceptions of service quality.(p.5)

A significant impact on judgments regarding the quality of a service is the outcome of that service transaction (Sasser et al., 1978; Gronroos, 1982a). What is also important is the manner in which the service is performed (Parasuraman et al., 1984; Sasser et al., 1978; Gronroos, 1978a; Lehtinen and Lehtinen, 1982). Several researchers have attempted to extract the specific dimensions of services which form the basis for quality judgments.

Sasser et al. (1978) proposed three dimensions of service quality. The tangible aspects of the service whose use was transferred to the consumer were classified as 'materials'. The physical location of the service provider and equipment used to provide the service were referred to as 'facilities'. Finally, the dimension of 'personnel' captured the delivery of the service and the personal interactions between consumer and service provider.

Gronroos (1982a) identified two different dimensions of service quality. The 'technical quality' dimension included those physical attributes and aspects of the service that have an impact on quality evaluation. There also exists a 'functional quality' dimension of the service; this includes the methods by which the technical quality is transferred to the consumer. 'Corporate image' has also been identified by Gronroos (1982b) as another dimension of service quality. In other research, however, he has stated that corporate image is a result of technical and functional quality (Gronroos, 1982a).

Lehtinen and Lehtinen (1982) focused on the notion that perceptions of service quality are influenced significantly by the interaction between consumers and the personnel of the service organization. They also identify three dimensions of service quality. 'Physical quality' includes those aspects of the service that are physical, or tangible, in nature. 'Interactive quality' is the result of the personal interactions a consumer becomes involved in during the service encounter. The image of the organization providing the service, and its importance, are captured in the dimension of 'corporate quality'. It seems there is very little difference between the classifications expressed by Lehtinen and Lehtinen (1982), Gronroos (1982a; 1982b), and Sasser et al. (1978).

Parasuraman et al. (1984; 1986) have done by far the most comprehensive work to date in the area of quality of service. Preliminary research identified ten dimensions along which consumers base their evaluations of the quality of the services they have received (Parasuraman et al., 1984). A breakdown of these dimensions, their labels, and a discussion of their contents, follows:

1) Reliability:	Consistency of performance and dependability.
2) Responsiveness:	Willingness and readiness of employees to provide the service.
3) Competence:	The possession of the required skills and knowledge to perform the service.
4) Access:	The approachability and ease of contact.
5) Courtesy:	The politeness, respect, consideration, and friendliness of personnel.
6) Communication:	Keeping customers informed in language they can understand, and listening to customers.
7) Credibility:	Trustworthiness, believability, honesty, and having the customer's best interests at heart.
8) Security:	Freedom from danger, risk, or doubt.
9) Understanding/Knowing:	Efforts made to understand the customer's needs.
10) Tangibles:	The physical aspects of the service such as facilities, equipment, and appearance of personnel.

Further analysis resulted in five statistically distinct dimensions which capture facets of the original ten dimensions (Parasuraman et al., 1986). A result of this process is the development of a multiple-item scale for measuring consumers' perceptions of service quality (SERVQUAL). A breakdown of the five dimensions, their labels, and their revised content, follows:

1) Tangibles:	Physical facilities, equipment, and appearance of personnel.
2) Reliability:	Ability to perform the promised service dependably and accurately.
3) Responsiveness:	Willingness to help customers and provide prompt service.

4) Assurance: Knowledge and courtesy of employees, and their ability to convey trust and confidence.

5) Empathy: Caring, individualized attention the firm provides its customers.

The categories of services sampled in developing SERVQUAL include appliance repair and maintenance, retail banking, long distance telephone, securities brokerage, and credit card organizations.

The discussion thus far has focused on subjective measures of the quality of a service. Many objective tools have also been employed in the past, including such measures as the number of complaints received, the time spent by customers in queues waiting for service, the total time spent in the service system, and direct feedback from the customers when questioned about their satisfaction level. These objective measures can usually be reduced to numerical statistics with evaluations and comparisons made on these results.

A problem with employing such objective measures exclusively, and the reason the present research is based more in subjective areas, is that an association between objective numerical results and the quality of a service has not been specified nor documented. For example, a consumer may spend considerable time waiting in a queue for the service process to begin and not have that wait have any bearing on his evaluation of the quality of service that he has received.

If the dimensions of the quality construct are determined, investigation can then be conducted to determine which factors influence evaluation of those dimensions and how important they are, and objective measures of the factors can then be developed. This is a more logical rationale than simply measuring constructs and assuming

that they are the factors on which a customer bases his opinions of service quality.

In examining the literature on what comprises service quality, several things become evident. Important aspects of the process are identified by several researchers; these aspects include both the quality of the tangible attributes and the importance placed on personal interactions. What is also apparent is the fact that, before Parasuraman et al. (1984; 1986), there had been very little empirical research done to determine the criteria by which consumers evaluate the quality of a service.

The developers of SERVQUAL believe that the scale can be used to assess the quality of organizations in a wide range of service applications. The scale has not been tested, however, outside of the five service categories in which it was developed, and has not been tested within the tourism industry. The determination of its applicability in different service industries will be valuable in determining the scale's generalizing properties and its potential use as a tool for managing service quality

2.3 HOW CONSUMERS EVALUATE SERVICE QUALITY

As important as the issue of determining what constitutes service quality is the question of how consumers evaluate the quality of such a service. Gronroos (1978) stressed the importance of understanding how a service will be evaluated by consumers, and suggested that when

this is understood, it will be possible to influence these evaluations in a positive manner.

Before becoming involved in the discussion of how consumers evaluate service quality, however, it will be useful at this time to note the important distinction between the actual dimensions of quality and potential cues of these dimensions. The dimensions of service quality are those aspects of a service process which have an affect on the quality level of the transaction. For example, the reliability associated with a service provider, his ability to provide a service dependably and accurately, may be one specific dimension of service quality. This reliability would then have an effect on the level of service that is provided the consumer.

Cues, on the other hand, are indicators that consumers use to evaluate the performance of service providers on the relevant dimensions of service quality. Referring to the above example, a consumer may evaluate, either consciously or unconsciously, the degree to which the service provider has done what he agreed to do, and this evaluation helps make a determination regarding the reliability of the service provider. While the dimensions of quality service determine the level of service provided, it is usually the cues of the dimensions that the consumers use to evaluate the service transaction. This leads to the conclusion that it is important to know what dimensions constitute quality service, and after that, which cues are indicators of these dimensions.

In considering how consumers evaluate the quality of service, Sasser et al. (1978) identified seven attributes of a service that may play a role in the quality evaluations consumers make. The attributes

include security, consistency, attitude, completeness, condition, availability, and timing. The researchers discuss three models which they believe form the basis for the evaluation of service quality. The first model proposes an evaluation based on one overpowering attribute; that is, the performance of the service organization on one factor will determine the consumer's evaluation. A second model operates on the assumption that the evaluation is being based on the performance of one attribute, with threshold minimum performance levels for the other attributes. As long as these minimum levels are met, these attributes play no role in the overall evaluation with the entire evaluation based on the most important attribute. The third model proposes a theory whereby service quality evaluations are based on a weighted average of all the attributes. The weights assigned each attribute would be a function of and influenced by the consumer.

Churchill and Suprenault (1982) modelled the process of a consumer's comparison of expectations with actual physical product performance with a disconfirmation paradigm. This paradigm maintains that a consumer's expectations are confirmed when a product performs as expected, negatively disconfirmed when a product performs more poorly than expected, and positively disconfirmed when a product performs better than expected. It is then hypothesized that consumer satisfaction is related to the size and direction of the disconfirmation phenomena. When expectations are confirmed or positively disconfirmed, a consumer will experience satisfaction; when expectations are negatively disconfirmed, dissatisfaction will result.

Smith and Houston (1982) adopted the disconfirmation paradigm in their research, which suggests that the size and direction of the

confirmation/disconfirmation experience is based on the degree to which a service transaction follows a predisposed cognitive script. It is hypothesized that a service transaction may be conceptualized by a consumer as an event, composed of a set of actions, actors, and objects. A consumer, through repeated involvement in service transactions, may develop a cognitive script for the service process on which he will base his expectations of the event. The extent to which the script-defined expectations are met determine the consumer's level of satisfaction with the service transaction.

Some researchers have also suggested that the price of a service may become a major cue in the evaluation of quality in situations where other information may not be available (McConnell, 1968; Olander, 1970; Zeithaml, 1981; Tull et al., 1964). In a service transaction involving very little interaction between the service provider and the consumer, this could be the case. Also, service transactions that could be considered very routine or very minor may be evaluated based solely or primarily on the cue of price.

Most models that attempt to explain how consumers evaluate the quality of a service agree on one point: consumers base their evaluations of the quality of a service on a comparison of expectations with performance. The result of this comparison is perceived quality (Sasser et al., 1978; Gronroos, 1982a; Lehtinen and Lehtinen, 1982; Lewis and Booms, 1983; Parasuraman et al., 1984; 1986). Consumers form expectations about the future performance of a product when it is acquired (Engel et al., 1973). As the product is consumed, consumers compare the performance of the product to their prior expectations (Swan and Combs, 1976). Gronroos (1982a) states:

It is thus reasonable to state that the perceived quality of a given service will be the outcome of an evaluation process, where a consumer compares his expectations with the services he perceives he has got; i.e., he puts the perceived service against the expected service. The result of this process will be the perceived quality of service.(p.9)

Similarly, Parasuraman et al. (1986), in developing SERVQUAL, discussed the process used by consumers in evaluating a service:

...service quality, as perceived by consumers, stems from a comparison of their expectations of the services they will receive with their perceptions of the performance of firms providing the services. Accordingly, perceived service quality is the degree and direction of the discrepancy between consumers' perceptions and expectations.(p. 5)

There arises here a conflict between attitude and satisfaction. Parasuraman et al. (1986) have at one point postulated that service quality is an attitude relating to the global superiority of a service and that satisfaction is transaction specific. They further stated that incidents of satisfaction over time result in perceptions of service quality. However, in proposing and developing the SERVQUAL scale, the researchers have not restricted their definition of service quality as being a global judgement made up of incidents of satisfaction over time. Instead, quality is simply defined as being

the degree and direction of the discrepancy between consumers' perceptions and expectations.

In order to attempt to measure perceived service quality for an individual service transaction, it is necessary to make the assumption that consumers make service quality evaluations after and about each transaction. This is certainly not an unreasonable assumption. What is also evident from the different opinions regarding attitude as a global judgement of service quality is that consumers also carry with them opinions about the quality of service that organizations provide. This would be a more enduring judgement made up of incidents of quality service evaluation over time. Given this, the short term or immediate quality service evaluations are related to the overall satisfaction a consumer has with an individual service transaction, while a more enduring quality service judgement made up of repeated individual transactions would be related to an attitude.

The basis of the process whereby consumers evaluate quality is the comparison of perceptions and expectations. Whether the perceived quality is based on specific attributes of the service, or in the absence of these, on a cue such as price, the commonality of the perception-expectation comparison remains the same. In addition, Parasuraman et al. (1984) point out that in a service transaction, expectations may be viewed as the desires and wants of consumers. Thus, the perceived quality of a service is based on an evaluation process where consumers compare the service they have received with their wants and desires from the transaction. Along with the theoretical background supporting such a postulation, this is an intuitively appealing proposition.

2.4 QUALITY OF SERVICE IN THE TOURISM INDUSTRY

The theories regarding a tourist's perceptions of a tourism experience are remarkably similar to those expressed in the general services marketing literature. As explained by Graefe and Vaske (1987), tourists engage in tourism activities with the expectation that their actions will lead to certain rewards. The specific expectations that tourists hold can be influenced by both individual and environmental factors. In evaluating a tourism service transaction, tourists compare the outcomes they actually experienced with the rewards they expected from the procedure. The overall evaluation of the quality of a tourism service will be determined by the extent to which the experience matches the expected outcome.

As has been previously discussed, several researchers have attempted to generate consumer influenced sets of items that can be used to gauge, among other things, satisfaction levels. O'Brien et al. (1977) identified eight factors along which airline companies may be evaluated by consumers. These factors included:

- -helpfulness
- -reliability
- -inconvenience
- -scheduling
- -comfort
- -safety
- -meals
- -on-time arrivals

At the time, the approach used in this study was unique, as it generated and employed subject selected attributes. The researchers also noted that they did not believe their results to be conclusive;

rather, the intent of the research was to demonstrate that the procedure employing consumer-selected attributes could be applied across different classes of services.

Etherington and Var (1984) adopted this procedure in trying to determine which criteria consumers use to select airline companies. An interpretation of the results could be applied to a quality-criteria framework, as the attributes a consumer/tourist uses in selecting an airline could be argued to be similar to or the same in nature to those used in assessing the quality of the service offered. The criteria extracted from this process included:

    -convenient schedules
    -handling at the airport
    -services in flight
    -prices
    -airline employees

Each of these factors was also ranked in importance, and compared across non-business and business travellers. The researchers concluded that the importance of factors shifted from economic constraints for non-business travellers to time constraints for business travellers.

Makens and Marquardt (1977) generated consumer input to evaluate the importance of fifteen factors in air travel that were identified after consultation with the marketing departments of several major commercial airlines. These factors were:

    -ticket price
    -overall seating comfort
    -overall spaciousness
    -length of flight

        -noise level
        -meal and table service
        -menu selection
        -prompt deplaning
        -attention given by hostess
        -privacy
        -luggage allowance
        -storage space at seat
        -drink price
        -meeting people socially
        -making business contacts

The importance of these factors was also compared across first class and coach travellers, with the major difference occurring in the importance associated with ticket price.

Ritchie et al. (1980) solicited consumer input to aid in the extraction of eight dimensions that affect the choice of airline services. The eight dimensions identified were:

        -flight schedules
        -aircraft characteristics
        -fare prices
        -absence of restrictions
        -flight related service
        -safety considerations
        -reservation conditions
        -auxiliary service

The researchers also found differences in importance between business travellers and vacationers in three areas: flight schedules were considered more important to business travellers, while vacation travellers tended to consider prices and safety considerations more important.

Pizam et al. (1978) attempted to extract factors tourists use when evaluating satisfaction with a destination area. Eight factors were derived:

- -beach opportunities
- -cost
- -hospitality
- -eating and drinking facilities
- -accommodation facilities
- -campground facilities
- -environment
- -extent of commercialization

These factors are very destination specific, and the researchers suggest that they would only be applicable to destination areas that are similar to Cape Cod, Mass., where the study was conducted. They also make a call for research into determining similar sets of factors for different types of destination areas.

Other specific factors have been identified that can, for example, have an effect on satisfaction levels within a destination area. Usage and density levels can affect a tourist's perceptions regarding a destination area (Hall, 1974; Gramann, 1982; Kaur, 1979). Also, evidence of human impacts on the environment has been shown to be detrimental to satisfaction levels (Lucas, 1979; Stankey, 1973).

The importance of the distinction between the actual dimensions of quality service and cues to those dimensions may be seen when reviewing the work done in this area. Some studies have attempted to extract specific dimensions, (O'Brien et al., 1977; Etherington and Var, 1984; Ritchie et al., 1980), while others tend to have discussed cues more than dimensions (Makens and Marquardt, 1977; Pizam et al., 1978). The disparity in the results can be substantial if the distinction between dimensions and cues is not noted.

What quickly becomes apparent when reviewing the research that has been done in the area of quality of service within the tourism industry is how fragmented it is. There has been some research

conducted within various segments of the industry, but there has been no coordinated effort toward developing an aid for the entire tourism industry. What is required at this point in time is the development or application of a scale or set of scales to help understand and measure quality of service within the tourism industry. An important result of the proposed application of SERVQUAL to the tourism industry will be an examination of similarities and differences between the tourism industry and other service industries, and a similar comparison across different segments of the tourism industry.

CHAPTER 3

METHODOLOGY

3.0 PURPOSE

The overall purpose of the research was to determine the extent to which the scale developed by Parasuraman et al. (1986), which was formulated to measure service quality across a wide range of service categories, can be used to measure and evaluate service quality within selected segments of the tourism industry.

3.1 OBJECTIVES

The specific objectives of the research were as follows:

1) To measure expectations and perceptions of services using the SERVQUAL scale for one of the categories previously studied by the developers of the scale, and for four tourism services. The category used by Parasuraman et al. (1986) whose service quality was measured in this research was:

    Banking services

The tourism services whose quality was measured were:

    Airline services

    Hotel services

    Ski Area services

    Restaurant services

2) To compare the results obtained from the Parasuraman et al. (1986) research with the results obtained in this research for the category of service that was common to both: banking services.

3) To compare the results obtained from the Parasuraman et al. (1986) research with the results obtained in this research across the tourism categories of services.

4) To formulate conclusions concerning the adequacy of the SERVQUAL scale to tourism services, and if it appears to be inadequate, to provide recommendations to improve the applicability of the scale to tourism services.

When these objectives are met, it will be easier for members of the tourism industry to manage service quality within their organizations and provide the best product offering possible.

3.2 HYPOTHESES

The following hypotheses were proposed:

H1: There will be no significant difference between the results obtained by Parasuraman et al. (1986) and those obtained by this

research with regard to the service category that was common to both: banking services.

H2: There will be no significant difference between the results obtained by Parasuraman et al. (1986) and those obtained by this research as a result of a comparison of the services used in developing the scale and the tourism services tested.

H3: There will be no significant difference between the results obtained by this research for the banking service category and the tourism service categories.

3.3 DATA COLLECTION

The data for the initial analysis of the SERVQUAL scale was gathered from a random sampling procedure and included 200 respondents for each of the five service categories. In order to be qualified to respond to the questionnaire, it was necessary for the respondent to be 21 years of age or older and a current or recent user of the service category. For the purposes of this research, a recent user was defined as being a user who had participated in a service transaction in the service category within the previous three months. The respondents were recruited from the Undergraduate and Graduate Management programs at the University of Calgary, the Adult Education Programs at Southern Alberta Institute of Technology and Mount Royal College, and the technical programs at the Alberta Vocational Centre

in Calgary. By sampling from these different schools, respondents encompassing a reasonably wide variation of demographic characteristics were surveyed.

The three part questionnaire was self administered by qualified respondents after an introduction to the research, explanation of what was desired, and notification of the right not to participate. Each respondent was assigned randomly to a service category, and if they were not a current or recent user of that service, reassigned randomly to another service category. This process ensured that a varied and random number of surveys for each category was completed during each session.

The first section of the survey required respondents to express their expectations of firms within the specific service category they had been assigned. The second section entailed the respondent naming the firm with which they had had the most recent experience within the service category, and providing information regarding his or her perceptions about that particular firm's performance during that service transaction. The final section of the questionnaire included questions relating to the respondent's overall experiences with and perceptions about the firm being considered, and included a general question about the perceived adequacy of the survey. A copy of the questionnaire is included in the Appendix.

Mean expectation and perception scores were generated by combining the scores for the individual statement sets for each dimension and calculating the average of the total score. Individual dimensions that had one or more statements to which there was no response were assigned no mean value for that dimension. Quality

scores were calculated in the same fashion, and dimensions that had a missing response in a statement set for either expectations or perceptions were assigned no value.

Approximately 7% of the respondents interviewed also participated in a post-interview discussion with the survey administrator in order to generate some informal feedback regarding the survey instrument. These interviews were unstructured and ranged in length from 10 to 40 minutes, and were valuable for generating ideas that could be investigated quantitatively once the results from the research were tabulated. All of the surveys and the interviews were conducted by this researcher.

This methodological approach is similar to the one employed by Parasuraman et al. (1986) in developing the SERVQUAL scale. This consistency was desirable in order to maintain as much similarity between the studies as possible since cross study comparisons have been made and conclusions drawn based on these comparisons.

3.4 DATA ANALYSIS

The statistical analysis of the data in this study and the data reported by Parasuraman et al. (1986) were broken down into three phases.

3.4.1 PHASE 1

The statistical results of the research in terms of mean expectation, perception and quality scores were first computed. These results, along with variances and numbers of respondents, are presented in the results section. Frequencies are also reported for the questions that required yes/no or excellent/good/fair/poor responses. Examples of these questions include whether the respondent had reported problems with the service they were considering or what was their overall rating of the quality of the service and the quality of the survey.

3.4.2 PHASE 2

The second phase of the data analysis entailed the comparison of mean expectation, perception, and quality scores by means of T-tests and analysis of variance (ANOVA). In comparing the results of the research done by Parasuraman et al. (1986) with this research, multiple Student's T-tests were used. More sophisticated techniques, such as the use of ANOVA, were not applicable in this situation, as the original data used in developing SERVQUAL was not available. When comparing the results of tourism services with those of other service categories tested in this research, however, the use of ANOVA was applicable to determine whether the results varied significantly for all the service segments compared. When the results of the ANOVA indicated that there was a significant difference, multiple Student's

t-tests were used to isolate which comparisons contributed to that difference.

This phase of the data analysis allowed for the evaluation of the hypotheses that were presented. Decisions to reject or fail to reject the hypotheses were made and justified, and possible reasons for the results are discussed.

3.4.3 PHASE 3

The third phase of the data analysis entailed the examination of the data for specific trends, problems or abnormalities. Statistical methods that were applicable in analyzing the data for specific situations were used and the implications of the findings discussed.

CHAPTER 4

RESULTS

4.0 DIMENSION SCORES

Among the objectives of the research was the desire to compare the results obtained by the scale developed by Parasuraman et al. (1986) with those obtained by this research for the category of services that was common to both. In addition, comparisons of the results obtained within the tourism scales to those of a segment of the SERVQUAL research were also desirable. In order to facilitate these comparisons, however, it was first necessary to examine the individual results of the research for each of the categories of services; banking, airlines, hotels, restaurants, and ski areas.

The survey directly measured consumers' expectations and perceptions regarding a recent service transaction in which they had participated. Quality scores were arrived at by adopting an assumption of the developers of the SERVQUAL scale, that the difference between consumers' expectations and perceptions constituted the quality associated with that service transaction. Using this logic then, the higher that absolute quality score, or the less negative the score, the higher the quality of the service transaction.

4.0.1 BANKING

Table 4.0 shows a breakdown of the results obtained by the research with regard to the service category that was common to both research projects, banking services. Mean expectation scores ranged from a high of 6.68 for the dimension Reliability to a low of 5.67 for

TABLE 4.0

EXPECTATION, PERCEPTION AND QUALITY SCORES
OF BANKING SERVICES

Expectations

	Mean	Std. Dev	N
Tangibles	6.15	0.70	200
Reliability	6.68	0.38	198
Responsiveness	5.80	1.03	200
Assurance	6.58	0.43	195
Empathy	5.67	0.87	196
Combined Scale	6.21	0.43	189

Perceptions

	Mean	Std. Dev	N
Tangibles	5.72	0.89	199
Reliability	5.20	1.28	199
Responsiveness	4.31	1.48	198
Assurance	5.26	1.23	197
Empathy	4.22	1.35	197
Combined Scale	4.93	1.08	192

Quality

	Mean	Std. Dev	N
Tangibles	-0.43	0.94	199
Reliability	-1.46	1.31	197
Responsiveness	-1.50	1.69	198
Assurance	-1.31	1.22	192
Empathy	-1.42	1.55	194
Combined Scale	-1.27	1.10	186

the dimension Empathy. The dimension Assurance had the second highest mean expectation score of 6.58, followed by Tangibles at 6.15 and Responsiveness at 5.80.

The dimension Tangibles had the highest mean perception score, which was 5.72. This was followed by the dimensions Assurance at 5.26, Reliability at 5.20, and Responsiveness at 4.31. The Empathy dimension had the lowest mean perception score for banking services, at 4.22.

Mean quality scores ranged from a high of -0.43 for the Tangibles dimension to a low of -1.50 for the dimension Responsiveness. The second highest mean quality scores were obtained by the Assurance dimension at -1.31, followed by Empathy at -1.42 and Reliability at -1.46.

4.0.2 AIRLINE

A breakdown of mean expectation, perception and quality scores for airline services is presented in Table 4.1. Mean expectation scores varied from a high of 6.46 for the Reliability dimension to a low of 5.76 for the Empathy dimension. Assurance had the second highest expectation score of 6.44, followed by Tangibles at 6.39 and Responsiveness at 5.79.

The dimension Tangibles had the highest mean perception score, at 5.90. This was followed by Assurance at 5.53, Reliability at 5.21, and Responsiveness at 5.07. Empathy had the lowest mean perception score at 4.95.

TABLE 4.1

EXPECTATION, PERCEPTION AND QUALITY SCORES
OF AIRLINE SERVICES

Expectations

	Mean	Std. Dev	N
Tangibles	6.39	0.58	197
Reliability	6.46	0.48	200
Responsiveness	5.79	0.85	191
Assurance	6.44	0.50	194
Empathy	5.76	0.87	199
Combined Scale	6.18	0.43	185

Perceptions

	Mean	Std. Dev	N
Tangibles	5.90	0.87	198
Reliability	5.21	1.16	197
Responsiveness	5.07	1.28	197
Assurance	5.53	1.10	197
Empathy	4.95	1.14	199
Combined Scale	5.35	0.92	186

Quality

	Mean	Std. Dev	N
Tangibles	-0.49	0.92	195
Reliability	-1.26	1.23	197
Responsiveness	-0.70	1.44	188
Assurance	-0.87	1.09	191
Empathy	-0.83	1.36	197
Combined Scale	-0.79	0.96	175

The highest mean quality score belonged to the Tangibles dimension, at -0.49. This was followed by the dimensions Responsiveness at -0.70, Empathy at -0.83, and Assurance at -0.87. The lowest mean quality score for airline services was the dimension Reliability, with a score of -1.26.

4.0.3 HOTEL

Table 4.2 contains a breakdown of the mean expectation, perception, and quality scores for hotel services. The range of mean expectation scores varies from a high of 6.43 for the dimension Reliability to a low of 5.71 for the dimension Empathy. Assurance represented the second highest mean expectation level with a score of 6.38, followed by Tangibles at 6.15 and Responsiveness at 5.76.

The highest mean perception score belonged to the dimension Tangibles, with a score of 5.49. This was followed by the dimensions Assurance, Reliability, and Empathy, with scores of 5.36, 5.26, and 5.05 respectively. The lowest mean perception score was attributable to the dimension Responsiveness, which had a score of 5.01.

Quality scores ranged from a high score of -0.66 in the case of the dimension Tangibles to a low of -1.16 for Reliability. The second highest score belonged to the dimension Empathy with a score of -0.67, followed by Responsiveness at -0.75 and Assurance at -1.00.

TABLE 4.2

EXPECTATION, PERCEPTION AND QUALITY SCORES
OF HOTEL SERVICES

Expectations

	Mean	Std. Dev	N
Tangibles	6.15	0.69	198
Reliability	6.43	0.56	200
Responsiveness	5.76	0.80	200
Assurance	6.38	0.50	198
Empathy	5.71	0.83	200
Combined Scale	6.11	0.45	196

Perceptions

	Mean	Std. Dev	N
Tangibles	5.49	1.09	198
Reliability	5.26	1.03	200
Responsiveness	5.01	1.29	200
Assurance	5.36	1.10	196
Empathy	5.05	1.20	199
Combined Scale	5.24	1.03	193

Quality

	Mean	Std. Dev	N
Tangibles	-0.66	1.03	198
Reliability	-1.16	1.12	200
Responsiveness	-0.75	1.28	200
Assurance	-1.00	1.05	194
Empathy	-0.67	1.26	199
Combined Scale	-0.85	0.95	191

4.0.4 RESTAURANT

A breakdown of mean expectation, perception and quality scores for restaurant services is detailed in Table 4.3. The mean of expectation scores ranged from a high of 6.33 in the case of the Assurance dimension to a low of 5.42 for the Empathy dimension. The second highest mean expectation score resulted from the Reliability dimension with a score of 6.18, followed by Tangibles at 6.03 and Responsiveness at 5.47.

For the service category of restaurants, the dimension Tangibles had the highest mean perception score at 5.69. This was followed by Assurance with a score of 5.43, Reliability at 5.17, and Empathy at 4.91. The dimension Responsiveness had the lowest mean perception score at 4.74.

The highest mean quality score belonged to the dimension Tangibles, with a score of -0.35. The dimension Reliability had the lowest mean quality score at -1.00. Tangibles was followed in order by the dimension Empathy at -0.59, Responsiveness at -0.74, and Assurance at -0.88.

4.0.5 SKI AREA

Table 4.4 shows a breakdown of the results obtained by the research with respect to the service category of ski areas. Mean expectation scores ranged from a high of 6.31 for the dimension Assurance to a low of 5.28 for the dimension Empathy. The dimension

TABLE 4.3

EXPECTATION, PERCEPTION AND QUALITY SCORES
OF RESTAURANT SERVICES

Expectations

	Mean	Std. Dev	N
Tangibles	6.03	0.69	195
Reliability	6.18	0.63	200
Responsiveness	5.47	1.17	198
Assurance	6.33	0.50	197
Empathy	5.42	0.92	200
Combined Scale	5.91	0.53	193

Perceptions

	Mean	Std. Dev	N
Tangibles	5.69	1.01	198
Reliability	5.17	1.10	195
Responsiveness	4.74	1.37	198
Assurance	5.43	1.10	197
Empathy	4.91	1.23	198
Combined Scale	5.22	0.99	187

Quality

	Mean	Std. Dev	N
Tangibles	-0.35	1.10	194
Reliability	-1.00	1.21	195
Responsiveness	-0.74	1.55	196
Assurance	-0.88	1.12	190
Empathy	-0.59	1.31	198
Combined Scale	-0.70	1.03	182

TABLE 4.4

EXPECTATION, PERCEPTION AND QUALITY SCORES
OF SKI AREA SERVICES

Expectations

	Mean	Std. Dev	N
Tangibles	5.84	0.76	200
Reliability	6.11	0.55	196
Responsiveness	5.67	0.89	199
Assurance	6.31	0.48	195
Empathy	5.28	0.90	198
Combined Scale	5.87	0.50	188

Perceptions

	Mean	Std. Dev	N
Tangibles	5.41	1.02	199
Reliability	5.11	0.75	199
Responsiveness	4.79	1.09	199
Assurance	5.36	0.79	200
Empathy	4.66	0.93	200
Combined Scale	5.08	0.73	197

Quality

	Mean	Std. Dev	N
Tangibles	-0.44	0.44	199
Reliability	-0.99	0.74	195
Responsiveness	-0.88	1.23	199
Assurance	-0.94	0.80	195
Empathy	-0.62	1.02	198
Combined Scale	-0.77	0.74	186

Reliability had the second highest mean expectation score of 6.11, followed by Tangibles at 5.84 and Responsiveness at 5.67.

The dimension Tangibles had the highest mean perception score, with a score of 5.41. This was followed by the dimensions Assurance with a score of 5.36, Reliability at 5.11, and Responsiveness at 4.79. The dimension Empathy had the lowest mean perception score at 4.66.

Mean quality scores for ski area services ranged from a high of -0.44 for the dimension Tangibles to a low of -0.99 for the dimension Reliability. Following Tangibles and preceding Reliability are the dimensions Empathy, Responsiveness, and Assurance, with mean quality scores of -0.62, -0.88, and -0.94 respectively.

4.1 SERVICE TRANSACTION SATISFACTION LEVELS

In order to be able to compare the scores of people who had received satisfactory service to those of people who had not received satisfactory service, several questions regarding the level of satisfaction with the service transaction were asked. These questions inquired of the respondent whether they had ever reported a problem with the services they had received from the service giver, whether they would recommend the service giver to a friend, and what they felt the overall quality of the service transaction was.

4.1.1 REPORTED PROBLEM

The first question asked the respondent whether they had ever reported a problem with the services they had received from the particular service provider that they had been considering when responding to the survey. Table 4.5 presents a breakdown of the responses by service category. Problems were encountered trying to relate these results with other data as there was no question asking the respondent how often he had used the service provider. Given the lack of such a question and that there was a much smaller likelihood of having a problem reported in the past if the frequency of usage was low, the results did not facilitate comparisons with other data to attempt to extract trends. Examined on their own, however, the results are still valid and interesting.

As is illustrated in the table, respondents completing the banking survey reported the highest incidence of problem occurrence, at 39.9% of the respondents. This was followed by restaurant services at 24.2%, hotel services at 21.6%, and airline services at 20.7%. Respondents completing ski area services surveys reported the lowest incidence of problem reporting, at 8.6%. In comparison to these figures, the average percentage of respondents reporting problem occurrence for the entire series of surveys is 23.0%

TABLE 4.5

FREQUENCY AND PERCENTAGE OF RESPONDENTS
REPORTING PROBLEM

	REPORTED PROBLEM		NO REPORTED PROBLEM	
	N	%	N	%
Banking	79	39.9%	119	60.1%
Airline	40	20.7%	153	79.3%
Hotel	43	21.6%	156	78.4%
Restaurant	46	24.2%	144	75.8%
Ski Area	17	8.6%	180	91.4%
Total	225	23.0%	752	77.0%

4.1.2   RECOMMENDATION

The next question asked respondents if they would recommend to a friend the particular service giver they were considering in responding to the survey. Unlike the question regarding whether the respondent had ever reported a problem, this question is less vulnerable to the frequency of usage issue and can be later used to draw comparative conclusions. Table 4.6 shows a breakdown of its responses by service category.

Within the service category of ski areas, 86.7% of respondents stated they would recommend this service giver to a friend; this was the highest for all the surveyed segments. This was followed by restaurant services at an 83.6% recommendation level, airline services at 80.8% and hotel services at 79.0%. Banking services had the lowest level of recommendations at 65.5%. Overall, the average percentage of respondents who would recommend the service giver they were considering to a friend was 79.1%.

4.1.3   OVERALL QUALITY OF SERVICE

Respondents were also asked to rate the service giver on an overall quality of service dimension. A breakdown of the responses to this question is provided in Table 4.7.

For the banking service sector, good service quality was reported by 50.0% of respondents. This was followed by the evaluations fair at 26.8%, excellent at 17.2%, and poor at 6.1%.

TABLE 4.6

FREQUENCY AND PERCENTAGE OF RESPONDENTS
RECOMMENDING SERVICE PROVIDER

	WOULD RECOMMEND		WOULD NOT RECOMMEND	
	N	%	N	%
Banking	127	65.5%	67	34.5%
Airline	156	80.8%	37	19.2%
Hotel	154	79.0%	41	21.0%
Restaurant	158	83.6%	31	16.4%
Ski Area	169	86.7%	26	13.3%
Total	764	79.1%	202	20.9%

TABLE 4.7

OVERALL QUALITY OF SERVICE
RATING FREQUENCIES

	EXCELLENT	GOOD	FAIR	POOR
Banking	17.2%	50.0%	26.8%	6.1%
Airline	23.2%	59.8%	10.3%	6.7%
Hotel	31.7%	43.7%	23.6%	1.0%
Restaurant	28.8%	48.2%	14.7%	8.4%
Ski area	22.8%	62.9%	12.7%	1.5%
Total	24.7%	52.9%	17.7%	4.7%

Note: Totals may not equal 100.0% due to rounding.

For each of the tourism services sampled, the order of ranking categories was consistent; good was the category most often chosen by the respondents, followed in order by excellent, fair and poor. Airline services were rated good by 59.8% of respondents, while hotel services had good ratings by 43.7%. 48.2% of restaurant service respondents rated the service they had received as good, and 62.9% had the same rating for ski areas.

Excellent service quality was reported by 23.2% of airline respondents, as compared to 31.7% of hotel respondents and 28.8% of restaurant respondents. Ski area respondents reported excellent service with a percentage frequency of 22.8%.

Airline services had the lowest percentage of respondents reporting fair service, at 10.3%. This was followed by hotel services, the highest at 23.6%, restaurant services at 14.7%, and ski area services at 12.7%.

For all the service categories, banking included, the poor service category had the lowest average response rate. Respondents considering airline services reported they had encountered poor service at a 6.7% rate. Hotel services had an average percentage frequency for this category at 1.0%, while restaurant service respondents answered in this category 8.4% of the time. Ski area respondents reported poor service at a rate of 1.5%.

4.2 QUALITY OF SURVEY

The respondents tested in this research program were asked to give their rating of the quality of the measuring instrument used in the survey. Based on the degree to which they felt the survey captured their expectations and perceptions of the service transaction, respondents answered in one of the four categories used to reflect their impressions of the quality of the service they had received.

A breakdown of the responses to this question, differentiated by service segment, is shown in Table 4.8. Of the respondents answering the banking service survey, 12.9% felt that it did an excellent job capturing their expectations and perceptions regarding banking services. 59.8% of respondents felt the survey did a good job in this regard, while 25.3% felt it did a fair job and 2.1% thought it was a poor instrument.

In evaluating the airline segment survey, 6.8% of the respondents felt it did an excellent job. 59.2% of respondents answered good when asked the question, while 31.4% answered fair. 2.6% of respondents felt the survey did a poor job measuring their expectations and perceptions.

Of the respondents completing the hotel service survey, 4.0% felt that it was an excellent instrument. For 53.3% of the respondents the response to this question was good, while it was fair for 36.2% and poor for 6.5% of those answering it.

The restaurant service segment survey was reported as being and excellent instrument by 6.9% of the respondents completing it. 58.2%

TABLE 4.8

QUALITY OF SURVEY
RATING FREQUENCIES

	EXCELLENT	GOOD	FAIR	POOR
Banking (n=194)	12.9%	59.8%	25.3%	2.1%
Airline (n=191)	6.8%	59.2%	31.4%	2.6%
Hotel (n=199)	4.0%	53.3%	36.2%	6.5%
Restaurant (n=189)	6.9%	58.2%	30.2%	4.8%
Ski area (n=192)	2.6%	60.4%	33.3%	3.6%
Total (n=965)	6.6%	58.1%	31.3%	3.9%

Note: Totals may not equal 100.0% due to rounding.

of respondents felt it was a good tool, while 30.2% felt it was fair. A total of 4.8% of those surveyed in this segment felt that the instrument was poor.

In responding to this question in the ski area service survey, 2.6% of those interviewed felt that the instrument was excellent. The response good was given by 60.4% of respondents, while 33.3% felt that it was fair. 3.6% of those surveyed using this instrument classified it as being a poor tool.

The combined results for this question for all the service segments show that 6.6% of respondents felt the survey was an excellent instrument. 58.1% of respondents felt it was good, while 31.3% felt it was fair. Overall, 3.9% of the total respondents felt that the survey was a poor instrument.

Conclusions regarding how well the survey seemed to work reflecting levels of quality will be drawn during the examination of the results presented in the next section.

CHAPTER 5

EXAMINATION OF RESULTS

5.0 DIMENSION SCORES - COMPARISONS TO SERVQUAL BANKING RESULTS

One of the objectives of this research was the comparison of its results with those obtained by the SERVQUAL research. Specifically, two categories of comparisons can be made: comparison of the results obtained for the service category that was common to both studies, banking services, and comparison of the results obtained for the tourism service categories tested with those obtained by the SERVQUAL research. For the purposes of this second set of comparisons, the results obtained for banking services will be those used to represent the Parasuraman research.

The comparisons will be made using Student's t-tests, as the unavailability of the raw data for the SERVQUAL research prohibits the use of any other applicable statistical approach. Other approaches will be employed in later comparisons and examinations.

5.0.1 BANKING

A comparison of the expectation, perception, and quality scores obtained by the SERVQUAL research and those obtained by this research for the common service segment banking is shown in Table 5.0. The values were compared using Student's two sided t-tests and significant differences at alpha equal to 0.95 and 0.99 are reported.

For the expectations section of the scale, significant differences in the results are apparent for the Reliability, Responsiveness, and Empathy dimensions, as well as the total combined

TABLE 5.0

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES OF BANKING SERVICES

Expectations

	SERVQUAL			FICK		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.05	0.90	197	6.15*	0.69	200
Reliability	6.55	0.76	197	6.68*	0.38	198
Responsiveness	5.48	1.49	197	5.80*	1.03	200
Assurance	6.45	0.83	194	6.58**	0.43	195
Empathy	5.30	1.42	196	5.67**	0.87	196
Combined Scale	5.97	0.74	189	6.21**	0.43	189

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.52	1.08	195	5.72*	0.89	199
Reliability	5.47	1.14	197	5.20*	1.28	199
Responsiveness	4.56	1.57	196	4.31	1.48	198
Assurance	5.38	1.15	194	5.26	1.23	197
Empathy	4.40	1.41	194	4.22	1.35	197
Combined Scale	5.07	1.00	185	4.93	1.08	192

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.52	1.14	194	-0.43**	0.94	199
Reliability	-1.07	1.27	196	-1.46**	1.31	197
Responsiveness	-0.93	1.76	195	-1.50*	1.69	198
Assurance	-1.03	1.33	190	-1.31**	1.22	192
Empathy	-0.90	1.77	193	-1.42**	1.55	194
Combined Scale	-0.88	0.98	177	-1.27**	1.10	186

*   Significant at .95
** Significant at .99

scale score. The only significant difference in the results for the perceptions section of the survey occurs for the dimension Reliability, while for the quality section, differences are apparent in almost all the dimensions; Reliability, Responsiveness, Assurance, and Empathy, as well as the total combined scale score.

Since the quality construct is made up of the values reported for perceptions less the values reported for expectations, it is no surprise that differences existing in the expectations construct are apparent in the quality construct, given that there is only one significant difference within the perceptions construct. For this service segment, this is the case for the dimensions Responsiveness, Empathy and the total scale score. The significant difference seen within the perceptions construct, that of the dimension Reliability, serves to further enhance the significant difference already apparent within the expectations construct for this dimension within the quality construct. Furthermore, the existence of the significant difference for the dimension Assurance within the quality construct is made up of combined differences for that dimension within the expectation and perception, neither of which is significant individually.

5.0.2 AIRLINE

Table 5.1 contains the comparisons of the data generated for the airline segment of the tourism industry with the banking results obtained by SERVQUAL. Again, the values were compared using Student's

TABLE 5.1

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES OF SERVQUAL BANKING SERVICES WITH TOURISM AIRLINE SERVICES

Expectations

	SERVQUAL			FICK		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.05	0.90	197	6.39**	0.58	197
Reliability	6.55	0.76	197	6.46*	0.48	200
Responsiveness	5.48	1.49	197	5.79*	0.85	191
Assurance	6.45	0.83	194	6.44**	0.50	194
Empathy	5.30	1.42	196	5.76**	0.87	199
Combined Scale	5.97	0.74	189	6.18	0.43	185

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.52	1.08	195	5.90**＊	0.87	198
Reliability	5.47	1.14	197	5.21**	1.16	197
Responsiveness	4.56	1.57	196	5.07**	1.28	197
Assurance	5.38	1.15	194	5.53**	1.10	197
Empathy	4.40	1.41	194	4.95**	1.14	199
Combined Scale	5.07	1.00	185	5.35**	0.92	186

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.52	1.14	194	-0.49	0.92	195
Reliability	-1.07	1.27	196	-1.26	1.23	197
Responsiveness	-0.93	1.76	195	-0.70	1.44	188
Assurance	-1.03	1.33	190	-0.87	1.09	191
Empathy	-0.90	1.77	193	-0.83	1.36	197
Combined Scale	-0.88	0.98	177	-0.79	0.96	175

* Significant at .95
** Significant at .99

two sided t-tests and significant differences are reported.

For the section of the scale that measured expectations, significant differences exist for several dimensions, specifically the dimensions Tangibles, Responsiveness, and Empathy. For perceptions, almost all of the dimensions of the airline results differ significantly from the SERVQUAL banking; Tangibles, Reliability, Responsiveness, and Empathy. The total combined scale score for perceptions is also significantly different. However, for the quality construct, no dimension exhibits a significant difference.

It appears that here a cancelling effect is exhibited within the quality construct. Differences in Tangibles, Reliability, and Empathy common to both expectations and perceptions seem to negate each other, leaving the differences in quality scores for these dimensions small enough to be insignificant. The significant differences in Reliability and the total combined scale score that exists when considering perceptions are reduced enough when combined with expectations to eliminate their significance within the quality construct.

5.0.3 HOTEL

A comparison of the expectation, perception, and quality scores of the results from the hotel service segment as compared to SERVQUAL banking services is illustrated in Table 5.2. As is the case for all the segments, Student's two sided t-tests were used to test for significant differences within the mean scores.

TABLE 5.2

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES
OF SERVQUAL BANKING SERVICES WITH
TOURISM HOTEL SERVICES

Expectations

	SERVQUAL			FICK		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.05	0.90	197	6.15	0.69	198
Reliability	6.55	0.76	197	6.43	0.56	200
Responsiveness	5.48	1.49	197	5.76*	0.80	200
Assurance	6.45	0.83	194	6.38**	0.45	198
Empathy	5.30	1.42	196	5.71**	0.83	200
Combined Scale	5.97	0.74	189	6.11*	0.45	196

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.52	1.08	195	5.49	1.09	198
Reliability	5.47	1.14	197	5.26**	1.03	200
Responsiveness	4.56	1.57	196	5.01**	1.29	200
Assurance	5.38	1.15	194	5.36**	1.10	196
Empathy	4.40	1.41	194	5.05**	1.12	199
Combined Scale	5.07	1.00	185	5.24	1.03	193

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.52	1.14	194	-0.66	1.03	198
Reliability	-1.07	1.27	196	-1.16	1.12	200
Responsiveness	-0.93	1.76	195	-0.75	1.28	200
Assurance	-1.03	1.33	190	-1.00	1.05	194
Empathy	-0.90	1.77	193	-0.67	1.26	199
Combined Scale	-0.88	0.98	177	-0.85	0.95	191

*   Significant at .95
** Significant at .99

For the expectations section, significant differences are apparent for the dimensions Responsiveness and Empathy and for the total combined scale score. For the perceptions construct, the same dimensions Responsiveness and Empathy exhibit significant differences, but the total combined scale score for this construct does not. Within the quality construct, however, there are no significant differences for any dimension nor the total combined scale score.

The cancelling effect that was discussed in the previous section seems to become important here as well. Differences apparent in Responsiveness and Empathy for both the expectations and perceptions constructs appear to cancel each other out, leaving neither significantly different within the quality construct. The difference in total combined scale score apparent in the perceptions construct also seems to become insignificant when combined with the expectations construct total combined scale score to make up the score for the quality construct.

5.0.4 RESTAURANT

Table 5.3 contains a comparison of the SERVQUAL banking results with the results obtained by the restaurant service segment surveys. Two sided t-tests were used to test for significant differences between the means.

For the expectations section of the survey, only the dimension Reliability showed any significant difference from the SERVQUAL results. The dimensions Reliability and Empathy were the only

TABLE 5.3

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES OF SERVQUAL BANKING SERVICES WITH TOURISM RESTAURANT SERVICES

Expectations

	SERVQUAL			FICK		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.05	0.90	197	6.03	0.69	195
Reliability	6.55	0.76	197	6.18**	0.63	200
Responsiveness	5.48	1.49	197	5.47	1.17	198
Assurance	6.45	0.83	194	6.33	0.50	197
Empathy	5.30	1.42	196	5.42	0.92	200
Combined Scale	5.97	0.74	189	5.91	0.53	193

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.52	1.08	195	5.69**	1.01	198
Reliability	5.47	1.14	197	5.17**	1.10	195
Responsiveness	4.56	1.57	196	4.74	1.37	198
Assurance	5.38	1.15	194	5.43**	1.10	197
Empathy	4.40	1.41	194	4.91**	1.23	198
Combined Scale	5.07	1.00	185	5.22	0.99	187

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.52	1.14	194	-0.35	1.10	194
Reliability	-1.07	1.27	196	-1.00	1.21	195
Responsiveness	-0.93	1.76	195	-0.74	1.55	196
Assurance	-1.03	1.33	190	-0.88*	1.12	190
Empathy	-0.90	1.77	193	-0.59	1.31	198
Combined Scale	-0.88	0.98	177	-0.70	1.03	182

* Significant at .95
** Significant at .99

dimensions significantly different within the perceptions section of the scale, while only Empathy showed significant differences within the quality construct for this service segment.

It appears that for the restaurant service segment, the differences exhibited in the Reliability dimension in both the expectations and perceptions construct are cancelled in the quality construct. The difference in the Empathy dimension within the perceptions construct is carried over into the quality construct.

5.0.5 SKI AREA

A comparison of the results for the dimension scores obtained by the SERVQUAL banking and the ski area service segment is presented in Table 5.4. The mean scores of each dimension are compared using Student's two sided t-tests and significant differences are reported.

Significant differences in mean expectation scores can be seen for three of the dimensions: Tangibles, Reliability and Assurance. The same is true for Reliability in perception scores, while Empathy is significantly different in this construct as well. Quality scores for ski areas, as compared to the SERVQUAL banking results, differ significantly in no dimension.

The fact that none of the dimensions within the quality construct were significantly different than the SERVQUAL banking results can be attributed to several factors. Differences common to both, in this case for the dimension Reliability, cancel each other out to leave the difference for the dimension within the quality construct insignificant. For dimensions that were significantly different

TABLE 5.4

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES OF SERVQUAL BANKING SERVICES WITH TOURISM SKI AREA SERVICES

Expectations

	SERVQUAL			FICK		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.05	0.90	197	5.84*	0.76	200
Reliability	6.55	0.76	197	6.11**	0.55	196
Responsiveness	5.48	1.49	197	5.67	0.89	199
Assurance	6.45	0.83	194	6.31*	0.48	195
Empathy	5.30	1.42	196	5.28	0.90	198
Combined Scale	5.97	0.74	189	5.87	0.50	188

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.52	1.08	195	5.41**	1.02	199
Reliability	5.47	1.14	197	5.11**	0.75	199
Responsiveness	4.56	1.57	196	4.79	1.09	199
Assurance	5.38	1.15	194	5.36	0.79	200
Empathy	4.40	1.41	194	4.66*	0.93	200
Combined Scale	5.07	1.00	185	5.08	0.73	197

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.52	1.14	194	-0.44	1.19	199
Reliability	-1.07	1.27	196	-0.99	0.74	195
Responsiveness	-0.93	1.76	195	-0.88	1.23	199
Assurance	-1.03	1.33	190	-0.94	0.80	195
Empathy	-0.90	1.77	193	-0.62	1.02	198
Combined Scale	-0.88	0.98	177	-0.77	0.74	186

*   Significant at .95
** Significant at .99

within only one of the constructs, in this case Tangibles, Assurance, and Empathy, insignificant differences for each of these dimensions within the other construct appear to have been enough to reduce the differences to insignificant proportions when combined to produce the quality construct. There is not any cumulative effect for the dimension Responsiveness, which had exhibited no significant differences within the expectations and perceptions constructs.

5.0.6 IMPLICATIONS

In order to meet one of the requirements of effectiveness as a measuring tool, the scale should be able to perform consistently within service segments and be able to distinguish between different service segments. While the scale did not perform consistently within the banking industry for the two research studies, it did distinguish between the tourism service industries and the SERVQUAL banking results where many significant differences in mean scores occurred.

In comparing the banking results between the research projects, it could be theorized that expectation scores would be relatively consistent while perception scores could vary more as a result of the different banking establishments that exist in the two regions where the studies were undertaken. If this were true, quality scores, since they are derived as a difference of expectation and perception scores, could also vary and not be viewed as proof that the scale is not consistent.

The expectation scores of the banking service segment results did vary significantly, however, in three of the five dimensions; Reliability, Responsiveness, and Empathy, as well as with the combined total scale score. While this may seem to indicate that the scale is not exhibiting any consistency while attempting to measure constructs within the same service category, there are also several other factors that should be considered. The results may be affected due to the fact that the SERVQUAL research was conducted in the United States, while this project was undertaken in Canada. The differences in the banking industry between the two countries are quite extensive, and could play a significant role in forming consumers' expectations about banking establishments. There has also been a time period of approximately 2 years between the collation of the results between the two projects, and it is possible that events have occurred in that time period that may have had an effect on consumers' expectations of banking establishments.

The question may then arise as to whether choosing the banking industry as the industry that would be common to both studies was a good choice. This possibility will be examined in the section exploring limitations of the research.

Despite the existence of these extraneous factors, the results support a rejection of the first hypothesis, that there would be no significant different between the results obtained by Parasúraman et al. (1986) and those obtained by this research with regard to banking services. The resultant conclusion is that, in this case, the scale does not perform consistently when applied to one of the same service segments in which it was developed.

With regard to the tourism service segments, mean expectation scores differed significantly from the SERVQUAL banking scores for some dimensions within each segment. This ranged from a high of three dimensions for airline services and ski area services to a low of one dimension for restaurant services. Hotel services had a total of two of the five dimensions having mean expectation scores which were significantly different.

The fact that each tourism service category had at least one dimension significantly different from the SERVQUAL banking dimensions results suggests a conclusion that the scale has some ability to distinguish between different service segments. Given this, it would be appropriate to reject the second hypothesis, that there would be no significant difference between the results obtained by Parasuraman et al. (1986) and those obtained by this study within the services used in developing the scale and the tourism services. It would then be appropriate to conclude that, in this case, the scale was able to distinguish between service segments. The existence of those outside factors remains, however, so confidence in this conclusion is not as high as it could be.

In order to further examine the ability of the scale to make these distinctions, comparisons between the mean scores of banking service expectations and tourism service expectations will be made. By examining these comparisons, the necessity of considering the extraneous factors that were mentioned as possible contributors to the differences illustrated when comparing SERVQUAL banking results is eliminated. For this research study, the segments of the survey were conducted in the same geographical area during the same time frame.

5.1 DIMENSION SCORES - COMPARISONS TO THIS STUDY'S BANKING RESULTS

After examining the results of the comparisons between the results obtained by this research with those obtained by the SERVQUAL research, it is evident that a comparison of the results obtained by this research for the dimension of banking with those obtained for the tourism service categories would be valuable. This should facilitate a determination of the scale's ability to distinguish between service segments, especially those belonging to a significantly different industry.

5.1.1 ALL SERVICE SEGMENTS

Before undertaking a series of Student's t-tests to determine if the mean expectation, perception and quality scores for this study's banking results are significantly different than those of the tourism service segments, a one-way analysis of variance was conducted to determine if the difference in scores was significant across all the service segments. Table 5.5 shows the resultant F-scores and their significance for this statistical procedure.

All of the mean expectation dimension scores exhibit significant differences among the service categories, while the dimensions Tangibles, Responsiveness and Empathy exhibit differences for the

TABLE 5.5

ANALYSIS OF MEAN SCORE VARIANCE BETWEEN
ALL SERVICE SEGMENTS IN THIS STUDY

Expectations

		F-score	Significance
Tangibles	(n=989)	16.45	0.00**
Reliability	(n=993)	38.85	0.00**
Responsiveness	(n=987)	4.26	0.00**
Assurance	(n=978)	9.92	0.00**
Empathy	(n=992)	11.27	0.00**

Perceptions

		F-score	Significance
Tangibles	(n=991)	7.84	0.00**
Reliability	(n=989)	0.53	0.72
Responsiveness	(n=991)	10.45	0.00**
Assurance	(n=982)	1.63	0.16
Empathy	(n=990)	15.53	0.00**

Quality

		F-score	Significance
Tangibles	(n=984)	2.42	0.05*
Reliability	(n=983)	5.72	0.00**
Responsiveness	(n=980)	10.52	0.00**
Assurance	(n=961)	5.43	0.00**
Empathy	(n=985)	14.67	0.00**

*   Significant at .95
**   Significant at .99

perceptions construct. For the quality construct, all the dimensions exhibit significant differences within the service categories. These results indicate that, for the most part, the mean dimension scores are not the same across all service categories. This does not, however, allow for the rejection of the hypothesis that there are no significant differences between the results from this study's banking segment and the tourism segments. In order to make a firm conclusion about the hypothesis, a series of Student's t-tests will be conducted to determine if the differences apparent from the analysis of variance exist between the banking results and the tourism results or within the tourism segment results.

5.1.2 AIRLINE

A comparison of the expectation, perception, and quality scores obtained by this research for banking services with those obtained for airline services is illustrated in Table 5.6. The values were compared using Student's two sided t-tests and significant differences at alpha equal to 0.95 and 0.99 are reported.

Within the expectations construct, significant differences in mean scores can be seen for the dimensions Tangibles, Reliability, and Assurance. Significant differences in the perceptions section are evident for the Tangibles, Responsiveness, Assurance, and Empathy dimensions, as well as with the total combined scale score. The dimensions Responsiveness, Assurance, and Empathy, along with the total combined scale score, were significantly different within the

TABLE 5.6

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES
OF THIS STUDY'S BANKING SERVICES WITH
THIS STUDY'S AIRLINE SERVICES

Expectations

	BANKING			AIRLINE		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.15	0.70	200	6.39**	0.58	197
Reliability	6.69	0.38	198	6.46**	0.48	200
Responsiveness	5.80	1.03	200	5.79	0.85	191
Assurance	6.58	0.43	195	6.44**	0.50	194
Empathy	5.67	0.87	196	5.76	0.87	199
Combined Scale	6.21	0.43	192	6.18	0.43	185

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.72	0.89	199	5.90*	0.87	198
Reliability	5.20	1.28	199	5.21	1.16	197
Responsiveness	4.31	1.48	198	5.07**	1.28	197
Assurance	5.26	1.23	197	5.53*	1.10	197
Empathy	4.22	1.35	197	4.95**	1.14	199
Combined Scale	4.93	1.08	192	5.35**	0.92	186

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.43	0.94	199	-0.49	0.92	195
Reliability	-1.46	1.31	197	-1.26**	1.23	197
Responsiveness	-0.50	1.69	198	-0.70**	1.44	188
Assurance	-1.31	1.22	192	-0.87**	1.09	191
Empathy	-0.42	1.55	194	-0.83**	1.36	197
Combined Scale	-1.27	1.10	186	-0.79**	0.96	175

*   Significant at .95
** Significant at .99

quality construct.

The significant differences that were apparent in the dimension Tangibles for both the expectations and perceptions constructs appeared to have cancelled each other out and caused no significant difference to be apparent in the quality construct. The difference in Reliability in the expectations construct appears to have been negated by an insignificant difference in the same dimension within the perceptions construct, so that there is no significant difference within the quality construct. Dimension scores that had only significant differences in the perceptions construct, these being Responsiveness and Empathy, carried that difference over into the quality construct. The same is true for the total combined scale score. Finally, significant differences apparent for the Assurance dimension in both the expectations and perceptions construct failed to cancel each other out, and combined to make this dimension significantly different for the quality construct as well.

5.1.3 HOTEL

A comparison of the expectation, perception, and quality scores obtained by this research for banking services with those obtained for hotel services is shown in Table 5.7. Student's two sided t-tests were used to test for significant differences within the mean dimension scores.

Between this study's banking results and hotel results, the dimensions Reliability, Assurance, and the total combined scale score

TABLE 5.7

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES
OF THIS STUDY'S BANKING SERVICES WITH
THIS STUDY'S HOTEL SERVICES

Expectations

	BANKING			HOTEL		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.15	0.70	200	6.15**	0.69	198
Reliability	6.68	0.38	198	6.43**	0.57	200
Responsiveness	5.80	1.03	200	5.76**	0.80	200
Assurance	6.58	0.43	195	6.38**	0.50	198
Empathy	5.67	0.87	196	5.71	0.83	200
Combined Scale	6.21	0.43	192	6.18*	0.45	196

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.72	0.89	199	5.49*	1.09	198
Reliability	5.20	1.28	199	5.26**	1.03	200
Responsiveness	4.31	1.48	198	5.01**	1.29	200
Assurance	5.26	1.23	197	5.36**	1.10	196
Empathy	4.22	1.35	197	4.05**	1.20	199
Combined Scale	4.93	1.08	192	5.24**	1.03	193

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.43	0.94	199	-0.66*	1.03	198
Reliability	-1.46	1.31	197	-1.16**	1.16	200
Responsiveness	-0.50	1.69	198	-0.75**	0.75	200
Assurance	-1.31	1.22	192	-1.00**	1.00	194
Empathy	-0.42	1.55	194	-0.67**	1.26	199
Combined Scale	-1.27	1.10	186	-0.85**	0.95	191

*   Significant at .95
** Significant at .99

exhibited significant differences within the expectations construct. For the perceptions construct, significant differences can be seen for the dimensions Tangibles, Responsiveness, and Empathy. The total combined scale score was also significantly different within this construct. All of the dimensions and the total combined scale score were significantly different within the quality construct.

Each of the dimensions was significantly different in one and only one of the two tested constructs, expectations and perceptions. Those differences were carried over in each case to the quality construct. The total scale score was significantly different in both the expectations and perceptions construct, however, these differences did not cancel each other out and the score remained significantly different for the quality construct.

5.1.4 RESTAURANT

Table 5.8 contains a comparison of the mean expectation, perception, and quality scores of this study's banking results with the results from the restaurant service category. Significant differences between the means were detected using Student's two sided t-tests.

Within the expectations section of the survey, significant differences exist for the dimensions Reliability, Responsiveness, Assurance, Empathy, and for the total combined scale score. The dimensions Responsiveness and Empathy are significantly different within the perceptions construct, as well as the total combined scale

TABLE 5.8

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES
OF THIS STUDY'S BANKING SERVICES WITH
THIS STUDY'S RESTAURANT SERVICES

Expectations

	BANKING			RESTAURANT		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.15	0.70	200	6.03**	0.69	195
Reliability	6.68	0.38	198	6.18**	0.63	200
Responsiveness	5.80	1.03	200	5.47**	1.07	198
Assurance	6.58	0.43	195	6.33**	0.50	197
Empathy	5.67	0.87	196	5.92**	0.92	200
Combined Scale	6.21	0.43	192	5.91**	0.53	193

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.72	0.89	199	5.69	1.01	198
Reliability	5.20	1.28	199	5.17**	1.10	195
Responsiveness	4.31	1.48	198	4.74**	1.37	198
Assurance	5.26	1.23	197	5.43**	1.10	197
Empathy	4.22	1.35	197	4.91**	1.23	198
Combined Scale	4.93	1.08	192	5.24**	0.99	187

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.43	0.94	199	-0.31**	1.10	194
Reliability	-1.46	1.31	197	-1.00**	1.21	195
Responsiveness	-0.50	1.69	198	-0.74**	1.55	196
Assurance	-1.31	1.22	192	-0.88**	1.12	190
Empathy	-0.42	1.55	194	-0.59**	1.31	198
Combined Scale	-1.27	1.10	186	-0.70**	1.03	182

* Significant at .95
** Significant at .99

score. All of the dimensions in the quality construct with the exception of Tangibles are significantly different from the results obtained by the banking service segment. In addition to the dimensions Reliability, Responsiveness, Empathy, and Assurance, the total combined scale score was also significantly different within the quality construct.

In this situation, the cancelling effect did not have any influence on the existence of significant differences within the quality construct. Any dimension that was significantly different within either the expectations construct or the perceptions construct, or both, was significantly different within the quality construct.

5.1.5 SKI AREA

A comparison of the mean dimension scores for the expectations, perceptions, and quality constructs between this study's banking services and ski area services is contained in Table 5.9. The mean scores for each dimension are compared using Student's two sided t-tests and significant differences reported.

The dimensions Tangibles, Reliability, Assurance, Empathy, and the total combined scale score are significantly different between the banking and ski area results within the expectations construct. Examining the perceptions construct, the dimensions Tangibles, Responsiveness, and Empathy are significantly different. Within the quality construct, the mean scores for the dimensions Reliability, Responsiveness, Assurance, Empathy, and the total combined scale score

TABLE 5.9

COMPARISON OF EXPECTATION, PERCEPTION AND QUALITY SCORES
OF THIS STUDY'S BANKING SERVICES WITH
THIS STUDY'S SKI AREA SERVICES

Expectations

	BANKING			SKI AREA		
	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	6.15	0.70	200	5.84**	0.76	200
Reliability	6.68	0.38	198	6.11**	0.55	196
Responsiveness	5.80	1.03	200	5.67**	0.89	199
Assurance	6.58	0.43	195	6.31**	0.48	195
Empathy	5.67	0.87	196	5.28**	0.90	198
Combined Scale	6.21	0.43	192	5.87**	0.50	188

Perceptions

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	5.72	0.89	199	5.41**	1.02	199
Reliability	5.20	1.28	199	5.11**	0.75	199
Responsiveness	4.31	1.48	198	4.79**	1.09	199
Assurance	5.26	1.23	197	5.36**	0.79	200
Empathy	4.22	1.35	197	4.66**	0.93	200
Combined Scale	4.93	1.08	192	5.08	0.73	197

Quality

	Mean	Std. Dev.	N	Mean	Std. Dev.	N
Tangibles	-0.433	0.941	199	-0.436**	1.185	199
Reliability	-1.458	1.307	197	-0.994**	0.741	195
Responsiveness	-0.500	1.691	198	-0.878**	1.229	199
Assurance	-1.308	1.221	192	-0.941**	0.799	195
Empathy	-0.424	1.550	194	-0.623**	1.020	198
Combined Scale	-1.265	1.096	186	-0.6773**	0.736	186

*   Significant at .95
** Significant at .99

are significantly different between samples.

Other than the dimension Tangibles, the cancelling effect makes no impression on the dimensions that are significantly different within the quality construct; these dimensions being Reliability, Responsiveness, Assurance, and Empathy. The total combined scale score is also not affected. The dimension Tangibles, however, is not significant because the significant differences that exist within both the constructs expectations and perceptions cancel each other out, leaving an insignificant difference within the quality construct.

5.1.6 IMPLICATIONS

This set of comparisons was designed to determine if the scale is able to distinguish between service segments that were quite dissimilar, those being banking services and tourism services. The scale did distinguish between the results for the banking service segment and those of the tourism service segments, as evidenced by the fact that many significant differences in mean scores occurred.

In comparing the results of the banking service segment and the tourism service segments, it is not expected that any of the constructs would perform consistently. Most importantly, to prove its ability to distinguish between services, it would be expected that expectation scores would vary to some significant degree. The perception scores and quality scores, which are a function of the perception scores, are also expected to vary between the service

segments, but that variance cannot be viewed as support for any proposed hypothesis.

The expectation scores of the banking services for this study did vary from the expectation scores of the tourism services in all of the dimensions and for the total scale score. These differences ranged from a high of four dimensions in the case of restaurant and ski area services to a low of two dimensions for hotel services. Airline services had significant differences in expectation scores from banking services for three dimensions.

Given the existence of these differences, it would be appropriate to reject the third hypothesis, that there would be no significant difference between this study's banking results and its tourism results. The scale seemed to have an ability to differentiate between tourism services and banking services, two types of services that are quite dissimilar in nature. This could also be further generalized to indicate that the scale appears to have the ability to distinguish between different service segments if the segments that are being compared are quite dissimilar.

It should be noted, however, that while the scale does appear to have the ability to differentiate between service segments, this does not imply that it is accomplishing the function that it has been designed for, measuring service quality. The ability to distinguish between different service segments is a desirable one, but this alone does not validate the conclusion that service quality is being measured by the scale. The question of what the scale is measuring, along with some other issues determined to be important, will be discussed in the next section.

5.2 THE PERFORMANCE AND FUNCTIONING OF THE SERVQUAL SCALE

Conclusions have been drawn about the applicability of the SERVQUAL scale developed by Parasuraman et al. (1986) to tourism service segments and the consistency of the results within a common service segment, banking. An examination will now be conducted into other factors more directed to the performance and functioning of the scale. This section will examine two aspects of the SERVQUAL scale and make conclusions based upon the results of these examinations.

5.2.1 NEGATIVELY WORDED DIMENSIONS

In developing the SERVQUAL scale, Parasuraman et al. (1986) made a decision to word approximately half of the statement sets negatively in order to conform to scale construction procedures outlined by Churchill (1979). Each of the statement sets in the Responsiveness and Empathy dimensions were worded negatively and the resulting scores reversed to be comparable with the positively worded statement scores.

A disturbing trend becomes evident when the scores for the negatively worded dimensions are examined and compared with those of the positively worded dimensions. For both the SERVQUAL study and for this study, the mean expectation and perception scores for the Responsiveness and Empathy dimensions are lower than for the Tangibles, Reliability, and Assurance dimensions. This is true across all service segments examined. The quality scores, since they are made up of the difference between perception and expectation scores,

do not exhibit this trait as the trend of lower scores for negatively worded statements appears to affect both components of the quality score construct.

The fact that the expectation and perception scores for the negatively worded dimensions were consistently lower than those of the positively worded dimension suggests two possible conclusions. The first possibility is that consumers' expectations of and perceptions about the quality of service they received were genuinely lower for the dimensions Responsiveness and Empathy than for the dimensions Tangibles, Reliability and Assurance. The second possible conclusion is that some of the respondents became confused or less likely to answer in the extreme when considering a negatively worded statement. If this were the case, the mean expectation and perception scores for the negatively worded dimensions would be lower than for the positively worded dimensions, and would not accurately represent their true expectations and perceptions of the quality of service.

In order to determine which of these conclusions seems more appropriate, an examination of the variance associated with each of the dimensions is desirable. This will determine if the variance of the mean expectation and perception scores is higher for the negatively worded dimensions than for the positively worded dimensions.

Snedecor and Cochran (1967) discuss a statistical test which may be used in determining whether non-independent variances are significantly different. In order to apply this test, however, an examination of the correlations between positively and negatively worded expectation and perception scores should be conducted to

determine if the assumption of non-independence is valid. Table 5.10 through Table 5.14 illustrate the correlations between the positively and negatively worded expectation and perception scores for each of the five service segments. Correlations significant at alpha equal to .95 and .99 are reported.

As can be seen in the tables, most of the correlations between positively and negatively worded expectation scores are significant, while all of the correlations for the perception scores are significant. These correlations are a validation of the assumption of non-independence of the expectation and perception scores, and supports the use of the comparison of variance test described by Snedecor and Cochran (1967).

Table 5.15 contains an examination of mean expectation score variance, and compares the variances of the negatively worded dimensions with those of the positively worded dimensions using this test for determining whether variances vary significantly. The superscripts attached to the negatively worded dimensions indicate variance levels significantly higher than the positively worded dimensions at an alpha equal to .95 .

For every service segment's expectations scores, the negatively worded dimensions' variances were significantly higher than the variances of the positively worded dimensions. This suggests that the scale scores of the negatively worded dimensions were more varied than the scale scores of the positively worded dimensions. This result lends credibility to the statements made by many of the respondents interviewed after completing the survey who felt that the negatively worded statements were confusing and hard to understand.

TABLE 5.10

CORRELATION BETWEEN POSITIVELY AND NEGATIVELY WORDED
EXPECTATION AND PERCEPTION SCORES
FOR BANKING SERVICES

Expectations

Positively Worded   **Negatively Worded**

	Responsiveness	Empathy
Tangibles	-0.029 (n=200)	0.238** (n=196)
Reliability	0.201** (n=198)	0.288** (n=195)
Assurance	0.201** (n=195)	0.223** (n=193)

Perceptions

Positively Worded   **Negatively Worded**

	Responsiveness	Empathy
Tangibles	0.485** (n=200)	0.475** (n=196)
Reliability	0.576** (n=198)	0.641** (n=195)
Assurance	0.611** (n=195)	0.647** (n=193

*   Significant at .95
** Significant at .99

TABLE 5.11

CORRELATION BETWEEN POSITIVELY AND NEGATIVELY WORDED
EXPECTATION AND PERCEPTION SCORES
FOR AIRLINE SERVICES

Expectations

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.105 (n=188)	0.272^{**} (n=196)
Reliability	0.159^{*} (n=191)	0.217^{**} (n=199)
Assurance	0.207^{**} (n=188)	0.162^{*} (n=193)

Perceptions

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.292^{**} (n=188)	0.311^{**} (n=196)
Reliability	0.551^{**} (n=191)	0.533^{**} (n=199)
Assurance	0.626^{**} (n=188)	0.550^{**} (n=193)

*   Significant at .95
** Significant at .99

TABLE 5.12

CORRELATION BETWEEN POSITIVELY AND NEGATIVELY WORDED EXPECTATION AND PERCEPTION SCORES FOR HOTEL SERVICES

Expectations

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.028 (n=198)	0.109 (n=198)
Reliability	0.294** (n=200)	0.319** (n=200)
Assurance	0.106 (n=198)	0.254** (n=198)

Perceptions

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.549** (n=198)	0.629** (n=198)
Reliability	0.769** (n=200)	0.769** (n=200)
Assurance	0.803** (n=198)	0.805** (n=198)

* Significant at .95
** Significant at .99

TABLE 5.13

CORRELATION BETWEEN POSITIVELY AND NEGATIVELY WORDED EXPECTATION AND PERCEPTION SCORES FOR RESTAURANT SERVICES

Expectations

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.275^{**} (n=194)	0.177^{**} (n=195)
Reliability	0.336^{**} (n=198)	0.275^{**} (n=200)
Assurance	0.318^{**} (n=196)	0.257^{**} (n=197)

Perceptions

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.353^{**} (n=194)	0.446^{**} (n=195)
Reliability	0.506^{**} (n=198)	0.612^{**} (n=200)
Assurance	0.609^{**} (n=196)	0.661^{**} (n=197)

*   Significant at .95
** Significant at .99

TABLE 5.14

CORRELATION BETWEEN POSITIVELY AND NEGATIVELY WORDED EXPECTATION AND PERCEPTION SCORES FOR SKI AREA SERVICES

Expectations

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.352^{**} (n=199)	0.216^{**} (n=198)
Reliability	0.464^{**} (n=195)	0.351^{**} (n=194)
Assurance	0.289^{**} (n=194)	0.229^{**} (n=193)

Perceptions

Positively Worded	Negatively Worded	
	Responsiveness	Empathy
Tangibles	0.320^{**} (n=199)	0.382^{**} (n=198)
Reliability	0.673^{**} (n=195)	0.623^{**} (n=194)
Assurance	0.565^{**} (n=194)	0.592^{**} (n=193)

*   Significant at .95
** Significant at .99

TABLE 5.15

COMPARISON OF VARIANCE OF EXPECTATION SCORES
BETWEEN POSITIVELY AND NEGATIVELY WORDED DIMENSIONS
FOR ALL SERVICE SEGMENTS

	Banking Variance	Airline Variance	Hotel Variance	Restaurant Variance	Ski Area Variance
Positively Worded					
Tangibles	0.49 (n=200)	0.34 (n=197)	0.47 (n=198)	0.48 (n=195)	0.57 (n=200)
Reliability	0.14 (n=198)	0.23 (n=200)	0.31 (n=200)	0.39 (n=200)	0.30 (n=196)
Assurance	0.18 (n=195)	0.25 (n=194)	0.25 (n=198)	0.25 (n=197)	0.23 (n=195)
Negatively Worded					
Responsiveness	$1.06^{a,b,c}$ (n=200)	$0.73^{a,b,c}$ (n=191)	$0.65^{a,b,c}$ (n=200)	$1.15^{a,b,c}$ (n=198)	$0.79^{a,b,c}$ (n=199)
Empathy	$0.75^{a,b,c}$ (n=196)	$0.76^{a,b,c}$ (n=199)	$0.69^{a,b,c}$ (n=200)	$0.84^{a,b,c}$ (n=200)	$0.81^{a,b,c}$ (n=198)

a = significantly higher than Tangibles at .95
b = significantly higher than Reliability at .95
c = significantly higher than Assurance at .95

Given this confusion, it is reasonable to state that the expectation scores derived for the negatively worded dimensions are not a true representation of the expectations of the respondents for these dimensions.

An examination of the mean perception score variances is contained in Table 5.16. Again, variances significantly higher at alpha equal to .95 are indicated by superscripts. The variance associated with negatively worded dimensions is significantly higher than for the positively worded dimensions for many of the comparisons. Although the effect is not as pronounced as was observed for the mean expectation score variances, it is evident that the same problem with confusion exists to some degree for the measurement of perceptions about quality services. Again, it is reasonable to state that the perception scores derived for the negatively worded dimensions are not directly comparable to the positively worded statements and are not a true representation of the perceptions of the respondents for these dimensions.

The existence of this confusion as a result of the negatively worded statements casts doubt onto one of the conclusions made by Parasuraman et al. (1986). They suggested that the order of importance of the expectations could be interpreted from the relative mean scores for each of the dimensions when compared with each other. This is quite significant, as the concept of differing importance for dimensions is one not addressed by the scale. Unfortunately, the problems resulting from the confusion surrounding the negatively worded questions serve to invalidate any interpretations that may be made about the relative importance of the different dimensions.

TABLE 5.16

COMPARISON OF VARIANCE OF PERCEPTION SCORES
BETWEEN POSITIVELY AND NEGATIVELY WORDED DIMENSIONS
FOR ALL SERVICE SEGMENTS

	Banking Variance	Airline Variance	Hotel Variance	Restaurant Variance	Ski Area Variance
Positively Worded					
Tangibles	0.80 (n=199)	0.76 (n=198)	1.18 (n=198)	1.02 (n=198)	1.04 (n=199)
Reliability	1.65 (n=199)	1.34 (n=197)	1.07 (n=200)	1.22 (n=195)	0.56 (n=199)
Assurance	1.51 (n=197)	1.20 (n=197)	1.20 (n=196)	1.21 (n=193)	0.62 (n=200)
Negatively Worded					
Responsiveness	$2.20^{a,b,c}$ (n=198)	$1.65^{a,c}$ (n=197)	$1.65^{a,b,c}$ (n=200)	$1.87^{a,b,c}$ (n=198)	$1.19^{b,c}$ (n=199)
Empathy	1.83^{a} (n=197)	1.29^{a} (n=197)	1.44^{b} (n=199)	1.52^{a} (n=198)	$0.86^{b,c}$ (n=200)

a - significantly higher than Tangibles at .95
b - significantly higher than Reliability at .95
c - significantly higher than Assurance at .95

5.2.2 THE QUALITY CALCULATION

The major assumption made by Parasuraman et al. (1986) in the development of the SERVQUAL scale was that the quality of a service could be quantified by subtracting the survey's expectation dimension scores from its perception dimension scores. This assumption that the degree to which perceptions exceeded expectations constitute quality is one that, for the purposes of testing the scale, was accepted without qualification.

Upon examining the functioning of the scale, the inadequacies of this assumption become apparent. In the case of both the SERVQUAL research and this research, respondents were asked to respond on a 7 point Lickert scale the degree to which they agreed or disagreed with a statement. The quality of the service was calculated from the difference between the perceptions and expectations, and usually resulted in a relatively small negative amount. The greater, or less negative, the number, the higher the quality of service.

This is not a reasonable approach due to the fact that should a respondent have high expectations of a dimension, when that expectation is subtracted from the perception score, the resultant quality score would be artificially restricted. Inversely, should a respondent have low expectations of a particular dimension, the smaller score to be subtracted from the perception score would allow for a artificially higher quality score.

To clarify further, if a respondent chose 7 on a 7 point scale to represent the fact that he strongly agreed with an expectation statement, because the highest possible score achievable within the

corresponding perception statement is 7, the highest possible quality score obtainable would be 0.

PERCEPTION - EXPECTATION = QUALITY

7 - 7 = 0

Thus, if a respondent had high expectations of a dimension and answered at the extreme high end of the scale for its expectation measuring items, there was a ceiling of 0 on the quality score that could be obtained. This ceiling would be in effect regardless of how superior the perception of the service would be.

Similarly, should a respondent choose 1 on the seven point scale to reflect that his expectations were not high for that dimension, even a moderate perception score would result in an artificially high quality score.

PERCEPTION - EXPECTATION = QUALITY

4 - 1 = 3

Thus, if a respondent had low expectations of a dimension and answered at the extreme low end of the scale for its expectation measuring items, the computed quality score could be easily inflated even with a moderate perception score.

The assumption that quality is made up of the difference between perception and expectation scores has the effect of placing a ceiling on dimensions with high expectations and artificially inflating dimensions with low expectations. It is for this reason that

attempting to measure quality as a difference between the two is not feasible and does not reflect accurately the actual perceived quality of the service. The next section examining where the quality of the service is measured will explore this contention further.

5.2.3 THE QUALITY MEASUREMENT

The previous section documented the existence of ceilings and the possibility of artificially inflating the service quality scores of dimensions having extremely high or low expectations on the part of the service users. Experience with the scale has suggested that the items categorized as measuring perceptions of the service are actually measuring the performance of the firm on those items. In effect then, the scale items purported to be measuring perceptions may be measuring the actual quality of the service. Given this, it would be expected that the correlations between perception scores and the question asking respondents directly to evaluate the quality of the service would be higher than the correlations between the calculated quality of service score and the quality of service evaluation.

An examination of the correlations between both the computed quality score and the perception construct score with the quality measuring evaluation for the banking service segment is contained in Table 5.17. Pearson product-moment correlation coefficients are reported for each of the five dimensions.

For the banking service segment, all 5 dimensions have stronger correlations between the perception construct scores and the quality

TABLE 5.17

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN COMPUTED QUALITY SCORES AND PERCEPTIONS WITH ACTUAL QUALITY EVALUATIONS FOR BANKING SERVICES

DIMENSIONS	COMPUTED QUALITY	PERCEPTIONS
Tangibles	0.334 (n=197)	0.466 (n=197)
Reliability	0.668 (n=195)	0.693 (n=197)
Responsiveness	0.520 (n=196)	0.596 (n=196)
Assurance	0.639 (n=190)	0.675 (n=195)
Empathy	0.557 (n=192)	0.594 (n=195)

evaluations than between the calculated quality scores and the quality evaluations. No test for significant differences between correlation coefficients is possible, so only the existence of a difference and the magnitude and direction of that difference may be discussed.

Table 5.18 contains a similar analysis of the correlations for the airline service segment. Of the five dimensions, three result in stronger correlations between the perception scores and actual quality evaluations. These three dimensions are Reliability, Responsiveness, and Assurance. Two of the dimensions, Tangibles and Empathy, have stronger relationships between the computed quality scores and the quality evaluations, although the difference for Empathy is very slight.

The same result is apparent when examining the hotel service segment as illustrated in Table 5.19. The correlations between the calculated quality scores and the quality evaluations is lower in every dimension than the correlation between the perception construct scores and the quality evaluations.

Table 5.20 contains an examination of these correlations for the restaurant service segment. Only the dimension of Tangibles had a stronger correlation between the computed quality scores and the quality evaluations than between the perception construct scores and the quality evaluations. The dimensions Reliability, Responsiveness, Assurance and Empathy all had stronger correlations between the perception construct scores and the quality evaluations.

An examination of the relevant correlation coefficients for the ski area service segment is contained in Table 5.21. All the dimensions had stronger correlations between the perception construct

TABLE 5.18

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN
COMPUTED QUALITY SCORES AND PERCEPTIONS WITH
ACTUAL QUALITY EVALUATIONS
FOR AIRLINE SERVICES

DIMENSIONS	COMPUTED QUALITY	PERCEPTIONS
Tangibles	0.492 (n=189)	0.397 (n=192)
Reliability	0.702 (n=191)	0.723 (n=191)
Responsiveness	0.479 (n=184)	0.517 (n=191)
Assurance	0.690 (n=185)	0.761 (n=191)
Empathy	0.476 (n=191)	0.471 (n=191)

TABLE 5.19

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN
COMPUTED QUALITY SCORES AND PERCEPTIONS WITH
ACTUAL QUALITY EVALUATIONS
FOR HOTEL SERVICES

DIMENSIONS	COMPUTED QUALITY	PERCEPTIONS
Tangibles	0.603 (n=197)	0.700 (n=197)
Reliability	0.622 (n=199)	0.673 (n=199)
Responsiveness	0.584 (n=199)	0.639 (n=199)
Assurance	0.639 (n=193)	0.708 (n=195)
Empathy	0.566 (n=198)	0.682 (n=198)

TABLE 5.20

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN
COMPUTED QUALITY SCORES AND PERCEPTIONS WITH
ACTUAL QUALITY EVALUATIONS
FOR RESTAURANT SERVICES

DIMENSIONS	COMPUTED QUALITY	PERCEPTIONS
Tangibles	0.566 (n=185)	0.564 (n=189)
Reliability	0.677 (n=186)	0.736 (n=186)
Responsiveness	0.496 (n=187)	0.627 (n=189)
Assurance	0.728 (n=181)	0.811 (n=184)
Empathy	0.536 (n=189)	0.610 (n=189)

TABLE 5.21

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN COMPUTED QUALITY SCORES AND PERCEPTIONS WITH ACTUAL QUALITY EVALUATIONS FOR SKI AREA SERVICES

DIMENSIONS	COMPUTED QUALITY	PERCEPTIONS
Tangibles	0.328 (n=196)	0.493 (n=196)
Reliability	0.527 (n=192)	0.576 (n=196)
Responsiveness	0.468 (n=196)	0.572 (n=196)
Assurance	0.571 (n=192)	0.661 (n=197)
Empathy	0.515 (n=195)	0.535 (n=197)

scores and the quality evaluations than between the calculated quality scores and the quality evaluations.

Overwhelmingly, the perception construct scores are more highly correlated with the actual evaluation of quality service received than the calculated quality scores. This suggests the conclusion that the expectation construct is certainly not the best instrument to measure consumers' evaluations of service quality. It could also be suggested that the correlations between the computed quality of service and the quality of service evaluation is due almost entirely to the perceptions construct.

An initial reaction to these suggestions is the conclusion that expectations have nothing to do with the evaluation of the quality of service. Many researchers (Sasser et al. (1978), Gronroos (1978a), Lehtinen and Lehtinen (1982), and Parasuraman et al. (1984)) however, have identified service quality as the result of a comparison between the expectations of a service they will receive and the performance of the firm providing the service. More likely, then, is the conclusion that the SERVQUAL perceptions construct is actually measuring the perceived quality of service, after consumers have gone through a process of expectation and perception evaluation. This conclusion is one strongly suggested and supported by the analysis of correlations previously discussed.

5.2.4 SATISFACTION WITH THE SURVEY

The survey instrument included a question designed to examine how satisfied the respondents were that the survey had captured their

perceptions and expectations regarding quality of service. The results of this question were presented in Table 4.8.

In order to determine if there was any difference in this satisfaction level between those respondents completing the banking survey and those completing the tourism segment surveys, an analysis comparing these two segments was conducted. The results of that analysis are presented in Table 5.22.

For those respondents considering banking services, 12.9% characterized the survey instrument as excellent, 59.8% felt it was good, and 25.3% felt it was fair. Only 2.1% of the respondents felt the instrument was a poor one.

The survey instrument was characterized as excellent by 5.1% of those respondents considering tourism services, while 57.7% felt it was good. It was characterized as fair by 32.8% of the respondents and poor by 4.4%.

When a Chi-Square statistic was calculated on the cross-tabulation of these variables, a value of 19.4 with 3 degrees of freedom, was obtained. This significant value indicates that there is a difference with the satisfaction level with the survey between those respondents considering banking services and those considering tourism services.

Given the existence of this significant difference, it is also useful to note that the trend is for the respondents considering banking services to answer excellent or good more often than those respondents considering tourism services. Alternately, those tourism respondents, when asked their satisfaction with the survey, answered fair or poor more often than banking respondents. The conclusion that

TABLE 5.22

COMPARISON OF SATISFACTION WITH SURVEY INSTRUMENT
BETWEEN BANKING AND TOURISM RESPONDENTS

	Banking Survey (n=194)	Tourism Surveys (n=771)
Satisfaction Response		
Excellent	12.9%	5.1%
Good	59.8%	57.7%
Fair	25.3%	32.8%
Poor	2.1%	4.4%

Note: Totals may not equal 100.0% due to rounding.

can be drawn from this trend is that the survey was felt to be a better instrument by those respondents considering banking services than it was by those respondents considering tourism services.

CHAPTER 6

CONCLUDING DISCUSSION

6.0 SUMMARY

The overall purpose of the research was to determine the extent to which the scale developed by Parasuraman et al. (1986), developed to measure service quality across a wide range of service categories, can be used to measure and evaluate service quality within selected segments of the tourism industry. The first stage in making this determination was the investigation of the scale to determine if it performed consistently within a common service segment and had the ability to distinguish between dissimilar service segments.

6.0.1 THE RESULTS

The SERVQUAL scale did not perform consistently within the banking service segment for the two research projects. Several possible reasons for this inconsistency were illustrated, but the first hypothesis, predicting consistency within a common service segment, was rejected.

When the results for the tourism service segments were compared to the SERVQUAL banking results and this study's banking results, differences in the mean expectation score results were found in both cases. This result supported a rejection of the second and third hypotheses, which predicted that no differences would be found.

The resultant conclusion from the analysis of the hypotheses proposed is that the scale did not perform consistently within common service segments, but did appear to have the ability to distinguish

between different service segments. During the course of the research project, however, several problems with the SERVQUAL instrument did become apparent. These problems were investigated within the examination of results.

6.0.2 NEGATIVELY WORDED DIMENSIONS

The developers of the SERVQUAL scale decided to word approximately half of the statement sets designed to measure expectations and perceptions negatively, in order to conform to scale construction procedures outlined by Churchill (1979). This study used the same wording for the same statement sets in order to maintain consistency. For every service segment's expectation scores, the negatively worded dimensions' expectation scores were lower than those of the positively worded dimensions' expectation scores. In addition, the variances associated with the negatively worded dimensions were higher than those of the positively worded dimensions.

The conclusion that can be drawn from these observations is that the expectation and perception scores from the negatively worded statement sets are not directly comparable to those of the positively worded dimensions, and that it is likely that the scores of the negatively worded dimensions are not a true representation of the expectations and perceptions consumers have of these dimensions. This may also be true of the positively worded dimensions, but there is no way available to determine that with the collected data. Finally, Parasuraman et al. (1986) suggested that the importance of the

dimensions could be interpreted from the results of the expectations dimensions. That is, the dimension with the highest mean expectation score could be considered the most important, while the dimension with the lowest mean expectation score could be considered the least important. This could be an extremely useful conclusion, since the issue of importance is not addressed at all in the SERVQUAL scale. Unfortunately, due to the problems resultant from the negatively worded dimensions, such an interpretation no longer seems appropriate.

6.0.3 THE QUALITY CALCULATION

The main premise of the SERVQUAL scale is that the perceived quality of a service can be calculated by subtracting the expectation score of an item from the perception score of that item. The resulting amount would be a quality score. The major problem with this assumption is the notion that quality scores may be artificially inflated or have ceilings placed on them as a result of very low or very high expectation scores.

With a very low expectation score, a dimension scoring only moderately well in the perception measuring item will have an artificially high quality score. Should a very high expectation score be observed, a ceiling is placed on the possible quality score. This calls into question the very premise of the scale and leads to the possible conclusion that the quality calculation within SERVQUAL is not measuring the perceived quality of the service.

6.0.4 THE QUALITY MEASUREMENT

An examination instigated by the possibility of ceilings and artificially inflated quality scores led to an interesting discovery. It was determined that in almost every situation, the perception scores correlated more highly with the actual determination of the reported quality of service a consumer had received than the computed quality scores. It was then theorized that the correlation that existed between the quality scores and the reported quality of service was likely due almost entirely to the perceptions construct.

The conclusion that was drawn from these results is that the perception construct is actually a better measure of the perceived quality of service a consumer has received. The consumer is thus responding to the items measuring 'perceptions' after he has undergone his own internal comparison of expectations and perceptions.

It should then be determined what, if anything, the expectations construct is measuring. Although little subjective evidence exists at this point, it is theorized here that the items purported to be measuring 'expectations' are actually measuring some sort of importance associated with a particular dimension. It has been quite evident throughout the research that, while the SERVQUAL scale treats all the dimensions equally with respect to importance, this is not an accurate nor appropriate treatment.

Some dimensions are more important than others, and this importance can shift when evaluating different service segments. This is particularly evident in the tourism service segments, where different factors may be important in different situations. The same

is also true of services in general; it is not possible to assume that the importance of different dimensions will remain consistent across all service segments.

6.0.5 SATISFACTION WITH THE SURVEY

After completing their survey, respondents were asked their overall satisfaction level with that survey, in order to determine if there would be different satisfaction levels evident among respondents considering different service categories. Respondents expressing dissatisfaction with the instrument could feel that the instrument did not do an adequate job capturing their expectations and perception regarding their service experiences.

An examination of these results showed that banking service respondents tended to have higher levels of satisfaction with the survey instrument than tourism service respondents. Given the problems with the instrument already discussed, this discrepancy could suggest the conclusion that the scale is not as suited for tourism services as it is for the service segment banking, and more generally, for the service segments in which it was developed.

6.1 LIMITATIONS OF THE RESEARCH

There were several limitations of the research that became apparent as the project proceeded. The choice of banking as the common service segment between the SERVQUAL research and this research may have been unfortunate, as there was increasing focus on and awareness of banking operations by the Canadian Government during the data collection stages. It is also apparent that there are considerable differences between the banking systems of Canada and the United States, where the SERVQUAL research was undertaken. These factors may have caused some shift or discrepancy in the expectations and perceptions of the banking industry.

There is also a time difference of 2 to 3 years between the collation of the SERVQUAL results and this research's results. This time difference may have had an influence on the results due to the possibility of shifts in consumer values, thus casting some doubts as to the validity of direct comparisons between the two studies.

Finally, the lack of a specific question on the survey instrument asking how often people had used a particular service and service provider caused some problems in the use of the questions soliciting input as to whether a problem had ever been reported by the consumer and whether he would recommend the service provider to acquaintances. The lack of such frequency of use questions restricted the interpretation of the results of these questions to simply reporting the results. Any further interpretation of these results could have been misleading.

6.2 FURTHER RESEARCH

While the SERVQUAL scale does not appear to be a good instrument for measuring perceived service quality within the tourism industry and is questionable for service industries in general, it does provide a good basis for the determination of what dimensions consumers use in evaluating service quality. The research undertaken in the development of the scale provides a strong foundation for the continuation of research in this area.

Any attempt to determine the dimensions of perceived quality of service or to measure this quality must be based upon a fundamentally sound definition of perceived quality of service. The literature is quite clear on the opinion that perceived quality of service is the difference between a consumer's expectations and perceptions with regard to a service provider's performance. The current research focused on this notion of perceived quality of service, but questioned whether SERVQUAL's expectations and perceptions constructs were measuring these personal evaluations. The next stage in refining the SERVQUAL instrument should be research into a reliable and accurate method of determining a consumer's expectations and perceptions with regard to a service transaction. A successful research attempt in this area will allow an accurate determination of perceived quality of service for any dimensions that are proposed.

This research also questioned the use of both positively and negatively worded statement sets, and demonstrated that the responses to these statements may not accurately reflect the actual expectations and perceptions (as defined by SERVQUAL's developers) of respondents.

The question of whether both positively and negatively worded statements sets should be used within the same instrument and whether they should be applied consistently, as was the case with SERVQUAL, or in some random fashion, is one that further research should attempt to address. Another issue that arises from this point includes an attempt to determine the optimal method of wording statements requiring a respondent's evaluation in order to most accurately reflect reality. In addition, should it be determined that positively and negatively worded segment sets should not be used together, which type of statement sets, if either, provide the best representation of a consumer's feelings.

An additional issue that should be addressed in any attempt to measure perceived quality of service is one on which the literature is not too clear. It has been generally accepted, both within the literature and by the SERVQUAL developers, that <u>satisfaction</u> with a service experience is based upon a single iteration of the experience; <u>perceived service quality</u> is the result of repeated satisfaction determinations by a consumer and is a more enduring feeling. However, the scale specifically requires the respondent to consider an individual experience within a service category and respond to expectation and perception measuring statement sets. According to the definition of satisfaction, then, the scale is actually measuring satisfaction with the service experience and not explicitly measuring perceived quality of service. Further research should attempt to clarify the relationship between satisfaction and perceived quality of service in order to determine whether, as the SERVQUAL scale does, the equating of satisfaction with perceived quality of service is a valid

assumption. While examining this relationship, some consideration should also be given to the frequency of usage with respect to individual services. Frequently used services may have a different satisfaction level/perceived quality of service relationship than services that are used with a lower frequency. It might also be that cues, or other proxy indicators for service quality, are more important for services a consumer uses less frequently.

The issue of relative importance of dimensions is also one that is not well addressed, both within the service marketing literature and the SERVQUAL research. It is very important to first determine whether the importance of dimensions has an effect on the perceived quality of service, and how that importance level should be incorporated into the service quality measurement and the aggregate model of service quality. It is an intuitively appealing notion that service dimensions that are not perceived as being important relative to other dimensions should not be considered as heavily in the final service quality measurement; items that are considered important should be more heavily weighted. Another issue that must be considered is the construction of importance measuring items; the same consideration that goes into determining expectation and perception measuring items must be given to the determination of these importance measuring items.

It is not obvious that the SERVQUAL scale, which was an attempt to develop a perceived quality of service measuring instrument that was applicable across a wide variety of service categories, was itself developed using a representative group of service industries. Now that a series of dimensions have been extracted that appear to work

well within the service categories for which it was developed, application of the scale to other service categories should be undertaken to determine SERVQUAL's effectiveness. This research was an attempt to make that determination in a category of services generally considered as belonging to the tourism industry, and the scale did not function as well as it did for these other service categories. The scale should be tested in other new service segments in order to determine if the results apparent in this research are the result of some unique characteristics of tourism services or are the result of the scale not being as generally applicable as theorized. Should this be the case, and the results indicate that the SERVQUAL scale is not generally applicable to a degree that could be deemed desirable, examination of a wider variety of service categories might facilitate further development of the scale to broaden its application. Without a doubt, the SERVQUAL scale is a valuable contribution to the area of service quality; what must be determined by further research is how it can be improved to become a more valuable and applicable instrument.

The SERVQUAL scale is not a good instrument for measuring perceived service quality specifically within the tourism industry, but it is a good starting point for the determination of the dimensions of service quality for tourism services. Further research should specifically attempt to evaluate dimensions already identified and extract additional appropriate dimensions. Consideration must also be given to the points previously mentioned; concerns dealing with expectation and perception measurements, methodological concerns such as the use of negatively worded statement sets, satisfaction and

perceived quality relationships, and importance determination and measurement must all be considered for the tourism subset of services. The results of this research should also be considered in the light of general service categories, to determine if they can contribute to the further development of a general service quality measurement instrument.

REFERENCES

Berry, Leonard L. (1980) "Service Marketing is Different" <u>Business</u> 30 (May-June) 24-29.

Booms, Bernard H. and Mary J. Bitner (1981) "Marketing Strategies and Organization Structures for Service Firms" in J. Donnelly and W. George (eds) <u>Marketing of Services: 1981 Special Educators' Conference Proceedings</u> (Chicago: American Marketing Association).

---------- and Jody Nyquist (1981) "Analyzing the Customer/Firm Communication Component of the Services Marketing Mix" in J. Donnelly and W. George (eds) <u>Marketing of Services: 1981 Special Educators' Conference Proceedings</u> (Chicago: American Marketing Association).

Burkart, A. J. (1981) "Tourism: A Service Industry?" <u>International Journal of Tourism Management</u> 1 (March) 2-3.

Buzzell, Robert D. and Frederick D. Weirsema (1981) "Successful Share Building Strategies" <u>Harvard Business Review</u> (January-February) 135-144.

Carmen, James M. and Eric Langeard (1980) "Growth Strategies of Service Firms" <u>Strategic Management Journal</u> 1 (Jan-March) 7-22.

Churchill, Gilbert A. Jr. (1975) "A Paradigm for Developing Better Measures of Marketing Constructs" Journal of Marketing Research 16 (Feb) 64-73.

---------- and C. Suprenault (1982) "An Investigation into the Determinants of Customer Satisfaction" Working Paper Series, University of Wisconsin-Madison.

Crosby, Philip B. (1979) Quality is Free: The Art of Making Quality Certain (New York: The American Library Inc.).

Ebrahimpour, Maling (1985) "An Examination of Quality Measurement in Japan: Implications for Management in the United States" Journal of Operations Management 5 (August) 419-431.

Eiglier, Pierre E., E. Langeard, C. H. Lovelock, J. E. G. Brown and R. F. Young (1977) Marketing Consumer Services: New Insights (Cambridge, Ma.: Marketing Science Institute).

Engel, James F., David T. Kollat and Roger D. Blackwell (1973) Consumer Behavior 2nd ed. (New York: Holt, Rinehart and Winston).

Etherington, Lois and Turgut Var (1984) "Establishing a Measure of Airline Preference for Business and Non-Business Travellers" Journal of Travel Research 22 (Spring) 22-27.

Garvin, David A. (1983) "Quality on the Line" Harvard Business Review (Sept-Oct) 65-73.

Graefe, Alan R. and Terry J. Vaske (1987) "A Framework for Managing Quality in the Tourist Experience" Annals of Tourism Research 14 390-404.

Gramann, J. H. (1982) "Toward a Behavioral Theory of Crowding in Outdoor Recreation: An Evaluation and Synthesis of Research Leisure Sciences 5 109-126.

Gronroos, Christian (1978) "A Service Oriented Approach to the Marketing of Services" European Journal of Marketing 12 (No. 8) 588-601.

---------- (1982a) Strategic Management and Marketing in the Service Sector (Helsingfors: Swedish School of Economics and Business Administration).

---------- (1982b) "An Applied Service Marketing Theory" European Journal of Marketing Vol. 16 30-41.

Hall, J. (1974) "The Capacity to Absorb Tourists" Built Environment 8 (3) 392-397.

Holbrook, Morris B. and Kim D. Corfman (1985) "Quality and Value in the Consumption Experience: Phaedrus Rides Again" in J. Jacoby and

J. Olson (eds) <u>Perceived Quality</u> (Lexington, Ma: Lexington Books) 31-57.

Howard, John and Jagdish Sheth (1969) <u>The Theory of Buyer Behavior</u> (New York: John Wiley and Sons).

Kaur, J. (1979) "Re-examining Tourist Capacity of Naintal" in K. Sinnhuber and F. Jung (eds) <u>Studies in the Geography of Tourism and Recreation</u> (Vienna: Weiner Geographic Schriften) 108-114.

Knisely, Gary (1979) "Listening to Consumer is Key to Consumer of Service Marketing" <u>Advertising Age</u> 50 (February 19) 54-60.

Langeard, Eric, John E. Bateson, Christopher H. Lovelock and Pierre E. Eiglier (1981) <u>Service Marketing: New Insights from Consumers and Managers</u> (Cambridge, Ma: Marketing Science Institute).

Lehtinen, Jarmo R. (1983) "Customer Oriented Service System" Service Management Institute Working Paper, Helsinki, Finland.

Lehtinen, Uolevi and Jarmo R. Lehtinen (1982) "Service Quality: A Study of Quality Dimensions" unpublished working paper, Service Management Institute, Helsinki, Finland.

Lewis, R. C. and B. H. Booms (1982) "The Marketing Aspects of Service Quality", in Leonard L. Berry, G. Lynn Shostack and Gregory D.

Upah (eds)   Emerging Perspectives on Service Marketing   (Chicago:
American Marketing Association).

Lovelock, C. H.   (1980)   "Towards a Classification of Services", in C.
W. Lamb and P. M. Dunne   Theoretical Developments in Marketing
(Chicago: American Marketing Association).

----------   (1981)   "Why Marketing Management Needs to be Different
for Services"   in J. H. Donnelly and W. R. George (eds) Marketing of
Services: 1981 Special Educators' Conference Proceedings   (Chicago:
American Marketing Association) 5-9.

----------   (1983)   "Classifying Services to Gain Strategic Insights"
Journal of Marketing   47 (Summer) 9-20.

Lucas, R. C.   (1979)   "Perceptions of Non-Motorized Recreational
Impacts" in Proceedings of Recreational Impact on Wildlands   (Seat-
tle: U.S.D.A. Forest Service)   24-31.

Makens, James C. and Raymond A. Marquardt   (1977)   "Consumer Percep-
tions Regarding First Class and Airline Seating"   Journal of Travel
Research 16 (Summer) 19-23.

McConnell, J. D.   (1968)   "Effect of Pricing on Perception of Product
Quality" Journal of Applied Psychology 52 (August) 300-303.

O'Brien, Terrence V., Nat de Gennaro and George W. Summers (1977) "Customer Perceptions of Product Attributes in the Airline Industry" Journal of Travel Research 15 (Winter) 9-13.

Olander, F. (1970) "The Influence of Price on the Consumer's Evaluation of Products" in B. Taylor and G. Wills (eds) Pricing Strategy (Brandon: Systems Press).

Oliver, Richard (1981) "Measurement and Evaluation of the Satisfaction Process in Retail Settings" Journal of Retailing 57:3 25-48.

Olshavsky, Richard W. (1985) "Perceived Quality in Consumer Decision Making: An Integrated Theoretical Perspective" in J. Jacoby and J. Olson (eds) Perceived Quality (Lexington, Ma: Lexington Books) 3-29.

Parasuraman, A., Valarie A. Zeithaml and Leonard L. Berry (1984) "A Conceptual Model of Service Quality and Its Implications For Future Research" Report No. 84-106, Marketing Science Institute, Cambridge, Ma.

----------, Valarie A. Zeithaml and Leonard L. Berry (1986) "SERVQUAL: A Multiple Item Scale For Measuring Customer Perceptions of Service Quality" Report No. 86-108, Marketing Science Institute, Cambridge, Ma.

Perry, Aaron and S. Thomas Friedman (1973) "Dimensional Analysis of Attitudes Toward Commercial Flying" _Journal of Applied Psychology_ 58, 388-390.

Phillips, Lynn W., Dae R. Chang and Robert D. Buzzell (1983) "Product Quality, Cost Position and Business Performance: A Test of Some Key Hypotheses" _Journal of Marketing_ 47 (Spring) 26-43.

Pizam, Abraham, Yoram Neuman and Arie Reighel (1978) "Dimensions of Tourist Satisfaction With a Destination Area" _Annals of Tourism Research_ 5 (July-September) 314-322.

Rabin, Joseph H. (1983) "Accent is on Quality in Consumer Services This Decade" _Marketing News_ (March 4) 12.

Rathmell, John M. (1974) _Marketing in the Service Sector_ (Cambridge, Ma: Winthrop).

Regan, William J. (1963) "The Service Revolution" _Journal of Marketing_ 47 (July) 57-62.

Ritchie, J. R. Brent., Everett E. Johnston and Vernon J. Jones (1980) "Competition, Fares and Fences - Perspective of the Air Traveller" _Journal of Travel Research_ 18 (Winter) 17-25.

Sasser, W. E., R. P. Olson and D. D. Wyckoff (1978) _Management of Service Operations_ (Boston: Allyn and Bacon).

Shostack, G. Lynn (1977) "Breaking Free From Product Marketing" Journal of Marketing 41 (April) 73-80.

Smith, Ruth A. and Michael J. Houston (1982) "Script Based Evaluations of Satisfaction with Services" in Leonard L. Berry, G. Lynn Shostack and Gregory D. Upah (eds) Emerging Perspectives on Service Marketing (Chicago: American Marketing Association) 59-62.

Snedecor, George W. and William G. Cochran (1967) Statistical Methods (Iowa: Iowa State University Press).

Stankey, G. H. (1973) Visitor Perception of Wilderness Recreation Carrying Capacity Report No. INT-102 (Ogden, Utah: U.S.D.A. Forest Service).

Swan, John E. and Linda J. Combs (1976) "Product Performance and Consumer Satisfaction: A New Concept" Journal of Marketing (April) 25-33.

Takeuchi, Kirotaka and John A. Quelch (1983) "Quality is More Than Making a Good Product" Harvard Business Review (July-August) 139-145.

Thomas, D. R. E. (1978) "Strategy is Different in Service Business" Harvard Business Review (July-August) 158-165.

Tull, D., R. A. Boring and M. H. Gonsoir (1964) "A Note on the Relationship of Price and Imputed Quality" Journal of Business 37 (April) 186-191.

Upah, Gregory D. (1980) "Mass Marketing in Service Retailing: A Review and Synthesis of Major Methods" Journal of Retailing 56 (Fall) 59-76.

Zeithaml, Valarie A. (1981) "How Consumer Evaluation Processes Differ Between Goods and Services" in J. H. Donnelly and W. R. George (eds) Marketing of Services: 1981 Special Educators' Conference Proceedings (Chicago: American Marketing Association) 186-190.

----------, A. Parasuraman and Leonard L. Berry (1985) "Problems and Strategies in Services Marketing" Journal of Marketing 49 (Spring) 33-46.

APPENDIX

<u>QUALITY OF SERVICE</u>

1) EXPECTATIONS

Directions: This survey deals with your opinions of services. Please show the extent to which you think firms offering _____ services should possess the features described by each statement. If you strongly agree that these firms should possess a feature, circle the number 7. If you strongly disagree that these firms should possess a feature, circle 1. If your feelings are not strong, circle a number in the middle. There are no right or wrong answers - all we are interested in is a number that best shows your expectations about firms offering services.

		Strongly Agree						Strongly Disagree
1.	Their employees should get adequate support from these firms to do their jobs well.	7	6	5	4	3	2	1
2.	Customers' dealings with these firms should be very pleasant.	7	6	5	4	3	2	1
3.	Their employees should be well dressed and appear neat.	7	6	5	4	3	2	1
4.	They should provide their services at the time they promise to do so.	7	6	5	4	3	2	1
5.	They should have up-to-date equipment.	7	6	5	4	3	2	1
6.	When customers have problems, these firms should be sympathetic and reassuring.	7	6	5	4	3	2	1
7.	Their employees should be knowledgable.	7	6	5	4	3	2	1
8.	It is unrealistic to expect employees to know what the needs of their customers are.	7	6	5	4	3	2	1
9.	They should keep their records accurately.	7	6	5	4	3	2	1
10.	They shouldn't be expected to tell customers exactly when services will be performed.	7	6	5	4	3	2	1

	Strongly Agree						Strongly Disagree

11. Customers should be able to feel safe in their transactions with these firms' employees.

7 6 5 4 3 2 1

12. It is unrealstic to expect these firms to have their customers' best interests at heart.

7 6 5 4 3 2 1

13. It is okay if they are too busy to respond to customer requests promptly.

7 6 5 4 3 2 1

14. When these firms promise to do something by a certain time, they should do it.

7 6 5 4 3 2 1

15. Customers should feel secure in their dealings with these firms.

7 6 5 4 3 2 1

16. It is not realistic for customers to expect prompt service from employees of these firms.

7 6 5 4 3 2 1

17. It is okay if customers have to wait a long time to receive their services.

7 6 5 4 3 2 1

18. The appearance of the physical facilities of these firms should be in keeping with the type of service provided.

7 6 5 4 3 2 1

19. Their physical facilities should be visually appealing.

7 6 5 4 3 2 1

20. These firms should be dependable.

7 6 5 4 3 2 1

21. These firms should not be expected to give customers individual attention.

7 6 5 4 3 2 1

22. Their employees don't always have to be willing to help customers.

7 6 5 4 3 2 1

23. Their employees should be polite.

7 6 5 4 3 2 1

24. Customers should be able to trust employees of these firms.

7 6 5 4 3 2 1

25. They shouldn't be expected to have operating hours convenient to all their customers.

7 6 5 4 3 2 1

26. Employees of these firms cannot be expected to give customers personal attention.

7 6 5 4 3 2 1

2) PERCEPTIONS

Directions: Consider now an organization in the industry that you have had a recent (within 3 months) experience with. For the purposes of the remainder of this survey, call this organization Company XYZ.

The following set of statements relates to your feelings about XYZ. For each statement, please show the extent to which you believe XYZ has the feature described by the statement. Once again, circling a 7 means that you strongly agree that XYZ has that feature, and circling a 1 means that you strongly disagree. You may circle any of the numbers in the middle that show how strong your feelings are. There are no right or wrong answers - all we are interested in is a number that best shows your perceptions about Company XYZ.

		Strongly Agree						Strongly Disagree
1.	Employees get adequate support from XYZ to do their job well.	7	6	5	4	3	2	1
2.	Your dealings with XYZ are very pleasant.	7	6	5	4	3	2	1
3.	XYZ's employees are well dressed and appear neat.	7	6	5	4	3	2	1
4.	XYZ provides its services at the time it promises to do so.	7	6	5	4	3	2	1
5.	XYZ has up-to-date equipment.	7	6	5	4	3	2	1
6.	When you have problems, XYZ is sympathetic and reassuring.	7	6	5	4	3	2	1
7.	Employees of XYZ are knowledgeable.	7	6	5	4	3	2	1
8.	Employees of XYZ do not know what your needs are.	7	6	5	4	3	2	1
9.	XYZ keeps its records accurately.	7	6	5	4	3	2	1
10.	XYZ does not tell customers exactly when services will be performed.	7	6	5	4	3	2	1
11.	You feel safe in your transactions with XYZ's employees.	7	6	5	4	3	2	1
12.	XYZ does not have your best interests at heart.	7	6	5	4	3	2	1
13.	Employees of XYZ are too busy to respond to customer requests promptly.	7	6	5	4	3	2	1

	Strongly Agree						Strongly Disagree
14. When XYZ promises to do something by a certain time, it does so.	7	6	5	4	3	2	1
15. You feel secure in your dealings with XYZ.	7	6	5	4	3	2	1
16. You do not receive prompt service from XYZ's employees.	7	6	5	4	3	2	1
17. You have to wait a long time to receive XYZ's services.	7	6	5	4	3	2	1
18. The appearance of the physical facilities of XYZ is in keeping with the type of services provided.	7	6	5	4	3	2	1
19. XYZ's physical facilities are visually appealing.	7	6	5	4	3	2	1
20. XYZ is dependble.	7	6	5	4	3	2	1
21. XYZ does not give you individual attention.	7	6	5	4	3	2	1
22. Employees of XYZ are not always willing to help customers.	7	6	5	4	3	2	1
23. Employees of XYZ are polite.	7	6	5	4	3	2	1
24. You can trust employees of XYZ.	7	6	5	4	3	2	1
25. XYZ does not have operating hours convenient to all of its customers.	7	6	5	4	3	2	1
26. Employees of XYZ do not give you personal attention.	7	6	5	4	3	2	1

For statistical purposes only, would you indicate the name of the organization you were thinking of when answering questions in this section (Company XYZ)?

_____

3)   OVERALL QUALITY

1.   Have you ever reported a problem with the services you have
     received with XYZ?

          Yes   No

2.   Would you recommend XYZ to a friend?

          Yes   No

3.   How would you rate XYZ's overall quality of service?

     Excellent   Good   Fair   Poor

Thank you for taking the time to fill out this survey. Your input
will be valuable in furthering knowledge in the relatively
unexplored topic area of service quality. There are just two more
questions we would like you to answer.

How good an instrument do you feel this survey is for capturing
your feelings and experiences in the area of quality of service?

     Excellent   Good   Fair   Poor

What suggestions could you make to improve this instrument to
better capture your feelings and experiences in the area of
quality service?

_____

_____

_____

AGAIN, THANK YOU FOR YOUR INPUT!