The Development and Psychometric Assessment of a Measure of Numerical Business Acumen

Jocelyn Anso

Submitted in partial fulfilment of
Master of Arts Degree (Social Sciences)
Massey University at Albany, 1999

Acknowledgments

I would like to express my thanks to Sheffield Consulting Group, and in particular, Anthony Mitchell and Lisa Barber, for their generous contributions of time, efforts, support, and resources. The original concept for this research was suggested by Anthony, and I received infinite assistance in its design and facilitation. I am also grateful to Sheffield Consulting Group for the introductions they provided to many of the organisations that took part. I have gained a great deal of invaluable knowledge from Sheffield Consulting Group in this consultative process, and have thoroughly enjoyed the opportunity to work with them on this project.

I would also like to thank all of the managers within the participating organisations for taking part in this research - their assistance greatly contributed to the smooth facilitation of this aspect of the project. Accordingly, I am immensely grateful to all those people who chose to participate in my research, as without their participation, this study would not have been possible.

Finally, I would like to thank my supervisor, Dr Philip Voss, for his unstinting willingness to give his time to provide guidance, feedback and assistance in this project. His never-failing positive attitude was great (even when I had a stupid question to ask!). I hope that our association has been as painless for him as it was for me, and I was happy to ensure that mutual laughter was a regular occurrence! Thanks, Philip - it's been a great experience.

Abstract

The competency Numerical Business Acumen has been identified as a managerial behaviour that is crucial for keeping pace with a dynamic business environment. Numerical Business Acumen was defined as the ability to interpret numerical data within the wider organisational and external context. The primary components of this construct were hypothesised to be numerical critical reasoning and business acumen. No commercially available test appears to adequately measure this construct. Therefore, the main aims of the present research were to develop a scenario-based measure of the Numerical Business Acumen competency, and to assess its psychometric properties. Two parallel versions of the Numerical Business Acumen measure and a comparative measure of critical reasoning were administered to a sample of 46 participants from seven organisations. The results showed that both the measure's internal consistency (r =.36) and stability over time (r = .27) were relatively poor. The Numerical Business Acumen measure did not correlate with the comparative measure used, indicating that a construct distinct from numerical critical reasoning was being assessed. This construct was tentatively identified as an aspect of business acumen, one which might include a type of context-specific critical reasoning based on experiential learning. Suggested improvements to the measure consisted of further refinement of the existing questions, and the inclusion of additional questions to measure other components of the Numerical Business Acumen competency.

Table of Contents

	Page
The competency-based approach	1
Definition of competency	4
Commonalities and trends in research	6
New Zealand research	10
The Numerical Business Acumen competency	11
Critical Reasoning	13
Definition of critical reasoning	13
Measurement of critical reasoning	14
Critical reasoning and work performance	14
Critical reasoning within context	16
Limitations of critical reasoning and the rational approach	19
Critical reasoning in action	20
Business Acumen	24
Tacit knowledge	25
Job experience and tacit knowledge	29
Practical intelligence	32
Numerical Business Acumen	35
What is Numerical Business Acumen?	35
Numerical Business Acumen and work performance	39
Is Numerical Business Acumen generic?	40
Measurement of Numerical Business Acumen	40
Numerical Business Acumen uses in Human Resources Management	41
Aims of the present research	42

Methodology 43

Participants 43

Procedure 44

Design of the Numerical Business Acumen measure 46

Results 57

Data entry and quality assurance 57

Scoring 58

Psychometric properties 60

Descriptive statistics 66

Demographic characteristics 68

Relationships with demographics variables 71

Response accuracy 73

Qualitative comments 75

Discussion 77

How reliable is the Numerical Business Acumen measure? 77

How valid is the Numerical Business Acumen measure? 86

Theoretical implications 92

Practical implications 96

Conclusion 104

Appendices

A. Information sheets

B. Consent form

C. Administrator guide

D. Scenario booklet for Numerical Business Acumen measure (Form A)

E. Question booklet for Numerical Business Acumen measure

F. Scenario booklet for Numerical Business Acumen measure (Form B)

References

List of Tables

Table	Title	Page
Table 1	*Behavioural indicators of Numerical Business Acumen*	38
Table 2	*Demographic details*	43
Table 3	*Summary of item analysis for all measures*	61
Table 4	*Summary of item analysis for individual scenarios in Numerical Business Acumen measure*	63
Table 5	*Correlations between the Numerical Business Acumen measures and NMG2 scores*	65
Table 6	*Correlations between self-rankings and all total scores*	66
Table 7	*Descriptive statistics for all measures*	67
Table 8	*A comparison of the mean Numerical Business Acumen & NMG2 scores between organisations in different industries*	69
Table 9	*A comparison of the mean Numerical Business Acumen & NMG2 scores between different job roles*	70
Table 10	*A comparison of the mean Numerical Business Acumen & NMG2 scores between managerial levels*	71
Table 11	*Correlations between demographic variables and total scores for all measures*	72
Table 12	*Regression coefficients and standardised regression coefficients for the prediction of NMG2 scores*	73
Table 13	*Proportion correct of items attempted for the Numerical Business Acumen and NMG2 measures*	74
Table 14	*Frequency with which various rationales were used in answering Question 4*	76

List of Figures

Figure	Title	Page
Figure 1	*The Iceberg Model*	4
Figure 2	*Numerical Business Acumen model*	36
Figure 3	*Scatterplot of Form A and Form B total scores*	64
Figure 4	*Frequency distribution for total scores on Numerical Business Acumen measure (Form A)*	67
Figure 5	*Frequency distribution for total scores on Numerical Business Acumen measure (Form B)*	67
Figure 6	*Frequency distribution for total scores on NMG2 test*	68
Figure 7	*Accuracy scores for Numerical Business Acumen measure (Form A)*	74
Figure 8	*Accuracy scores for NMG2 test*	75

In today's highly competitive global marketplace, an organisation's competitive advantage may no longer be sustained merely by demonstrating superiority in their products or services. Changes in the external environment, such as technological advances and growth in the service sector, have led to organisations becoming more reliant on their human resources as their main competitive advantage. Intangibles, such as specialised knowledge and responsiveness to ever changing customer needs, are becoming more highly valued (McLagan, 1997). Organisations have responded to these changes by becoming more flexible in their organisational structure and by encouraging employees to be more flexible in the way they view their work and their responsibilities. Consequently, as organisations become more flexible, managers require a broader range of skills to be able to function effectively (Lillibridge & Williams, 1992). Many organisations use a pre-defined framework, which indicates what personal skills, characteristics and abilities are considered necessary for effective managers, in any decisions regarding their selection, promotion and development (Burgoyne & Stuart, 1978). Therefore, the validity of the framework of managerial attributes used is likely to be a key factor in the accuracy of such decision-making.

The competency-based approach

The traditional approach of analysing the person and the job separately and then attempting to fit the two together has proved to be inadequate in predicting managerial performance (Spencer & Spencer, 1993). The constant change in managerial responsibilities has also meant that traditional job descriptions are no longer flexible enough and are quickly outdated (McLagan, 1997). The competency-based approach addresses these deficiencies by analysing jobs in terms of the characteristics and behaviours of the people who perform well in them, rather than taking the traditional job analysis approach of analysing the elements of the job itself (Lawler, 1994; Mitrani, Dalziel, & Fitt, 1992).

The competency-based approach emphasises criterion validity, that is, the factors that actually predict superior performance in a job, as opposed to factors which describe all of an individual's characteristics, in the hope that some of these factors will relate to job performance (Spencer & Spencer, 1993). Superior performance is typically defined as one standard deviation above average performance, which equates roughly to the top decile in a given working context (Spencer & Spencer). This emphasis on identifying behavioural indicators of superior performance also distinguishes the competency approach from person specifications because, although person specifications list the personal characteristics necessary to perform a job, only a minimally acceptable standard is aimed for (Cascio, 1998).

It is becoming increasingly clear that the characteristics of a successful manager are manifold, rather than simply based on academic intelligence. Indeed, the ability to function intelligently in real-life situations is equally, if not more, important to work performance (Klemp & McClelland, 1986). In line with this direction of thinking, Sternberg (1984, cited in Klemp & McClelland, 1986) defined intelligence as the characteristics leading to successful adaptation in a real-world context. Leading on from this, one way to increase an organisation's understanding of the adaptive characteristics required for superior performance is to compare individuals who have had varying levels of success in adapting to a given environment, in order to identify the characteristics that differentiate them (Klemp & McClelland).

In this way, a competency model is developed by systematically studying and comparing the performance of superior and average individuals within a specific job and a specific organisation, and identifying the personal characteristics that are associated with superior performance (Holland & Williams, 1994). This is typically carried out using some form of critical-incident interview, which involves the collection of incidents from supervisors, job incumbents or others familiar with the job, which demonstrate the personal characteristics of both superior and average performers (Ballantyne & Povah, 1995; Cascio, 1998; Mitrani, et al., 1992).

The systematic identification of 'superior-differentiators' produces more than just a list of desired knowledge, skills and attributes, because it also takes into account the organisation's unique challenges, culture, and environment. Thus, the competencies identified as associated with superior performance in one organisation may or may not be relevant to another organisation. Although the competency name and the general description may be similar (such as 'analytical thinking' or 'results-orientation'), the specific behavioural indicators would most likely vary from organisation to organisation (Holland & Williams, 1994).

The need for organisations to respond to the competitive marketplace effectively has led to a growing interest in the competency approach. The Management Charter Initiative (MCI) was introduced in the United Kingdom in the 1980s, with the objectives of promoting the concept of competency and facilitating the introduction of a list of generic management standards into British organisations (Wills, 1993). In New Zealand, the competency-based Skill New Zealand initiative was recently introduced by the government to encourage businesses to invest in ongoing training for their employees, with the dual objectives of producing multi-skilled workers and a more competitive global edge (Kerr, 1994).

Many organisations have recognised that the individual competencies of their 'human capital', such as innovation and passion for excellence, can provide a unique advantage in the marketplace (McLagan, 1997). Following on from this, an organisation's core competencies (the behaviours that an organisation strives to be the best at, such as innovation or outstanding customer service) are dependent on those individual competencies, and it is therefore essential that the two levels of competencies are aligned (Maccoby, 1990; Zingheim, Ledford, & Schuster, 1996).

As such, identifying the aspects of individual performance that are valued by the organisation is imperative, and should drive the underlying strategies of Human Resources practice. Once the competency dimensions of the various roles within an organisation have been identified, they can serve as

a template for characteristics to search for in recruitment and selection processes (Woodruffe, 1993a). Bethell-Fox (1992) suggested that a competency-based recruitment and selection process can ensure that the pre-defined characteristics selected for are those that will enable new employees to perform at the required level, and are sufficiently performance-related to reduce gender and ethnic bias. Competency profiles have also been linked to performance-pay (Zingheim, et al., 1996), and the identification of training and development needs (Woodruffe, 1993a). Competency-based development plans can ensure that an individual's abilities are developed and guided to meet the strategic goals of the organisation.

Definition of competency

Although the notion of competency is now widespread, there appears to be some confusion over how it is defined (Woodruffe, 1993b). The term 'competency' is often confused with the lay definition of competence as 'adequate, but not outstanding, performance' (Pomeroy, 1995). By contrast, a competency can be defined as "an underlying characteristic of an individual which is causally related to ... superior performance in a job" (Boyatzis, 1982, *p.* 23). These underlying characteristics can be motives, traits, skills, aspects of self-concept, or knowledge (Spencer & Spencer, 1993). These characteristics may exist within the individual at various levels: motives at the unconscious level, self-image at the conscious level, and skills at the behavioural level (Boyatzis, 1982).

The 'Iceberg Model' (which portrays the individual as an iceberg) is often used to demonstrate this concept. As illustrated in Figure 1, surface knowledge, such as technical and industry knowledge and skill competencies are at the tip of the 'iceberg', indicating behaviours that tend to be visible, and can be taught easily. The base of the iceberg is formed by the individual's motives and traits, values, and drives, which underlie behaviours which may be harder to assess and develop because they are more hidden, 'deeper' and central to personality. Self-concept competencies, such as self-confidence, fall

in both areas, because it is possible to change them over a period of time (Spencer & Spencer, 1993).

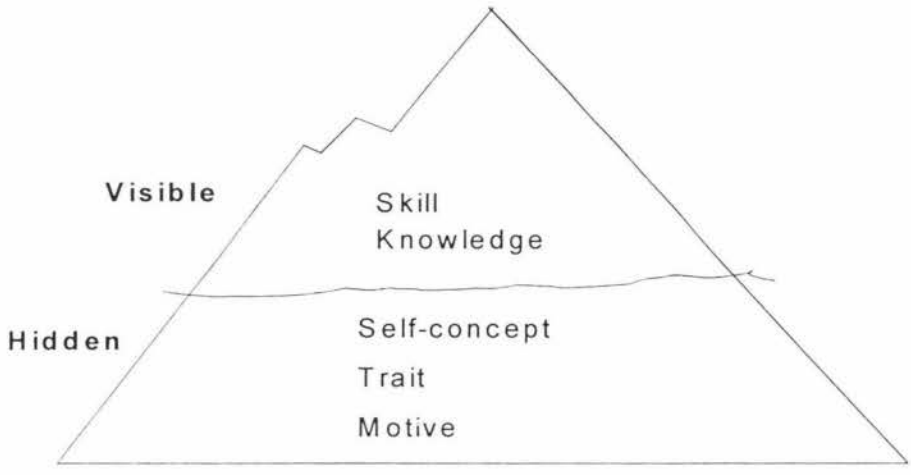


Figure 1
The Iceberg Model (Spencer & Spencer, 1993)

Thus, competencies may differ in the extent to which they can be learnt on the job. Knowledge and behavioural skills are relatively easy to develop, whereas attitudes and values are harder to alter. Core motive and trait competencies, which are central to personality, such as initiative and drive, are the most difficult to assess and develop (Spencer & Spencer, 1993). Yet it is these deeper competencies that can exert considerable influence over every-day behaviour, and they are therefore particularly important to identify for in the selection process. For example, poor leadership performance in executive management positions is seldom due to a lack of technical knowledge. More often than not, such inadequacy can be traced to the attitudes, personal characteristics, values, and motives those managers ascribe to when carrying out their tasks (Holland & Williams, 1994).

However, it would seem that, by defining competency in terms such a range of variables, Boyatzis (1982) created some confusion in what the term actually covers. Woodruffe (1993b) noted that the word *competency* seems to be used to cover any task, function or behaviour that might directly or indirectly affect job performance. Although Boyatzis's definition is the most comprehensive and not to be discounted, Woodruffe clarifies the issue by

defining *competency* simply as a discrete dimension of behaviour that is relevant to job performance. Still another definition of competency accommodates both the portability of behaviours from one situation to another, and the situation-specificity of other behaviours, by describing competency as "the ability to perform in a given context and the capacity to transfer knowledge and skills to new tasks and situations" (Wallace & Hunt, 1996, *p.* 38). This definition is the most up to date as it takes into account the portability of competencies in the changing world of work, as well as the necessary behavioural characteristics.

Although the relationship between competencies and superior performance is supposedly causal (Boyatzis, 1982), a set of competencies reflects only what an individual is capable of doing, not their willingness, nor the opportunity to perform. Thus, although a competency may integral to superior performance, it is also only one of several contributors (Pomeroy, 1995). The interaction of an individual's motives, the organisational environment, and the functional and situational demands of the job influences that individual's resulting work performance (Boyatzis, 1982).

Commonalties and trends in research

In addition to the inconsistencies in the definition of the term 'competency', there has also been disagreement about whether competencies are generic across industries, or are specific to an organisation. Because competencies are considered to be underlying characteristics, and are often rooted in intelligence and personality traits, they could be said to be generic (Boyatzis, 1982). Therefore, a generic characteristic may be apparent in many forms of behaviour, within many jobs in many organisations. Over the last 20 years, there has been extensive research investigating the relationship between managerial competencies and superior job performance, with an emphasis on discovering the competencies that are universal across different industries (Boyatzis, 1982; Burgoyne & Stuart, 1978; Klemp & McClelland, 1986; Spencer & Spencer, 1993; Wallace & Hunt, 1996; Wilkinson & Zwanenberg, 1994; Woodruffe, 1993b; Yukl, 1994; Zingheim, et al., 1996).

Initial studies focused on the administrative and intellectual skills required for effective managers (Klemp & McClelland, 1986; Sen & Das, 1991), with subsequent studies drawing attention to the interpersonal dimension that has emerged in modern management practice (Burgoyne & Stuart, 1978; Howard & Bray, 1988; cited in Sen & Das, 1991). Indeed, as organisations have become less reliant on centralised authority, interpersonal skills and competencies are increasingly being recognised as essential (Wallace & Hunt, 1996).

Boyatzis's (1982) study was a significant step in competency research with the formulation of an initial set of generic competencies that consistently distinguished superior managers. This set consisted of five competency clusters: Goal and action management, leadership, human resource management, focus on others, and directing subordinates. Boyatzis found that approximately one third of the variance in managers' performance could be accounted for by these generic competencies. The importance of job and organisation-specific competencies, such as specialised technical knowledge, and situational factors were acknowledged, with the remaining variance in performance being attributed to these factors. Hay/McBer (Spencer & Spencer, 1993) also developed a set of 21 generic competencies, resulting from 20 years of research, that appeared to differentiate superior from average performers across a range of occupational roles. These competencies fell into six main competency clusters: Achievement and action, helping and human service, influencing, managerial, cognitive, and personal effectiveness. These clusters provided a useful framework for understanding the broad range of managerial behaviours and activities across industries, and have been used extensively as a base for developing competency profiles in many organisations.

Competencies and management levels

Associated with the research on generic managerial competencies across organisations is the issue of whether these competencies are universal across entry, middle, and executive management levels, or are hierarchically specific. Thornton and Byham (1982) argued that some managerial competencies are

required at all levels with only the nature of the task and the degree of skill required changing. That is, as someone moves up the managerial hierarchy, higher levels of some competencies (such as strategic planning) and decreased levels of other competencies (such as technical knowledge), are required. However, the other side of the argument is that management skills do differ throughout the organisational hierarchy with different tasks, skills and leadership styles required. Clearly, there is some validity to both perspectives, because there will be some competencies that are required at all levels, but in differing degrees, and there may also be some competencies found predominantly in one managerial level.

Existing comparative research has established that there are differences between the managerial levels in terms of which competencies are relevant to each level, and how these competencies translate into different behaviours. Boyatzis (1982) found that an entry-level manager may use the competency 'developing others' to direct and guide his/her subordinates, whereas a middle-level manager may also use that competency, but to stimulate and develop the teams reporting to him/her. Thornton and Byham (1982) also noted that the consistent findings that certain communication patterns alter, and decision-making processes become more complex, at higher levels of management.

Ballantyne and Povah (1995) suggested that entry managers predominantly use technical knowledge, problem-solving and interpersonal skills (in a supervisory capacity); middle managers make greater use of decision-making skills, directive and collaborative skills (to achieve specific objectives through supervision of other groups), with less emphasis on technical skills; whereas executive managers predominantly use strategic planning and decision-making skills, using more external and organisational awareness abilities and almost no technical skills.

Current and future trends

Building on early research on managerial competencies which focused on informational and decisional skills (Sen & Das, 1991), the additional

dimensions of interpersonal skills, organisational and external awareness, and individual personality have been integrated into competency profiling only in the last decade. This reflects the changes which have taken place in managerial practice and in the external environment (Wilkinson & Zwanenberg, 1994; Woodruffe, 1993b; Zingheim et al., 1996). Recent research has included competencies such as flexibility, customer and market sensitivity (Wallace & Hunt, 1996; Wilkinson & Zwanenberg, 1994; Woodruffe, 1993b; Yukl, 1994; Zingheim et al., 1996). The value many organisations now place on their human resources as a source of competitive advantage (McLagan, 1997) is also reflected, with the development of employees, relationship building, and teamwork being commonly cited as important to managerial performance.

The importance of competencies being oriented towards the future has been noted by several researchers. Woodruffe (1993b) observed that competencies based on the past run the risk of missing changes required in response to organisational and environmental changes. This could be rectified by either including the competencies of 'changeability', such as sensitivity to external influences, and public awareness; or by ensuring that competency profiles are flexible enough to reflect any changes in the organisation's direction. Limerick (1990) identified a new emphasis on holistic and empathetic skills, such as intuition, maturity, and organisational awareness, being required for effective management performance. As organisations become more global, cross-cultural differences in how managers communicate with each other will also need to be addressed, with competencies such as teamwork and interpersonal sensitivity becoming important in dealings with Asian countries (Mo, 1986).

The generic competency debate

This focus on generic, as opposed to industry-specific, managerial competencies, has concerned several researchers (Woodruffe, 1993b; Wills, 1993). Generic competency models have been criticised for neglecting situation-specific issues such as the moral and ethical aspects of management, the changing nature of managerial boundaries, individual differences, and 'soft'

personal qualities, such as sensitivity and creativity, which are difficult to measure (Wills). Wills also argued that generic competency models focus on individual performance, which does not necessarily guarantee team and organisational effectiveness; and also that managerial performance is holistic and cannot be represented as separate competencies. However, these latter criticisms could apply to both generic and industry-specific competencies.

More recent generic competency models (e.g., Spencer & Spencer, 1993; Zingheim et al., 1996) address some of these criticisms with competencies such as flexibility, innovation, and team orientation being identified. Woodruffe (1993b) also suggests that generic competency lists were never intended to be seen as applicable in their entirety. It has been widely acknowledged that there are some competencies which are universal to managerial performance, and others that are situation or culture specific, and that many competencies may be irrelevant to any given job (Burgoyne & Stuart, 1978; Boyatzis, 1982, Spencer & Spencer, 1993). This implies that the optimal approach may be to derive a list that specifically fits a particular job in a particular organisation through a competency analysis, and to then use generic competencies either as a comparison or to speed up and/or enhance the accuracy of the process (Spencer & Spencer, 1993; Woodruffe, 1993b).

New Zealand research

Empirical research into generic managerial competencies in New Zealand is scarce. Wilson and Page (1993, cited in Elkin, 1995) found that most New Zealand managers emphasised task specific competencies rather than universal competencies. There was also a tendency to downplay personal characteristics and emphasise administrative tasks and traditional managerial functions such as the direction and supervision of employees. Similarly, Smith (1994) found that New Zealand managers were more concerned with managing their staff than with customer service or business awareness.

An exploratory study investigating managerial competencies across industries by Anso (1997) identified five competency clusters: Achievement

and action, business awareness, cognitive, leading and influence, and personal effectiveness. Although the results of this study were similar to overseas competency models (Boyatzis, 1982; Spencer & Spencer, 1993) in terms of the cognitive and interpersonal elements identified, there was also a significant difference. The previous competency models emphasised internal management issues, whereas the results of Anso's study additionally emphasised the importance of organisational and business awareness, and personal attributes relating to effectiveness. Anso's findings were consistent with the direction of recent overseas research into generic competencies mentioned previously, with competency profiles reflecting the changes taking place in the external environment (Wilkinson & Zwanenberg, 1994; Woodruffe, 1993b; Zingheim et al., 1996).

The emergence of competencies such as 'customer service orientation', 'organisational and external awareness', and 'flexibility' appear to indicate that New Zealand managers are becoming more conscious of the importance of awareness of and responsiveness to changes in the business environment. This need for business awareness, practical knowledge and skills, combined with generic 'soft' competencies such as personal characteristics, attitudes and values, may well reflect the constant change in both managerial responsibilities and the marketplace.

Following on from this, Spencer and Spencer (1993) suggest that, as managerial boundaries continue to change rapidly, context-specific knowledge may become less important in differentiating between superior and average performers. Instead, generic personal characteristics, such as innovation, learning agility, and self-understanding, may be more effective differentiators. These characteristics underlie a range of other skills that can only be developed if these 'meta-qualities' are present: Analytical thinking, pro-activity, sensitivity to change, and social skills (Pedler, Burgoyne, & Boydell, 1986, cited in Elkin, 1995).

The Numerical Business Acumen competency

The Numerical Business Acumen competency is an example of this combination of 'deeper' generic personal characteristics and resulting 'surface' knowledge and skills. This competency emerged as a result of conversations between Sheffield Consulting Group staff and a number of their large client organisations. These organisations had recognised the growing importance of employees not only being able to accurately analyse numerical data, but also being capable of applying those data in a meaningful way within the wider organisational and external context. An example of this would be to look beyond data illustrating that market share has decreased by three percentage points in the last six months, to determine what effect this may have on existing marketing strategies, departmental budgets, or the organisation's short and long-term strategic plans.

An extensive literature search revealed that no such competency has been identified in past research. This may be because the Numerical Business Acumen competency possibly has several underlying skills and attributes which are themselves competencies, such as data interpretation and business awareness. It could also be a reflection of the competencies that are now becoming important to managerial performance in keeping pace with changes in the external environment. After much discussion with 'experts' (e.g., Human Resources managers; Human Resources consultants) in this perceived competency, the following definition of Numerical Business Acumen competency was developed: Numerical Business Acumen is 'the knowledge, skills and attributes required to interpret numerical data in relation to their impact on the wider organisational and external environment, with a view to gaining optimum commercial benefit'.

With this definition in mind, two major components of this competency were identified: Critical reasoning ability (of which numerical reasoning is a sub-set) and business acumen. Each of these components will now be discussed in order to provide a thorough understanding of the theoretical basis of the Numerical Business Acumen competency.

Critical Reasoning

Reasoning is something that humans do everyday, when thinking about whether to do something and why we should do it, and in evaluating information we receive from others. However, not everyone can reason critically, that is, assess whether the conclusions drawn from the facts and evidence presented make sense. The terms 'critical reasoning' and 'critical thinking' are used synonymously in the academic literature, but for the purposes of clarity, this process will be referred to only by the term 'critical reasoning' in the present discussion.

Despite the availability of commercial tests which assess numerical critical reasoning, there is a surprising lack of academic research into this competency. Saville and Holdsworth (1991) defined numerical critical reasoning as the ability to understand and critically evaluate numerical data, for the purposes of their test of numerical critical reasoning. Critical reasoning itself can be based on different kinds of knowledge requiring different types of thinking. Therefore, for the present purposes, numerical critical reasoning is defined as the critical reasoning process specifically applied to the analysis and interpretation of numerical data. Because it is deemed to be an aspect of critical reasoning, the following background information will focus on critical reasoning, but should be presumed to also apply to numerical critical reasoning.

Definition of critical reasoning

The majority of the academic literature in this area defines critical reasoning as a process of events. That is, critical reasoning is a decision-making process which is reliant on information, background knowledge and previously accepted conclusions. As such, it consists primarily of understanding and evaluating the evidence and assumptions given in an argument, in order to judge whether the inferred conclusion makes sense (Cederblom & Paulsen, 1986; Haladyna, 1997; Norris & Ennis, 1986; Thomson, 1996).

Measurement of critical reasoning

An extensive review of the information available regarding commercial psychological tests (e.g., Conoley & Kramer, 1989; Impara & Conoley, 1996; Keyser & Sweetland, 1992; Sweetland & Keyser, 1991) indicated the availability of various critical reasoning tests. The content being measured in these tests is similar, with most measuring the ability to analyse, evaluate and make inferences from an argument or situation. However, the nature of the test items ranged from multiple-choice to essay questions. The focus of the multiple-choice tests was restricted to the evaluative aspects of critical reasoning; as opposed to the productive aspects, where test takers are able to generate conclusions by themselves rather than simply evaluate arguments. The other limitation noted for this type of test was that inference questions do not take account of test takers' varying backgrounds and knowledge of certain situations (Norris & Ennis, 1990). Thus, previous experience and knowledge, as well as the ability to generate conclusions, were not taken into account. The essay test addressed these shortcomings, but it had not been validated and was intended for developmental use only.

Surprisingly, although the majority of these critical reasoning tests had acceptable levels of reliability, little evidence of validation was presented. Although content validity was often addressed, the only evidence of criterion validity was for two different tests, showing moderate correlations with other cognitive tests. None of the test reviews provided any information regarding work-related criterion measures. However, many of the general cognitive ability tests available also include sub-sections of both critical reasoning and numerical ability[1].

Critical reasoning and work performance

The research on the relationship between critical reasoning and work performance is surprisingly slim. The fact that many critical reasoning texts provide guides on how to teach this skill to students (e.g., Cederblom &

[1] The usefulness of cognitive tests in predicting work performance will be discussed later in this chapter.

Paulsen, 1986; Norris & Ennis, 1990; Thomson, 1996), suggests that it is a valued skill. Indeed, Garrison (1992) claims that critical reasoning is a dominant theoretical framework in adult education.

Everyday observation also indicates that critical reasoning is a skill valued by organisations, because the ability to think critically often underlies crucial decision-making and problem-solving skills (Miller, 1997). Klemp and McClelland's (1986) meta-analysis of the intellectual capacities of senior managers found that superior performance was correlated with three specific competencies: Conceptual problem-solving, planning, and diagnostic information thinking.

As noted previously, one reason for the apparent lack of empirical research in this area may be that critical reasoning is often a component of general cognitive ability tests. Thus, the following summary of research in this area will be focused on the relationship between cognitive ability and job performance, with critical reasoning sub-tests mentioned where appropriate.

Extensive research has shown that general cognitive ability is a reliable predictor of job performance, yet it has its limitations. Measures of general intelligence usually include areas such as verbal and non-verbal reasoning, numerical ability, and inductive reasoning (Cascio, 1998; Thornton & Byham, 1982). Grimsley and Jarrett (1973, cited in Cascio, 1998) found that numerical ability (r =.42) and numerical reasoning (r =.41) significantly distinguished superior managers from average managers. Indeed, a test battery consisting of just verbal and numerical reasoning ability tests resulted in a multiple correlation of .52. These results led Grimsley and Jarrett to postulate that differences in intellectual ability were related to the level of managerial success in top levels of management.

The information available regarding the use of critical reasoning tests as reliable predictors of critical reasoning is inconclusive. The Watson-Glaser test was related to effectiveness in training in one study, with correlations of .20 to .30 (Thornton & Byham, 1982). Saville and Holdsworth's (1991) numerical critical reasoning tests produced mixed results as predictors of job

performance, with test scores moderately correlated with various performance criteria in several studies.

The above findings reflect the current debate regarding cognitive ability tests. According to some researchers, cognitive ability tests are unmatched as valid predictors of job performance for all job roles (Hunter, 1986; Hunter & Hunter, 1984; Ree, Earles & Teachout, 1994). These researchers advocate the use of a general cognitive ability which can be used to predict job performance in all settings. In response to this, Wagner (1994) presented the following empirical facts: The average validity coefficient between cognitive ability test scores and job performance is about .2, resulting in 4% of variance in job performance accounted for by those test scores. If that validity coefficient is corrected for measurement artifacts, the variance increases to approximately 25%. However, Wagner argued that validity coefficients provide at best limited and biased information about the true validity of cognitive tests because such coefficients overestimate the unique predictive ability of cognitive ability tests, as well as the magnitude of theoretical relations between cognitive ability and performance due to indirect causal effects.

Wagner (1994) suggested that the limitations of cognitive ability tests could be rectified by the use of a test battery, which would consist of both a cognitive test and as a measure of practical intelligence, such as tacit knowledge. This additional test would therefore measure the knowledge one gathers from both the context and content of every-day life. Similarly, Norris and Ennis (1990) suggested that critical reasoning should be taught in a variety of contexts, both subject-specific and general, to ensure that critical reasoning skills taught in academic environment were not limited to use only in similar situations.

Critical reasoning within context

The effectiveness of applying general principles of problem solving to any situation, irrespective of context and content, has also been criticised (Wagner, 1991). It appears to be a common view in the literature that critical

reasoning is primarily the analysis of argument, and is independent of context and content (Garrison, 1992; Little, 1980). Many books provide training in how to react to, and evaluate, evidence and arguments (Cederblom & Paulsen, 1986; Little, 1980; Thomson, 1996), yet relatively few aid in creating solutions. To expand on this, although reasoning may be an internal dialogue, it occurs within a certain context. The analysis of an argument is only one component of critical reasoning, as the evidence being analysed must also be considered within the context of our experiences. Thus, critical reasoning is a process of making sense of external experiences by analysing the information and issues at hand. There is a constant interplay between these internal and external processes, which results in the development of knowledge (Garrison, 1992). This often takes the form of an individual making sense of a situation by applying new ideas to previous knowledge and experience.

Brookfield (1987) identified five stages of the critical reasoning process which define it as both an internal and external process: A triggering event, an appraisal of the situation, an exploration to explain peculiarities, the development of alternative views, and the integration of those views into everyday life. This model begins and ends in the external context, with the internal process of reflection taking place in the middle three stages. Brookfield also identified four components of critical reasoning. Firstly, critical reasoning examines its own assumptions, beliefs and values. These assumptions come from past experiences and cultural norms, and are often hidden or unconscious. Critical reasoning evaluates these assumptions by examining patterns of behaviour, and verifies them for accuracy. Secondly, critical reasoning is never context-free. It is always influenced by the context of the 'thinker'; it is virtually impossible to ignore the circumstances and/or event that initiated the reasoning process. Thirdly, it is necessary to create alternative solutions and to evaluate them in terms of probable outcomes before accepting a proposed solution. Lastly, reflective scepticism should be encouraged. This moderate form of scepticism requires reasonable evidence to be shown before agreement is given.

Similarly, De Bono (1985) asserted that the standard concept of critical reasoning was lacking as a complete form of thinking in that it was reactive, rather than creative. He argued that the Western way of thinking which is based on dialogue and dialectical argument ignores the generative and the creative. As a result, De Bono developed his theory of six thinking hats. De Bono claimed that there are six different types of critical reasoning, which correspond to the roles taken on in everyday situations. These six types of thinking are: Neutral and objective (white hat), intuitive and emotional thinking (red hat), negative assessment and identification of potential problems (black hat), positive assessment and identification of potential benefits (yellow hat), creative (green hat), and objective thinking about the thinking process itself (blue hat). For effective critical reasoning, De Bono suggested that all of the six hats needed to be worn at some point in the overall process, because they are mutually complementary. In describing this type of 'action thinking', De Bono coined the term *operacy*, which is the skill of doing and the thinking that goes with it. These skills could be goal setting, evaluating priorities or generating alternatives.

Following on from this, Norris and Ennis (1990) suggested that critical reasoning consists of both evaluative and non-evaluative thinking. Creating alternative solutions to explain a problem is the non-evaluative part of critical reasoning, with the evaluation of those alternatives forming the evaluative part. However, Norris and Ennis also postulated that good reasoning is dependent on creative, as well as critical, reasoning. Creative reasoning consists of both reflective and non-reflective thinking. Brainstorming reasonable alternatives to a problem is an example of non-reflective thinking, with the conscious deliberation of those alternatives being reflective thinking. Thus, reflective creative reasoning overlaps with critical reasoning. Norris and Ennis hypothesised that both types of thinking are interdependent. Creative thinking requires evaluative critical reasoning before acceptance of the alternatives produced, and critical thinking requires the generation of alternative hypotheses before a potential solution can be critically evaluated.

Similarly, Scribner (1986) described the notion of practical thinking as the *mind in action*, which involves the acquisition and use of specific knowledge that is an important function of the larger activities in which problem-solving is embedded. This type of thinking is differentiated from the theoretical type of thinking, which is often involved in performing isolated mental tasks undertaken as ends in themselves, because it emphasises the dynamic process of thinking. Practical thinking is seen as being simultaneously adaptive to the constantly changing conditions in the world, and to the goals, values and knowledge of the individual (Scribner). Therefore, skilled practical thinking involves the ability to redefine a problem as well as the ability to solve it, the flexibility to develop a range of solutions which solve the 'same problem' but under varying circumstances, and taking into account the specific context of that problem. Scribner suggested that individual differences in this ability to think practically may affect cognitive performance.

Limitations of critical reasoning and the rational approach

Critical reasoning in its classic form exemplifies the rational approach to managerial problem solving. The hallmark of this approach is a sequence of principles: State the problem, describe it, identify differences between affected and unaffected standards of performance, identify associated changes with the problem, generate a list of likely causes, consider and verify the most likely cause of the problem (Wagner, 1991). The supposed advantages of this type of analytical approach include: (a) A reliance on principles of logic and scientific reasoning with the intended results of objective problem analysis and minimal bias; (b) their generality, in that the set of principles applies to any problem; and (c) their explicitness, resulting in this approach being easily communicated and taught (Wagner).

However, this rational approach to management was founded in Frederick Taylor's theory of scientific management (Yukl, 1994) which originated in the manufacturing industry. The move from manufacturing to service industries has resulted in many areas of managerial practice falling

outside of the bounds of rationality (Wagner, 1991). Indeed, there is evidence that successful managers rarely use the rational approach to problem solving (Mintzberg, 1973, cited in Wagner, 1991). Rather than dealing with problems in a systematic way, the managers in the study frequently dealt with problems with little idea of what the problem actually was, and what the optimal solution might be. The rational approach is often difficult for managers to apply because of existing biases inherent in the acquisition of information, the processing of information, and response selection (Hogarth, 1987, cited in Wagner, 1991). For example, many people have difficulty conceptualising problems in ways that go beyond their past experiences and knowledge. This corresponds with cognitive research that also describes many people as being seriously limited in their ability to understand and generate new solutions to common problems, especially in complex and abstract management situations (Bennett, Wheatley, Maddox, & Anthony, 1994). As a result, many individuals use standardised 'problem-solving' sets to cope with abstract situations, resulting in mildly innovative solutions that are based on a bounded set of values, rules and set of solutions learnt from previous situations.

Critical reasoning in action

However, perhaps the most consistent deviation of managerial practice from the principles of the rational approach is the bias for action (Wagner, 1991). It has been suggested that there is a lack of critical reasoning in business because of the rapid changes occurring in the business environment, coupled with an overload of information. This results in a need for rapid decisions, with managers focused on taking action, as opposed to gathering information and processing it sufficiently to take thoughtful action (Pascarella, 1997).

Indeed, Wagner (1991) described research showing that experienced managers took action very early in the problem-solving process. The managers' analytical skills were cursory and were based on personal experience, rather than more formal principles of critical reasoning and problem solving. Similarly, Schon (1983) found that many managers tended

to act almost spontaneously, with decisions based more on intuition than on rationality. Often when managers were asked to justify their decisions, they were unable to adequately explain reasons for their behaviour. Schon suggested that this was because "our knowing is ordinarily tacit, implicit in our patterns of action and in our feel for the stuff with which we are dealing. It seems right to say that our knowing is *in* our action" (*p.* 49).

Thus, one possible reason for this bias for action is that a large portion of the knowledge managers use in solving problems is tacit, and only becomes available indirectly in the context of action. Schon (1983) refers to this process as reflection-in-action: Managers critically analyse and restructure their intuitive understanding of a situation while taking action, using their past experiences to make sense of the situation. Therefore, this bias for action results in the process of reflection taking place concurrent with taking action, rather than as a precursor to action, as with the critical reasoning and problem-solving processes. Schon also noted that, although managers do reflect-in-action, they rarely reflect on their reflection-in-action. Consequently, their reasons for their actions are seldom discussed with others, and this important dimension of their managerial performance remains tacit, and inaccessible to others.

The role of intuition in critical reasoning

There is a growing acceptance of the use of intuition, or gut-feel, in management, in response to the need to keep pace with the rapid changes in the corporate and economic climate. Managers are, and will be, frequently required to make decisions where complete information is not available, and there is insufficient time to analyse the situation in a logical and ordered way. However, intuition has been conceptualised in various ways, hampering efforts to analyse and evaluate its role in management decisions. Intuition has been conceptualised as a sixth sense, a personality trait (such as the Myers-Briggs Type Indicator), and an unconscious process (Behling & Eckel, 1991).

The following definitions of intuition are most congruent with the current research, and may have more validity as they are based on cognitive

processes. Intuition has been defined as a set of observable actions based on the ability to gather information effectively and knowing how to use it in a meaningful way, and on the individual's knowledge of the relevant business (Behling & Eckel, 1991). Similarly, intuition may be defined as a distilled experience, involving variations on a relatively small number of decisions made previously. Thus, cues in a situation are recognised as being similar to ones previously encountered, and actions are selected or modified based on those that proved effective in that situation in the past (Behling & Eckel).

Intuition is a skill that becomes more prevalent as one goes up the management hierarchy (Agor, 1984; Mintzberg, 1973, cited in Wagner, 1991). Vasilash (1997) described a study investigating the critical reasoning skills used by several thousand staff in several organisations. The results showed that 26% of the managers surveyed made decisions strictly by gut-feel. Agor studied 2,000 managers and found that intuition was one of the skills most frequently relied on by top managers to make the right decisions. In addition, senior managers most commonly practiced an integrative thinking style. These managers were comfortable working with both facts and figures, and gut-feelings, for input into the decision-making process. However, the final decision was usually guided by their intuition. This corresponds with McKenzie's (1992) suggestion that to ignore any intuitive prompts which arise in the name of rationality is to disregard the experience that one has accumulated.

Despite the growing popularity and acceptance of the use of intuition in management (Agor, 1984), most researchers agree that the rational approach still has its place. Intuition is difficult to both put into practice, and to communicate to others, and is not in itself a complete solution to effective decision-making (Balle, 1994). Rather, many suggest that organisational effectiveness may lie in a blend of both critical reasoning and intuitive vision (Watts, 1993). This integrated style of thinking uses both analytical and intuitive skills interchangeably as the situation demands. An integrated thinker makes decisions guided by gut feel, after scrutinising the available facts and

information and receiving input from other members of the organisation (Agor, 1984).

In summary, the traditional view of critical reasoning has been primarily one of analysis, evaluation and inference of an argument. The majority of the available commercial critical reasoning tests seem to be based on this premise. Nevertheless, there appears to be a growing awareness that this view is only one component of critical reasoning, and that any analysis must be considered within the context of our experiences. The shortcomings of the rational approach to management thinking also reflect this, with the role of intuition and tacit knowledge being increasingly accepted as a valid way of contributing to managerial problem solving.

Business Acumen

Business acumen is often described by terms such as 'being street-wise', 'business sense', 'knowing the tricks of the trade', 'street smarts', and of course 'gut-feel'. Business acumen is commonly employed in the management literature as a supposedly technical term; however, it is surprisingly ill-defined, with numerous meanings given to it. After reviewing the 153 management articles published over the last 13 years (discovered using the ABI/Inform database) which mentioned the term business acumen, a distinct pattern emerged. In this literature, business acumen was frequently emphasised as being crucial to both individual and organisational success, yet often with no clear definition of what business acumen actually was. The definitions that were provided were often 'fuzzy', with business acumen being defined as either an understanding of, and experience in, business (Longenecker, Simonetti, & Mulias, 1996; O'Connor & Smallman, 1995); the ability to make strategic business decisions (Johnsson, 1991); business savvy or expertise (Filipczak, 1994; Johnsson, 1991); or a broader perspective and firm grasp of industry issues (Filipczak, 1994; Zuckerman, 1995). Indeed, possession of business acumen was often alluded to as contributing to organisational growth and competitive advantage, with the lack of this quality equally linked to stagnating or failing companies (O'Connor & Smallman, 1995). There were also numerous mentions of certain individuals' business acumen being associated with their outstanding success in their particular field. These observations were usually anecdotal, because there is a distinct lack of empirical research in the management field measuring the skills associated with business acumen.

By comparison, the term business acumen is not commonly found in the psychological research literature. However, practical competence has long been defined as a type of intelligence (Wagner, 1994). An aspect of practical intelligence is tacit knowledge, which is knowledge that is typically not openly expressed or taught (Wagner & Sternberg, 1985). This knowledge, combined with an individual's job experience (which in turn, is internalised by the individual to become tacit knowledge), appears to be equivalent in meaning to

the definitions found for business acumen. Because empirical research has been conducted on both tacit knowledge and job experience, these two areas will be used to form a theoretical basis for the business acumen component of Numerical Business Acumen.

Tacit knowledge

What is tacit knowledge?

An individual uses many types of knowledge every day in a job. This knowledge could be a factory worker's understanding of production methods, a lawyer's technical know-how, or knowledge embodied in patents, blueprints, and organisational structures and procedures. However, these types of knowledge, which can be expressed explicitly in words and numbers, represent just the tip of the iceberg of the entire body of possible knowledge required and used in any job. Increasingly, academic research also attests to the importance of tacit knowledge.

Polanyi (1962) was one of the first to conceive of the notion of tacit knowledge, identifying a type of knowledge 'of what cannot be said' (*p.* 87). Polanyi made an initial distinction between distal and proximal knowledge. Distal knowledge can be communicated openly and articulated clearly, and can be captured in documents, manuals, blueprints or procedures in a way that is clear and unambiguous to everyone concerned. In comparison, proximal knowledge (which is now more commonly known as tacit knowledge) is everything that distal knowledge is not. It is knowledge that which cannot be documented, formalised or reproduced through documents, techniques and models.

Wagner and Sternberg (1985) proposed that one product of the ability to gain informal learning on the job was the acquisition and use of tacit knowledge. They defined tacit knowledge as knowledge that is typically not openly expressed or taught. Hence, some knowledge that typically is tacit could be directly taught, such as knowing which general manager to approach with a new idea, but much tacit knowledge seems to be informal, unstructured and

somewhat inaccessible, making it potentially unsuitable for direct teaching. It is not that this knowledge is unteachable or not spoken of, but more that it is not taught directly to individuals as part of their jobs.

Recently, Myers and Davids (1992, 1993) summarised tacit knowledge as being distinguished from more formally developed skills in two ways: It is informal, and/or it is usually difficult to articulate. Myers and Davids' definition of tacit knowledge emphasised the practical knowledge gained through experience, leading to the development of 'know-how' or 'tricks of the trade'. This practical knowledge is usually untaught, increases with experience on the job, is largely context-dependent, is expressed in the form of 'rules of thumb', and is an automatic action in response to a specific situation.

Tacit knowledge may also include the intuition, beliefs and values that people form as a result of their experiences in life. These beliefs and assumptions form our individual mindsets which, in turn, influence decision making and the patterns of behaviour for everything we do (Saint-Onge, 1996). Following on from this, Nonaka (1994) suggested that tacit knowledge involves both cognitive and technical elements. The cognitive element consists of the schemata, paradigms, beliefs and views that help an individual to define and perceive their world, whereas the technical element consists of the know-how, crafts, and skills that apply to specific contexts.

Thus, tacit knowledge is the underlying knowledge that tends to be taken for granted as managers respond to everyday situations, or take action in accord with the overall organisational direction. Examples of tacit knowledge are knowing which people within an organisation to approach with a new idea and what angle to use to win them over, being aware of the values and norms of the organisational culture, or knowing which projects to accept that will be most beneficial to that individual.

The tacit knowledge process

So how does tacit knowledge impact on work performance? It is worth noting that tacit knowledge is by no means all that is necessary for job success, but is just one, albeit important, element. The skills that are used in any job role

often involve making the right choice from various solutions, although these options may often be selected automatically and without the individual being aware that a choice is even being made (Nelson & Winter, 1982). Thus, a major component of the knowledge that underlies a skilful performance is often tacit, in the sense that the individual is not fully aware of all the sequences carried out and could find it difficult to articulate those details. In fact, a sign that one has successfully acquired a new skill is a decreasing need to pay attention to detail. This corresponds with Polanyi's (1962) notion of two levels of awareness: Focal (conscious) and subsidiary (unconscious). As skill expertise grows, more reliance is placed on intuitive feeling to guide one's performance due to the skill sequence used becoming gradually relegated to an unconscious, or automatic level of awareness. In this way, personal knowledge becomes less easy to explain verbally, and more easy to explain through observation or direct experience.

In order for managers to be flexible in response to constant change and to provide innovative solutions to stay ahead of their competition, they must be able to take calculated risks which are often based on 'gut-feel' or their 'street-smarts'. This informal method of decision making relies on that person's tacit knowledge. Many managers have discovered through their own experiences that their use of tacit knowledge can be beneficial, resulting in a faster decision-making process, more effective decisions, and fewer pertinent factors necessary to make those decisions (Brockmann & Simmonds, 1997).

Tacit knowledge and work performance

Wagner and Sternberg (1985, 1987) have carried out extensive empirical research into the role of tacit knowledge in both academic and business settings. These researchers divided tacit knowledge into three categories, according to content: Tacit knowledge about managing oneself, which refers to the knowledge about how to manage oneself to maximise productivity; tacit knowledge about managing others, which refers to knowledge about managing both subordinates and other business relationships; and tacit knowledge about

managing career, which refers to knowledge about one's career reputation and influencing others.

Wagner and Sternberg's (1985) initial study produced three main results: Expert-novice differences in tacit knowledge were found for groups whose members differed in amount of experience and formal education; differences in tacit knowledge were strongly related to a wide variety of criterion measures of work performance; and tacit knowledge scores were unrelated to scores on verbal intelligence tests. Wagner and Sternberg also noted that the magnitude of correlations between tacit knowledge scores and criterion reference measures was approximately 1.5 to 2.5 times that of the correlations usually obtained between intelligence test scores and job performance measures (i.e., their correlations clustered around the .40 level, rather than the .20 level).

On the basis of these results, Wagner and Sternberg revised and extended the tacit knowledge framework to include two different orientations of tacit knowledge: A local orientation, which refers to being focused on short-term tasks and goals; and a global orientation, which refers to being focused on long-term goals when making decisions (Wagner, 1987; Wagner & Sternberg, 1986; Wagner & Sternberg, 1987). Thus, the three previously mentioned categories of tacit knowledge content were combined with these two orientations to form a more complete model of tacit knowledge.

Wagner and Sternberg (1986) developed a Tacit Knowledge Inventory for Managers from this revised framework, which consisted of a set of scenarios, each of which described a work-related situation associated with a set of response items presenting alternative courses of action. Further studies were carried out with participants who were required to rank the quality of these alternative courses of action. Performance on this measure was quantified by both identifying response alternatives that reliably distinguished experienced managers from those without management experience, and also by comparing participants' responses to a prototype devised from expert judgement. The two methods of quantification yielded almost identical results. The results of this study provided clear support for the three previous findings mentioned above.

That is, tacit knowledge increased with years of experience and, within any given level of experience, greater tacit knowledge was associated with greater career success. However, although the development of tacit knowledge appears to depend on years of experience, not all individuals appear to develop it to the same extent.

The results of these studies supported a tacit knowledge model which could be characterised as a general ability, as opposed to several independent abilities (Wagner, 1987). Correspondingly, the various kinds of tacit knowledge could be viewed as facets of this general ability as opposed to independent psychological constructs. For example, someone who is highly skilled in managing themselves often tends to be skilled at managing others and their tasks. Wagner (1987) suggested that a general tacit knowledge ability could imply that at least some of the tacit knowledge in one field is carried over to another. Following on from this, if someone has a high level of tacit knowledge in one industry, then they should be able to generalise at least some of it to another industry. It is possible that, however general the ability to acquire tacit knowledge is, the type of tacit knowledge is at least partially specific to the industry that that particular individual is currently employed in.

Wagner (1987) suggested that there are two possible sources of this general factor of tacit knowledge, which could also apply to the general factor for academic intelligence. One could be a manifestation of something inside an individual, such as a general ability to acquire tacit knowledge, or, it could reflect the nature and extent of their learning experiences. Exactly how this general factor is best conceptualised has not yet been clarified.

Job experience and tacit knowledge

Job experience refers to the length of time spent working in a particular occupation, whether it is with just one organisation or across several organisations. Therefore, the quantity of that experience is measured, but not the quality or diversity of the experience gained.

Wagner and Sternberg (1985) found that tacit knowledge was significantly correlated with managerial level (r =.34), years of tertiary education (r =.41), and salary (r =.46). Years of management experience was correlated with tacit knowledge to a lesser degree (r =.21), which suggested that tacit knowledge is not automatically acquired with each passing year. On the contrary, Wagner and Sternberg suggested that the strong relationship between tacit knowledge and salary and managerial level indicated that qualitative aspects of managerial experience are more important than quantity of experience based on the number of years served. However, the results of a later study (Wagner & Sternberg, 1986) were somewhat conflicting, with both tacit knowledge (r = -.30) and salary (r = -.21) negatively correlated to years of management experience, and no reliable relationships found between tacit knowledge and years of tertiary education and managerial level. This could be accounted for by the fact that Wagner and Sternberg (1986) did find expert-novice differences in tacit knowledge in groups of individuals whose members differed in amounts of training and experience and level of advancement. However, performance on the tacit knowledge measure also increased as a function of the level of professional advancement.

Barrette and Durivage's (1997) study found that tacit knowledge does not necessarily increase with length and variety of experience. Marginal differences in tacit knowledge were found between the three groups in their sample (novices, juniors and seniors) with the only significant differences found between the extremes (novices and seniors). This result was attributed to the possibility that what was measured was in fact an aptitude for practical judgement specific only to the situations presented in the test, as opposed to an aptitude for using tacit knowledge more generally in one's work. Thus, Barrette and Durivage suggested that their test did not measure so much the acquisition of tacit knowledge resulting from work experience, but rather the ability to make judgements independent of one's relative work experience.

Thus, job experience may contribute to the acquisition of other competencies, and how effectively an individual applies his/her knowledge

within a role (Klemp & McClelland, 1986). Brockmann and Simmonds (1997) suggested that although tacit knowledge is acquired through experience, how that tacit knowledge is used depends on the combination of experience and propensity to rely on intuition. The results of their study showed that industry-specific experience was significantly related to the use of tacit knowledge, whereas more general experience (measured either by age or years of tenure with the organisation) was nonsignificantly related. The explanation given for these findings was that industry dynamics were more applicable to analysing the business environment and therefore were of most significance to decision making.

Thus, the extant research outcomes are equivocal. Tacit knowledge does not always increase with years of job experience, and does not do so automatically with each passing year. Some people who are relative novices in a field may have a surprisingly high level of tacit knowledge, whereas other people who have been in a field for years may have a relatively low level of tacit knowledge. In other words, the level of tacit knowledge seems to depend less on the amount of experience one has, and more on how much one has learned from it.

Sternberg and Wagner (1989) suggested that the ability to internalise experience is dependent on three processes: Selective encoding (deciding what new information is relevant), selective combination (deciding how to use that information in a useful way), and selective comparison (identifying relationships between new and previously learnt information). The effectiveness of these processes can vary widely across industries, depending on how the relevant information is presented, the individual's prior knowledge base, and their motivation to learn (Sternberg & Wagner, 1989). Therefore, intra-individual as well as inter-individual differences in tacit knowledge can be expected.

Job experience and work performance

Selection methods based on the consistency principle (which assumes that past behaviour and/or performance predicts future behaviour and/or performance in jobs) have been found to have the highest predictive validities

(Russell, 1990). This has resulted in the widespread use of biographical information (of which past job experience is a major source) in selection (Cascio, 1998). However, as with tacit knowledge, it is not necessarily the experience itself, but what was learnt from the experience that is the supposed source of the predictive power of biographical information (Russell, 1990).

Although both the quality and quantity of an individual's past job experiences are commonly used in selection processes, the research conducted on the relationship between job experience and job performance may not justify this. Positive correlations have been found between job experience and job performance for all levels of both job experience and job complexity (McDaniel, Schmidt, & Hunter, 1988; Schmidt, Hunter, & Outerbridge, 1986). However, some moderator effects have been noted. The relationship between job experience and job performance is non-linear, as it has been typically found that the rate of increase in performance with increases in job experience decreases at higher levels of job experience (Schmidt, et al.). Also, correlations between job experience and job performance were higher for low-complexity jobs than for high-complexity jobs (McDaniel et al.; Schmidt, et al.). This effect was explained by the differences in available educational knowledge, as high-complexity jobs have more formal education sources through which to gain job knowledge. In this way, job experience was hypothesised to indirectly affect job performance through its impact on job knowledge.

Practical intelligence

Many managers and professionals agree that the knowledge that has made the greatest difference to their careers took place after their academic learning, and was gained on the job (Wagner & Sternberg, 1985). Indeed, many individuals who are highly successful in their jobs have very average histories of academic performance and, conversely, individuals who have histories of outstanding performance in formal education often are only moderately successful in their jobs.

Several researchers have noted that doing well in traditional intelligence tests does not guarantee that one will function intelligently in everyday life (Klemp & McClelland, 1986; Neisser, 1976; Scribner, 1986; Wagner & Sternberg, 1985). As Klemp and McClelland point out, problems in real-life are seldom as clear cut as those given in an intelligence test. Finding an effective solution to everyday problems faced on the job is rarely just a matter of logical aptitude, due to the possible influence of other variables such as the work environment and culture, ambiguity in the situation, and the complexity of interpersonal dynamics. By comparison, intelligence test problems have one correct solution, a limited amount of time to reach that solution, and limited relevance to real-life problems (Neisser, 1976). Wagner and Sternberg (1985) suggested that the lack of a relationship between academic success and success in real-world pursuits, such as job performance, is due to the fact that the intellectual demands in academic learning are actually just a subset of the intellectual demands of real-life situations. Thus, the tasks found in academic learning settings and in IQ tests can be described as measures of 'academic intelligence', which is just one component of intelligence in its broadest sense.

Many researchers have argued that there is another, equivalent component of intelligence - 'practical intelligence', or intelligent behaviour in natural settings (Barrette & Durivage, 1997; Klemp & McClelland, 1986; Neisser, 1976; Scribner, 1986; Sternberg & Wagner, 1985). Practical intelligence is defined as procedural knowledge that is relevant to everyday life (Sternberg & Wagner, 1989a).

After interviewing many individuals who were seen as having practical intelligence, Wagner and Sternberg (1987) noted three main perceptions. Firstly, IQ and ability test scores were not perceived as being predictive of success on the job. Secondly, it was generally agreed that although academic learning could be useful, it was not vital to success on the job. Thirdly, the factor considered to be crucial to success was the informal learning that took place on the job, and its contribution to an individual's practical intelligence.

Future research still needs to determine the relative contribution of both practical and academic intelligence to managerial performance. Wagner and Sternberg (1986) found that tacit knowledge scores accounted for two to five times more variance in job performance than did scores on traditional intelligence tests. A later study (Wagner, 1994) found that tacit knowledge scores consistently accounted for an additional 32% of criterion variance not accounted for by IQ scores. Barrette and Durivage's (1997) study showed that logical reasoning tests accounted for 12% of variance in work performance, and tacit knowledge accounted for an additional 18%.

Although research has shown mixed results for any significant relationship between tacit knowledge and academic intelligence tests, the notion that the two are inter-related makes good sense. We know that knowledge can be acquired through both formal training and through experience. Although any interdependence of the two kinds of knowledge may not be apparent, one is incomplete without the other. Formal academic knowledge needs to be validated in real-life experience and, similarly, experiential knowledge should have a base in formal knowledge (Sen & Das, 1991). For example, managers may need to give their 'gut-feel' decisions a theoretical base to validate the decisions and, correspondingly, management graduates need to relate and assess the utility of the theoretical knowledge they have learnt through practical experience.

Numerical Business Acumen

The previous sections have illustrated how the changes that have taken place in the socio-economic and business environments have impacted on the assessment of managerial skills. The rational approach to management, based on stable, predictable events, can no longer adequately keep pace with the rapidly changing external environment. Correspondingly, it is now recognised that the characteristics of a successful manager depend not only on formal learning but, to a larger extent, on the business acumen developed on the job. This business acumen results from the ability to successfully adapt to real-life contexts, using a combination of analytical skills and tacit knowledge. Thus, competencies which focus on responsiveness to those external dynamics, such as Numerical Business Acumen, are now seen as important constructs to assess.

What is Numerical Business Acumen?

The Numerical Business Acumen competency has been defined as the knowledge, skills and attributes required to interpret numerical data in relation to their impact on the wider organisational and external environment, with a view to gaining optimum commercial benefit. This definition will be broken down and discussed in two separate parts: The interpretation of numerical data and the interpretation of data in context.

As mentioned previously, there are different levels of critical reasoning. In the same way, there may be different levels of data interpretation: Numerical ability, analytical reasoning, and Numerical Business Acumen. Many people have a high level of numeracy, and are capable of performing mathematical calculations. Some of those people may also be able to use those numerical skills to analyse and interpret data, using effective critical reasoning skills. However, in order to be able to reason critically, an adequate level of numeracy is assumed.

Again, some of those people who are able to interpret and critically analyse numerical data may also have the ability to broaden that thinking to interpret the data in the context of organisational strategy, competitor activity

and economic influences. Thus, they are applying meaning obtained from the data for the optimum commercial benefit of the organisation. For example, sales figures for the last several months may have declined steadily. Interpreting this in isolation, one might recommend that action be taken to rectify the situation. However, interpreting these figures in a wider context might result in a recommendation of no action, as additional market research may have indicated that consumer demand for that particular product is declining and unlikely to reverse. Therefore, a certain amount of business acumen is necessary to critically analyse any given situation at this level.

Components of Numerical Business Acumen

Numerical Business Acumen is a collection of knowledge, skills and attitudes, which are grouped under two major headings: Numerical critical reasoning ability and business acumen. These two components can themselves be broken down further into underlying sub-characteristics (see Figure 2).

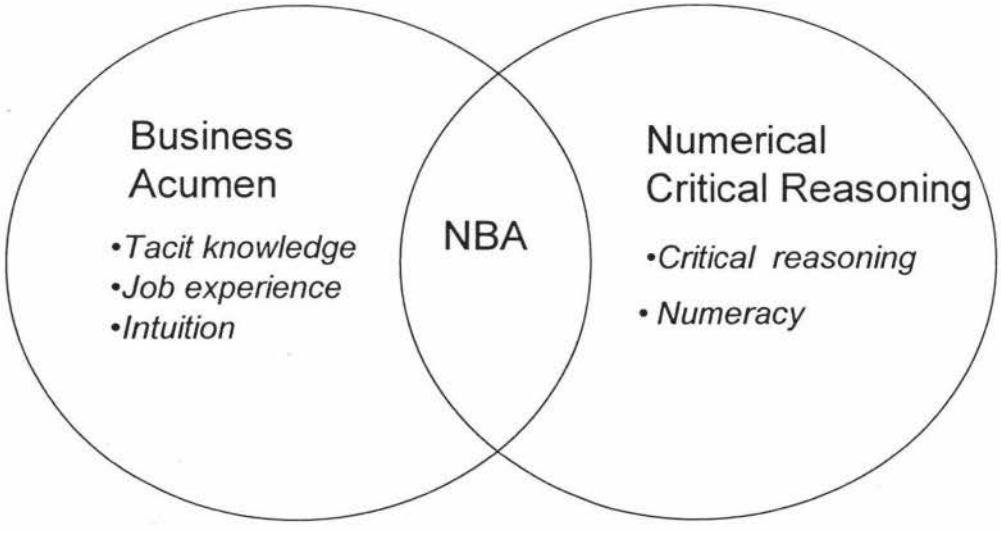


Figure 2

Numerical Business Acumen model

Numerical critical reasoning is defined as the ability to understand and critically evaluate numerical data (Saville & Holdsworth, 1991). The underlying skills involve a combination of critical reasoning principles and

numeracy. As previously noted, these skills do not include any consideration of the context provided by past experience. By including business acumen as an underlying characteristic, an individual's previous job experience, amount of tacit knowledge and use of intuition are also included. The resulting competency bundle is therefore a combination of aspects of an individual's practical intelligence, represented by the business acumen component, and academic intelligence, represented by the numerical critical reasoning component.

With reference to the Iceberg model (see Figure 1), Numerical Business Acumen is a combination of 'deeper' generic personal characteristics and resulting 'surface' knowledge and skills. As discussed previously, tacit knowledge and intuition are usually 'hidden', and may underlie the process of how that individual internalises their experiences. These deeper characteristics result in another aspect of business acumen, organisational and industry knowledge, corresponding to surface knowledge. Both numeracy and critical reasoning are also surface level skills in that they tend to be visible, and can be taught easily.

Behavioural indicators

Numerical critical reasoning and business acumen were further analysed to produce specific behaviours demonstrating below average, average and superior performance. These behavioural indicators were organised in an ascending scale from Level 1 to Level 6 (see Table 1). This rank ordering is known as a 'just noticeable difference scale' (Bethell-Fox, 1993). In other words, the incremental change in behaviour associated with moving from one level to the next should be distinct enough to allow that change to be easily observable and measurable.

The resulting indicators were based on behavioural descriptions used in Hay/McBer's research (Spencer & Spencer, 1993) for the Initiative, Information Seeking, Analytical Thinking, Conceptual Thinking, and Technical Expertise competencies. Additional behavioural indicators, such as the level of numeracy

and awareness of the marketplace, were derived by the researcher to complete the overall construct of Numerical Business Acumen.

Table 1

Behavioural indicators of Numerical Business Acumen

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Little understanding of numerical principles and limited numerical skills Only uses 'rules of thumb', common sense and past experience to identify problems Sees only essential similarities between current and past situations Responds to situations without analysing impact of actions	Uses basic numeracy skills Observes basic discrepancies, trends, and interrelationships in data Sees crucial similarities and differences between current and past situations Recognises the implications of patterns and trends shown in data to overall organisational performance	Understands and uses basic statistical analysis and established methodologies Can see multiple relationships among several parts of the situation Recognises several likely causes of events, and consequences of actions between parts of the data Distinguishes between relevant and irrelevant data Applies past knowledge to analyse current situations Determines current problems or opportunities based on the data, with a basic understanding of possible impact on organisational performance	Uses statistical analysis to manipulate and interpret data Makes long chains of causal connections, by identifying cause and effect and seeing multiple relationships between parts of the problem Identifies key issues in a complex situation Considers the advantages, risks, and implications of any decisions made on the organisation, in evaluating and choosing alternative solutions Recommends possible courses of action to maximise possible opportunities and minimise potential problems, with the objective of increasing profit.	Uses statistical tools to manipulate data to reveal trends or predictions not evident without analysis Sees multiple relationships among several parts of a problem, recognises several causes of events and several consequences of actions Identifies useful relationships among complex data from unrelated areas, and pulls together unrelated pieces into a comprehensive whole Extracts larger meaning from data to recommend possible courses of actions to opportunities/problems working within current organisational objectives and long-term strategy Identifies key opportunities/ problems that could impact on the organisation in the near future	Uses complex and multi-variate numerical analysis to manipulate and interpret data Can deduce correlation and causation; generates and tests multiple hypotheses for a situation; identifies useful relationships among complex data from unrelated areas Identifies several alternative solutions to key opportunities/ problems, developed to correspond with political / competitive environments Uses data to compare organisational performance to that of competitors; and to recommend ways of increasing competitive advantage Uses industry knowledge and skills to recommend possible courses of actions working within organisational objectives and long-term strategy

The three identified levels of data interpretation were associated with the following levels (with some degree of overlap): Numeracy was described by the behavioural indicators in Level 1; numerical critical reasoning was described by Levels 1 to 3; and Numerical Business Acumen was described by Levels 3 to 6. Therefore, superior performance in the Numerical Business

Acumen competency would be defined by those who demonstrated behaviours listed in Level 6.

Of course, these behavioural indicators are simply a guide. Thus, any definitions of superior and average performance made according to this ranking scale would need to be modified to meet the objectives of the job role in question, and the culture of the organisation. Ballantyne and Povay (1995) also noted that examples of behaviour should be given which are neither so general that they apply to most of the population, nor so specific that they only apply to a small part of a person's job. Individual behaviours should be clustered together in such as way that they have meaning to anyone likely to be demonstrating that competency.

Numerical Business Acumen and work performance

The relative contribution of tacit knowledge and academic knowledge to managerial performance has been discussed. Research has shown that, although the correlation between cognitive ability tests and work performance averages between .2 and .3, tacit knowledge measures appear to add approximately 20% to 35% of explained variance in predicting work performance (Barrette & Durivage, 1997; Wagner & Sternberg, 1986). Because tacit knowledge and numerical critical reasoning are the main components of Numerical Business Acumen, there is likely to be a relationship between Numerical Business Acumen and work performance. However, this relationship will not be directly assessed in the present research due to privacy considerations associated with accessing work performance measures.

Numerical Business Acumen and managerial level

Existing comparative research has established that there are differences between the managerial levels in the relevance of various competencies, and how these competencies translate into behaviours (Boyatzis, 1982). As one moves up the management hierarchy, the type of competencies used moves from technical and specialised knowledge to strategic planning and decision-making (Ballantyne & Povah, 1995). Empirical research has also indicated that

intuition is a skill that becomes more prevalent as one goes up the management hierarchy (Agor, 1984; Mintzberg, 1973). Business acumen becomes crucial when making successful decisions at senior management level, whereas more analytical, technical skills are required at entry-level management. Therefore, having a high level of Numerical Business Acumen may be associated with movement up the management hierarchy.

Is Numerical Business Acumen generic?

A generic characteristic may be apparent in many forms of behaviour, within many jobs in many organisations. The need for context to be considered in critical reasoning has already been discussed and, as a result, one could argue that, if critical reasoning is largely context-bound, it cannot be generic. Wallace and Hunt (1996) have suggested that individuals use both specialist and generalist knowledge paradigms to make sense of their environment. These knowledge paradigms, which are essentially scripts and schemata, would be termed 'experience' by most managers. Therefore, managers who draw upon their existing experience to make sensible inferences about new issues or environments are more likely to perform competently under any altered context or changed circumstances. In other words, such managers use their business acumen to successfully adapt to new situations. Wallace and Hunt found that although a certain amount of business acumen is industry-specific, much of it is generic as it consists of core personality traits, and is therefore portable across situations.

Measurement of Numerical Business Acumen

A thorough review of commercially available tests revealed only one critical reasoning test designed to measure a construct similar to Numerical Business Acumen. The Critical Reasoning Tests (Smith & Whetton, 1992, cited in Conoley & Kramer, 1989) were designed to assess the ability to use verbal and numerical data to make decisions similar to those asked of a manager. The numerical test consisted of 30 multiple-choice items based on the types of decisions a manager would be required to perform every day. Interestingly,

although the test was designed to measure practical intelligence as opposed to academic intelligence, the only criterion validation measure used in the test's development was school grades, a measure of academic intelligence. Reviewers concluded that the evidence of validity and reliability was insufficient for the test to be recommended for any other than research purposes.

Current critical reasoning tests focus on the assessment of an individual's ability to analyse, evaluate and make inferences. The circumstances surrounding a situation are not considered, which limits their ability to predict performance in real-life situations. The present critical reasoning tests are therefore likely to be inadequate measures of the Numerical Business Acumen competency. Because Numerical Business Acumen involves interpreting numerical data within its organisational and external context, a new measure needs to be developed; one that is reliable and valid. In order to measure the business acumen component of the competency, the test will need to replicate real-life scenarios accompanied by questions designed to draw out tacit business knowledge.

Numerical Business Acumen uses in Human Resources Management

The Numerical Business Acumen competency has been identified by several New Zealand managers as an important skill (Anso, 1997). Because it is a combination of core traits and behavioural skills, this competency has potential uses in both selection and development. Aspects of Numerical Business Acumen, such as intuition and tacit knowledge, may be difficult to assess and develop, and should therefore be identified in the selection process. Other aspects, such as critical reasoning skills are relatively easy to develop. A measure of Numerical Business Acumen would also be a useful tool for assessing the quality, as opposed to quantity, of job experience as the internalisation of those experiences would become tacit knowledge.

Aims of the present research

Although Numerical Business Acumen has been identified as a competency which may be important to managerial performance, it is difficult to accurately assess its relative contribution, as there appear to be no commercially available tests which adequately measure this construct. Therefore, the main aims of the present research were to: (a) Develop a scenario-based, pencil-and-paper measure of the Numerical Business Acumen competency, and (b) assess its psychometric properties.

Methodology

Participants

A convenience sample of seven organisations in the Auckland area took part in the present research. Five of those organisations were identified from Sheffield Consulting Group's client database, with the researcher approaching the remaining two organisations. Each of these organisations had either previously expressed an interest in the Numerical Business Acumen construct, or was open to taking part in the research. All of the organisations were relatively large in size, ranging between 50 to 8500 employees. There were six industry groups represented: Retail, fast moving consumer goods, airline services, financial services, energy, and human resources consulting. Within those industry groups, the occupational role of individual participants varied widely with sales representatives, analysts, consultants, management trainees, information technology and marketing roles being represented. A total of 46 managers (25 males and 21 females) took part in the research, with the number recruited from each organisation ranging from 2 to 13 participants. A summary of demographic details for the sample is presented in Table 2.

Table 2
Demographic details (N=46)

	Frequency
Business experience (years)	
0-5	13
6-10	3
11-15	14
16-20	4
21+	12
Highest academic qualification	
School Certificate	3
University Entrance	3
Higher School Certificate	2
Polytechnic qualification	1
University (Bachelors degree)	26
University (Post-graduate degree)	11
Age	
20-25	15
26-30	11
31-35	11
36-40	7
41+	2

Industry	
Airline	13
Energy	3
Fast moving consumer goods	15
Financial services	4
Human resources consulting	9
Retail	2

The range of industry experience among participants was from 1 year to 20 years, with 28% of the participants having less than 5 years experience, another 37% having 6 to 15 years experience, and the remaining participants having over 15 years experience. Highest academic qualifications varied from School Certificate to post-graduate degrees, with 87% of participants possessing tertiary qualifications. Participants' ages ranged from 20 to 50 years, with 33% of participants aged between 20 and 25 years, 48% aged between 26 and 35 years, and the remaining participants aged 36 years and over.

Procedure

The Human Resources Manager in each organisation was initially contacted by the Sheffield Consulting Group consultant managing the relationship for that organisation. The purpose of this initial contact was simply to introduce the research project. Subsequently, the researcher contacted these managers to introduce herself and to arrange interview times. The purpose of the interviews was twofold:

- To gather background knowledge regarding Numerical Business Acumen (see the section on Scenario Development for details).
- To obtain the managers' consent to proceed with the research. If consent was obtained, their assistance was then requested in the recruitment of participants at management level within the organisation.

Of the eight Human Resource Managers approached, seven agreed to participate in the research. These seven managers were provided with Information Sheets for both themselves and potential participants. These sheets detailed the purpose of the research project, what would be required of

participants and the assurance that test results would have no bearing on their role within the organisation.

Dates for the first administration sessions were organised through the Human Resources Manager, and Information Sheets and Consent Forms (see Appendices A and B for examples) were then sent out to the Human Resources Manager for distribution to each prospective participant two weeks prior to the administration session. All administration sessions took place in a group setting, with the tests being administered simultaneously to the whole group. The sessions took place in various conference or meeting rooms made available by that organisation. Participants were asked to bring a calculator, but writing paper, pencils and erasers were provided.

At the first administration session, participants completed the self-ranking scale and demographic questions, and then Form A of the Numerical Business Acumen measure and a comparative measure (NMG2[1]) were administered. A short break was allowed between the two tests if the participants wished. Each test was 35 minutes long, with the total administration session taking approximately 80 minutes. Signed consent forms were collected by the researcher at the beginning of the session. The participants were reminded of their right to decline to participate, and that the test measure was independent of the requirements of their organisational role. Standardised administration instructions were used across sessions for both measures (see Appendix C for the Test Administration Guide). Participants were asked to select an identification code which would be used again at the second session. They were also told when they would receive feedback on the results of the measures, and the procedure for doing so.

The second administration session took place two to four weeks later, depending on when it was convenient to the organisation. The interval between administration sessions was determined by the time constraints of the research project, but was deemed long enough to offset the effects of practice. Thirty-two of the original 46 participants took part in the second administration

---

[1] Saville and Holdsworth's Numerical Critical Reasoning test.

session. In the second session, the parallel form of the Numerical Business Acumen measure, Form B, was administered (using the standardised test instructions), taking 35 minutes. At the end of this session, participants were thanked for their time and reminded again of feedback procedures.

Design of the Numerical Business Acumen Measure

Scenario Development

The first step in developing the Numerical Business Acumen measure was to interview the Human Resource Manager in each prospective organisation. The researcher's definition of Numerical Business Acumen was given, and the managers were asked to expand on that definition by describing the skills, knowledge and abilities present when an individual was perceived to possess Numerical Business Acumen. The managers were also asked to identify behaviours distinguishing superior performance from average performance in this competency, to describe typical work-related situations where this competency would be demonstrated, and to describe how an individual possessing that competency would act in that situation. Exploratory questions were also asked regarding performance measurement procedures in each organisation, and the manager's evaluation of the adequacy of current commercial Critical Reasoning psychometric tests in measuring the skills underlying Numerical Business Acumen. This additional information was gathered to provide insight into how the present measure should be structured and to identify possible criterion reference measures.

An initial set of six scenarios was developed from these descriptions, in consultation with Sheffield Consulting Group. Each scenario described a typical work-related situation which was broad enough in scope to be understood by most industry groups. As part of the description, a variety of numerical data were presented in graphical and tabular form, such as profit and loss statements, sales figures, market share figures, fee structures, and customer service ratings. The data were also accompanied by statements

providing additional information about the situation described, thus providing a detailed overall picture.

A scenario-based measure was chosen as a valid method of replicating situations which individuals could encounter regularly in their occupational roles. This type of measurement has been used extensively as a valid way of measuring tacit knowledge (Sternberg, 1994; Wagner, 1987; Wagner & Sternberg, 1987), which constitutes a major component of the underlying structure of Numerical Business Acumen. This method is based on the simulation approach, which consists of observing individuals as they handle tasks that simulate job performance. A well-designed simulation which models a clearly defined domain, such as the Numerical Business Acumen construct, can obtain a sample of behaviour which represents a larger behavioural domain, such as a manager's job (Thornton & Cleveland, 1990). However, due to resource constraints, a pencil-and-paper task designed to resemble a simulation was used, rather than an actual simulation. The six scenarios initially developed ensured that a sufficiently broad range of data types encountered in various industries was covered.

Question Development

Superior performance in Numerical Business Acumen is dependent on an individual's ability to analyse a situation, and identify and solve any current or future issues. This process involves multiple steps, so a single test item would be insufficient to adequately measure performance in this competency. A pictorial item set, where the corresponding scenarios act as stimuli for the correlated set of items, was used. This type of item set provides a basis for testing critical thinking and problem solving and has been highly recommended for this purpose (Haladyna, 1997).

Each scenario was associated with four questions in the item set. The first question was designed to measure the numerical critical reasoning component of Numerical Business Acumen, which would enable direct comparisons to be made between scores on this test item and scores on comparative measures. The remaining three questions were designed to

measure the business acumen (tacit knowledge) component of Numerical Business Acumen.

A selected-response format, which is defined by a written question and pre-constructed choices that the test taker can select from (Haladyna, 1997), was chosen. Selected response test items have been proven to be more reliable, easier to score, and more objective than constructed-response test items, due to the fact they are free from issues of inter-rater reliability and rater-bias (Anastasi & Urbina, 1997). Although it can be argued that constructed-response questions (where the test taker responds to a written question and the administrator develops a scoring system to measure the answers) elicit more in-depth responses and more naturally elicit what the participant knows, individual variance in reading and writing abilities, as well as the issues mentioned above, can contaminate such measurement (Haladyna). In addition, various researchers have found that both selected-response and constructed-response items produce comparable results when measuring mathematical knowledge or mental abilities (Anastasi & Urbina; Haladyna).

However, a constructed-response item was also included to further explore how each participant's underlying reasoning processes contributed to their competency in Numerical Business Acumen. This item questioned participants' rationales for their answers to one of the previous items. These answers were not included in the overall test score, but were used to aid the understanding of the tacit knowledge component of Numerical Business Acumen.

Of the four selected-response items, a combination of three multiple-choice questions and one ranking question was used. The multiple-choice items each consisted of four to five alternatives, consisting of one best answer and several incorrect choices (distractors). This format is consistent with traditional psychometric research practice (Anastasi & Urbina, 1997). Each scenario item set was also constructed so that the answer to a given item was relatively independent the previous items. The ranking question consisted of a set of five to six responses presenting critical steps involved in analysing a

situation. Participants were asked to rank those steps in order of execution, in order to assess their problem-solving ability. The rationale for this type of format was based on Wagner and Sternberg's (1986) Tacit Knowledge Inventory for Managers, which consisted solely of the ranking of alternative courses of action - the reasoning being that this question format would measure an individual's ability to assess both the content and context of a given situation.

The sequence of the questions within each scenario was in order of difficulty, with the first question being relatively easy and proceeding onto more difficult questions. Similarly, the scenarios were arranged according to difficulty, with the first scenario being slightly easier than the following scenarios. This procedure follows standard test practice and was designed to give participants confidence in sitting the test, and also to reduce the likelihood of their wasting time on items they found difficult at the expense of easier items they could complete correctly (Anastasi & Urbina, 1997).

A short demographic survey was also included to collect the following information: Job title, years of industry experience, highest academic qualification, the type of industry currently employed in, age and gender.

Pilot Testing

After extensive revision and amendments to the proposed measure, a pilot study was conducted to evaluate the measure's face and content validity, as well as issues of user-friendliness (e.g., ease of use, format, consistency, and language used). A convenience sample of four participants took part. These participants were not part of the final sample, but all held jobs at a similar level to the participants who would be administered the final measure. The pilot Numerical Business Acumen measure consisted of the six scenarios, with between four and five questions each. The object of including more questions than the test required was to provide a small item pool, from which the questions perceived to be of poorer quality could be eliminated. Each participant sat the measure individually, in a setting with minimal disruption, and was given unlimited time to complete the test. Each participant took

approximately an hour to complete the six scenarios. Once each participant had completed the measure, feedback was elicited through a series of open-ended questions, regarding any perceived ambiguity in either the scenarios or associated questions, any industry-specific jargon used, and general ease of use. Feedback was also elicited regarding the overall credibility of the scenarios and the types of skills the participants felt they used in completing the measure.

As a result of this feedback, several grammatical changes were made to the measure to increase clarity and consistency within and across scenarios, a number of response choices were amended or eliminated, and the number of questions associated with each scenario was reduced to four (by eliminating the questions deemed to be of poorer quality).

Prototype Development

Due to the subjective nature of the questions used in this measure, it was decided that the most appropriate method for determining the 'correct' answer to each question would be through expert consensus. This corresponds with Haladyna (1997), who suggests that an expert, using a graded response such as a rating scale, should judge more complex cognitive skills. Similarly, Wagner and Sternberg (1987) used an expert sample of highly successful and experienced managers to determine correct responses to their Tacit Knowledge Inventory, and used those responses to form a prototype which was used to assess participants' responses. Moreover, experts can provide more accurate estimates of employment test validities than small-sample criterion-related studies. Thus, the validity of a cognitive test can be accurately estimated by expert judges if criterion-related validity studies are impractical (Hirsh, Schmidt & Hunter, 1986; Schmidt, Hunter, Croll, & McKenzie, 1983).

The three experts used in this research project were chosen according to their reputation, in either the business or academic environment, for possessing the key attributes of the Numerical Business Acumen competency. One judge was the managing partner of a global accountancy firm, the second was the national manager of a high-profile retail chain, and the third judge was a university lecturer whose field of research is applied statistics. It was hoped

that the combination of the judges' background and experience would provide an even spread of numeracy and business acumen skills.

The 'judging pack' included a copy of the scenario booklet, the question booklet, an instruction sheet and an evaluation sheet. Each judge was asked to both judge and critique the test. The judging component involved determining hierarchies of correctness for each question. The critique component involved evaluating the test with regards to possible improvements as well as rating the quality and difficulty of each scenario. Once all judging packs were completed and returned, the judges' responses were collated and an expert prototype was developed. The judges' responses were unanimous for approximately half of the questions, with the correct answers for the remainder of the questions determined by the majority .

In response to the feedback received from both the pilot sample and the judges, the number of scenarios was reduced from six to three (see Appendices D and E for copies of the scenario booklet and question booklet for Form A). This was done to keep the test time at a limit which would be seen as realistic by participating organisations. A time limit of 35 minutes was set, determined by the fact that most participants in the pilot sample averaged 10 to 12 minutes to complete each scenario. The types of data included in the three scenarios chosen were still considered to be sufficiently varied to cover most industries.

Alternate Form

An alternate form (Form B) of the Numerical Business Acumen measure was developed, in order to test temporal reliability (see Appendices E and F for copies of the scenario booklet and question booklet for Form B). The alternate form was based on the same three scenarios but with sufficient data and additional information altered to ensure that test participants would have to re-assess their response to each question. The data were altered so that the original question item sets could be reused. The 'correct responses' to these questions were determined by the researcher, using the rationales and answers given by the original judging panel as a guide.

Psychometric Properties

In order to assess the psychometric properties of the present measure, the following aspects of reliability and validity were built into the research design:

Reliability

The delayed alternate-form method was used to estimate reliability, rather than the test-retest method. Test-retest reliability involves administering the same test to the same participants on two different occasions. However, this method is prone to practice effects, where the test taker may recall many of their former responses resulting in spuriously high correlations between scores on the two administrations of the test. It has been noted that this is particularly true of tests of reasoning because, once the principle involved and the mode of thinking has been grasped, a correct future response can be reproduced without re-evaluating the problem (Anastasi & Urbina, 1997). In order to reduce this practice effect, a parallel form of the measure was constructed. The resulting reliability coefficient measures a combination of temporal stability and the consistency of scores across test forms (Anastasi & Urbina, 1997).

The internal consistency of both forms of the Numerical Business Acumen measure was estimated, to measure the degree to which the individual test items were correlated. The Kuder-Richardson reliability coefficient formula was used, as it is appropriate for test items scored as correct or incorrect (Cascio, 1998).

Validity

The degree to which the Numerical Business Acumen measure actually measures the underlying construct of Numerical Business Acumen was addressed via the following methods of validation:

Content-related validity

Content-related validation refers to the examination of a measure to determine whether or not it contains a representative and fair item sample of

the behaviour domain being measured (Anastasi & Urbina, 1997). This type of validation is dependent on the behaviour domain being clearly defined, and the item sample used being defined with enough accuracy to enable the user to judge how adequately it represents the item universe concerned (Cascio, 1998). As such, content validity was built into the Numerical Business Acumen measure from the start by devising work-related situations based on the results of extensive consultation with 'subject-matter' experts - the Human Resource Managers from the participating organisations. The final iteration of the measure was also assessed by the expert judges, as part of the overall judging process.

Any deeper content validation was deemed impractical because, although the design of the measure was intended to simulate real-life situations on paper, a thorough job analysis of a wide range of occupations would be needed to demonstrate a close resemblance between the job activities and the measure (Anastasi & Urbina, 1997).

Criterion-related Validity

Criterion-related validity relates to the effectiveness of test scores in predicting an individual's performance in a specified activity (Anastasi & Urbina, 1997). In this research, only concurrent criterion measures were used, because obtaining predictive measures was impractical due to time restrictions, and out of respect for the anonymity and confidentiality of the test participants. Also, research has shown that concurrent validation studies can provide useful estimates of predictive validity (Cascio, 1998).

Performance ratings have been used extensively in the criterion validation of most occupational tests, and can represent a valuable source of criterion data (Anastasi & Urbina, 1997). Supervisor or peer ratings are most common, but were not used in this study in order to protect the anonymity of the test participants. Self-reported abilities have also been correlated with a wide range of criteria (Shore, Shore & Thornton, 1992), although there are obvious concerns that participants tend to inflate their true ability and are

incapable of making accurate self-assessments. Fisher (1989, cited in Shore et al.) found that self-rankings were influenced by pre-existing self-perceptions and, if self-enhancing schemas were most salient, led participants to integrate information about their own behaviour in a selective manner. However, it is possible that asking participants to assess their own strengths and weaknesses for developmental purposes may involve less self-inflation (Shore et al.). Therefore, a self-ranking scale was devised and included as a means of assessing participants' self-perceptions of competency in Numerical Business Acumen. The scale used the previously identified behavioural indicators (see Table 1) and consisted of six levels (Level 1 being the lowest and Level 6 being the highest).

Numerical Business Acumen is a unique and complex construct, being a combination of numerical critical reasoning skills and business acumen. As a result, no commercially available measures were ideal for convergent validation. Therefore, a commonly used critical reasoning test, Saville and Holdsworth's Numerical Critical Reasoning test (NMG2), was chosen because it measures numerical critical reasoning skills, and it was hoped that some level of convergence would be demonstrated between performance on the two measures.

There are two versions of the Numerical Critical Reasoning test available, with NMG2 and NMG3 equivalent in terms of content and difficulty. The NMG2 version was used in the present research, because Sheffield Consulting Group are licensed to administer it, and because it was designed to assess management graduates (as opposed to non-graduates) which corresponded with the sample used.

The NMG2 test is designed to measure an individual's ability to understand and interpret numerical data. The test consists of a booklet containing 35 questions and a data card containing facts and figures presented in various statistical tables [2]. Five or ten multiple-choice response options are

[2] The statistical tables used in the NMG2 contain a range of information about several unrelated companies. The order of the questions is designed to refer the test taker to each individual table in random order.

provided for each question. Test administration cards are provided to ensure standardisation of instructions.

Standardisation data for the NMG2 were collected in 1990 from a sample of 289 final year undergraduates at several tertiary institutions in the United Kingdom (Saville & Holdsworth Limited [SHL], 1991). The internal consistency reliability coefficients (based on Cronbach's Coefficient Alpha) for the NMG2 were an acceptable level for use in comparisons between individuals: .87 for the standardisation sample, and .85 for a sample comprising final-year undergraduates applying for industry roles (the sample size was not reported). The alternate-form reliability coefficient for the NMG2 and NMG3 was .73. The test-retest reliability coefficient for the NMG2 was .84, based on a sample of undergraduate students[3].

Unfortunately, published validation information for the NMG2 is more limited. With regard to content validity, Saville and Holdsworth Limited (1991) claim that job analysis of a wide variety of professional and managerial jobs has demonstrated that numerical critical reasoning is often a key criterion associated with successful job performance. The majority of the performance criteria used in the studies reported appeared to be based on job analysis and should therefore be relatively objective and valid.

Of the empirical criterion-validation studies available, no data were available on the NMG2. However, summaries of validation of the NMG1, a version that has now been replaced by the more difficult NMG3, and the NA4, which is claimed to be parallel to the NMG battery with regard to difficulty, format and content, are available (SHL, 1991). The relationship between NMG1 scores and various measures of job performance varied amongst the reported studies: No significant correlation was found between NMG1 scores and supervisor ratings for a sample of computer systems employees, a small correlation (.21) was found between NMG1 scores and a performance rating for Innovation in a study consisting of 65 surveyors, and moderate correlations

[3] Surprisingly, no information was given about the size of the sample or the time interval between the two administrations.

were found between NMG1 scores and performance ratings for Analytical Reasoning (.29) and Planning (.27) in a study consisting of 38 sales executives. Scores on the NA1 were moderately correlated (in .30 range) with both supervisor ratings of Analytical Ability in a study consisting of 440 bank managers, and with supervisor ratings of 'above average' performance in a study consisting of 93 junior bank managers.

Construct-related Validity

Construct validity is usually established through the gradual accumulation of data through a variety of sources, where a construct is defined simultaneously with the development of a instrument to measure it (Kaplan & Saccuzzo, 1997). Bearing in mind that providing evidence of construct validity is an ongoing process, the following techniques were built into the present research design to provide sources of evidence of construct validity:

- The item content of the Numerical Business Acumen measure was statistically analysed, using extensive item analysis.
- The Numerical Business Acumen measure scores would be correlated with scores on the established NMG2 measure, which would hopefully demonstrate a reasonable level of convergence between performance on the two tests; yet also demonstrate that the Numerical Business Acumen measure is measuring an additional dimension.
- Expert judgement regarding the appropriateness of the test content in representing the Numerical Business Acumen construct was taken into consideration in the development process and any necessary amendments were made.
- Test participants in both the pilot sample and final test sample were questioned after the administration sessions about their perception of the measure: What skills they felt they had used, how the Numerical Business Acumen measure differed from the NMG2 measure, and the credibility of the scenarios.

Results

Data Entry and Quality Assurance

Each participant was allocated a numerical code for identification, which was written on both versions of the Numerical Business Acumen measure, and the NMG2 test. The test responses were then entered into a spreadsheet as follows:

Numerical Business Acumen measure

Each participant's demographic details, and the organisation they belonged to, were coded and entered. Self-ratings of the Numerical Business Acumen competency were also entered into the spreadsheet. Each answer to the four questions for each scenario was coded and grouped separately, giving a possible total of 12 responses. The answers to both versions of this measure were recorded and grouped separately. A random sample of 30% of both versions of the measure were checked for accuracy of data entry. An error rate of 20% was found in the data entry for the ranking questions, and of 5% for the multiple choice questions. Therefore, all of the answers were checked individually for accuracy. All errors found in this process were corrected.

There were four instances of unclear answers given to questions, such as two options circled in a multiple-choice question, or a number being used twice in the ranking questions. It was decided to randomly allocate a single answer in these cases.

The qualitative information gained from the additional question included in each scenario, questioning participants' rationale for their answer to Question 4, was also analysed. Using a process of thematic analysis, eight possible rationales were identified for each scenario. These rationales were relatively specific to each scenario, so they were further clustered into three overall rationales which reflected the underlying process of analysis each participant seemed to use in responding to each scenario. These three rationales involved the analysis and consideration of the following areas: The organisation itself, in terms of past performance and strategies; the external marketplace, which included competitors' strategies and wider socio-economic

influences; and customers' needs and perceptions. These rationales were coded and entered into the spreadsheet. In many cases, participants had included several rationales in their answer, which were recorded separately. Thus, each participant could have from one to three possible responses.

Saville & Holdsworth's NMG2 test

The indicated answers to the NMG2 test were entered in the spreadsheet and linked to each participant's numerical code. The answers to the 35 questions were coded according to which one of the 10 possible answers was chosen. Again, a random sample of 30% of the tests received were checked for accuracy, with an error rate of 10% found which were then corrected. There were no ambiguous or unclear answers.

Scoring

Numerical Business Acumen measure

a) Item scores

The alternate form (Form B) was marked in an identical fashion to Form A. The multiple-choice and ranking questions for each scenario were marked separately. Using the prototype based on the judges' responses, there was one correct answer for each multiple-choice question. A correct response was scored as '1' and an incorrect response scored as '0'. This gave a potential score of 3 for each scenario, and 9 across all three scenarios.

The ranking questions consisted of either five or six possible options to be ranked in a particular order. It was quickly established that although a 'correct' sequence had been ascertained from the judging panel, the possibility of a participant selecting the correct sequence was very low. For example, there were 120 possible permutations for the questions with five possible ranking options, and 720 possible permutations for those with six possible options. Upon further consideration, it was decided that there were several steps which would probably, in real-life situations, be carried out simultaneously, rather than in sequential order. Therefore, several logical groupings were formed for each of the three questions. This gave the ranking

questions in Scenarios 1 and 3 a total of twelve possible 'correct' answers, and Scenario 2 a total of four possible 'correct' answers. Each correct answer was scored as '1', with an incorrect answer scored as '0'. This gave a potential score of 1 for each scenario, and 3 across all three scenarios.

b) Summary scores

Combining the multiple-choice and ranking questions, the total score was out of 12. The following summary scores were computed for both versions of the Numerical Business Acumen measure: The total number of correct answers for Question 1, the total number of correct multiple-choice answers, the total number of correct ranking answers, and the total number of correct answers overall. The number of items each participant attempted, and the proportion of attempted test items which were correct were also computed.

Saville & Holdsworth's NMG2 test

a) Item scores

The NMG2 test item responses were initially marked with the scoring key provided by Saville and Holdsworth, which simply provided a raw total score. However, the scoring procedure was also carried out on the spreadsheet data. In addition to being a double-check on the accuracy of scoring, this gave participants' performance on individual items.

b) Summary scores

The following summary scores were computed for the NMG2 test: The total number of correct item responses, the number of items each participant attempted, and the proportion of attempted test items which were correct.

These data were analysed using SPSS for Windows Version 7.5.

Psychometric properties

Reliability

Internal consistency

The internal consistency of both forms of the Numerical Business Acumen measure was estimated to measure the degree to which the individual test items are inter-correlated. The reliability coefficient for Form A of the Numerical Business Acumen measure was only .36, indicating that the construct was being inconsistently measured across questions.

In total contrast, the reliability coefficient for the alternate form (Form B) of the Numerical Business Acumen measure was high (.91) - possibly spuriously so.[1] In order to determine if this result was also influenced by a restricted range (participants who found the first set of tests difficult may have declined to participate in the second session), the mean Form A score for those who only sat Form A was compared with the mean Form A score for those who sat both Form A and Form B. However, similar means were found for the first group (6.7) and the second group (6.4). The range of total scores was also sufficiently broad, with Form A scores ranging from 4 to 11 with a standard deviation of 1.68, and Form B scores ranging from 3 to 10 with a standard deviation of 1.98.

The internal consistency for the NMG2 test was also computed, giving a reliability coefficient of .91. Interestingly, this result is slightly higher than those reported previously in Saville and Holdsworth's own validation studies.

Item analysis

Quantitative item analysis was carried out in an attempt to identify ineffective test items in both versions of the Numerical Business Acumen measure, with the NMG2 again being used as a comparison. Both item difficulty and item discrimination measures were computed, and the results are shown in Table 3.

---

[1] It is possible that participants' familiarity with the type of test items used may have contributed to the apparent increase in internal consistency.

Table 3

Summary of item analysis for all measures. 'Diff.' denotes item difficulty and is measured by calculating the proportion of participants who answered each question correctly (p). 'Discrim.' denotes item discrimination and is measured by point-biserial correlation (r_{pb}) which is computed between item scores and the overall total score. For each measure, the mean, standard deviation (SD), minimum and maximum scores are shown.

Item	Form A (N=46)		Form B (N=32)		NMG2 (N=46)	
	Diff. (p)	Discrim. (r_{pb})	Diff. (p)	Discrim. (r_{pb})	Diff. (p)	Discrim. (r_{pb})
Mean	.54	.36	.69	.49	.57	.48
SD	.24	.12	.20	.15	.24	.15
Min	.20	.14	.34	.20	.11	.21
Max	.96	.60	.97	.70	.98	.77
1	.78	.14	.71	.48	.96	.38
2	.46	.42	.69	.46	.61	.41
3	.33	.27	.34	.20	.83	.30
4	.48	.36	.95	.67	.96	.30
5	.96	.40	.94	.67	.72	.44
6	.33	.17	.54	.39	.65	.55
7	.20	.34	.44	.30	.70	.61
8	.59	.45	.59	.43	.37	.37
9	.93	.35	.97	.70	.67	.64
10	.43	.60	.71	.50	.98	.21
11	.61	.37	.75	.54	.57	.25
12	.39	.42	.69	.48	.54	.42
13					.74	.26
14					.61	.46
15					.80	.29
16					.87	.35
17					.52	.44
18					.78	.45
19					.65	.41
20					.83	.59
21					.57	.54
22					.63	.63
23					.46	.44
24					.52	.72
25					.46	.66
26					.61	.77
27					.43	.69
28					.30	.65
29					.24	.64
30					.17	.50
31					.28	.54
32					.20	.44
33					.26	.60
34					.20	.57
35					.11	.37

The difficulty of the test items used in the Numerical Business Acumen and NMG2 measures was measured by calculating the proportion of participants who answered each question correctly, and these results are shown in Table 3. Overall, all three measures met the requirements for maximum differentiation, which are including test items with a moderate spread of difficulty, but whose average difficulty level is approximately .50 (Anastasi & Urbina, 1997). However, as the items in Form B were highly correlated, the spread of item difficulty should ideally be wider.

As expected, a significantly higher proportion of the participants passed Question 1 in Form A of the Numerical Business Acumen measure than the other three questions. The ranking questions in both versions of the Numerical Business Acumen measure (Items 3, 7, and 11) were correctly answered by a significantly lower proportion of the participants. It was also interesting to note that the test items for the NMG2 were obviously arranged in approximate order of difficulty, as the standardisation sample for the NMG2 were given unlimited time, rather than the specified 30 minutes, to complete the test (Saville & Holdsworth Ltd, personal communication).

Item discrimination is the degree to which an item correctly differentiates among test takers in terms of the behaviour that the test is designed to measure (Anastasi & Urbina, 1997). Point-biserial correlation was used, because it is the appropriate measure of the degree of relationship between a dichotomous item response and a continuous variable (the total score). These results are also shown in Table 3. The mean point-biserial correlations for Form A was much lower (.36) than for both Form B (.49) and the NMG2 (.48), although the results for Form B should be viewed with caution because they might have been tainted by practice effects.

The item discrimination scores for the NMG2, which was subjected to extensive item analysis in its construction, ranged from .21 to .77. Therefore, these scores were used as an acceptable guideline for the Numerical Business Acumen measure, with all correlations below .30 noted. Individual test items in Form A that were poorly correlated with the total score were the first (.14)

and third (.27) questions in Scenario 1, and also the second question (.17) in Scenario 2. In addition, item-total statistics were calculated to show the improvement in coefficient alpha if individual items were deleted from the test, with the same three questions being the only items which would result in an increase in the overall reliability coefficient from .36 to .38 (Question 3, Scenario 1), .40 (Question 1, Scenario 1) or .41 (Question 2, Scenario 2), respectively.

In order to ascertain the effectiveness of the individual scenarios, item analysis was also performed for each scenario in both versions of the Numerical Business Acumen measure. The results of these analyses are shown in Table 4. Item difficulty scores show the proportion of test participants who correctly answered *all four* of the questions in each scenario. Interestingly, participants scored more highly on Scenario 3 in both versions of the measure, indicating that that scenario was relatively easier than the others - yet this scenario was perceived to be the most difficult in the test construction process.

Table 4

Summary of item analysis for individual scenarios of the Numerical Business Acumen measure. 'Diff.' denotes item difficulty and is measured by calculating the proportion of participants who answered all of the questions correctly in each scenario (p). 'Discrim.' denotes item discrimination and is measured by point-biserial correlation (r_{pb}).

Scenario	Form A (N=46)		Form B (N=32)	
	Diff. (p)	Discrim. (r_{pb})	Diff. (p)	Discrim. (r_{pb})
1	.06	.56	.12	.92
2	.04	.60	.03	.92
3	.19	.67	.41	.58

Alternate form

Both the temporal stability and the consistency of scores across test forms were measured by computing the correlation coefficient between the total scores for Form A and Form B (i.e., for the 32 participants who completed both forms). This reliability coefficient was .27, and was not statistically significant ($p > .05$). Similarly, no significant correlations were found between

versions for either the ranking questions (r =.29) or the multiple-choice questions (r =.30). These findings suggest that the content of Form B was too similar to Form A, resulting in a strong practice effect as participants appeared to be familiar with both the scenarios, and the type of questions asked. The relationship between total scores on both versions is shown in Figure 3. The total scores for Form B were, on average, 23% higher than the total scores for Form A. However, this increase was inconsistent across participants, with the majority having a substantial increase in scores, but others receiving either the same or lower scores on Form B than on Form A.

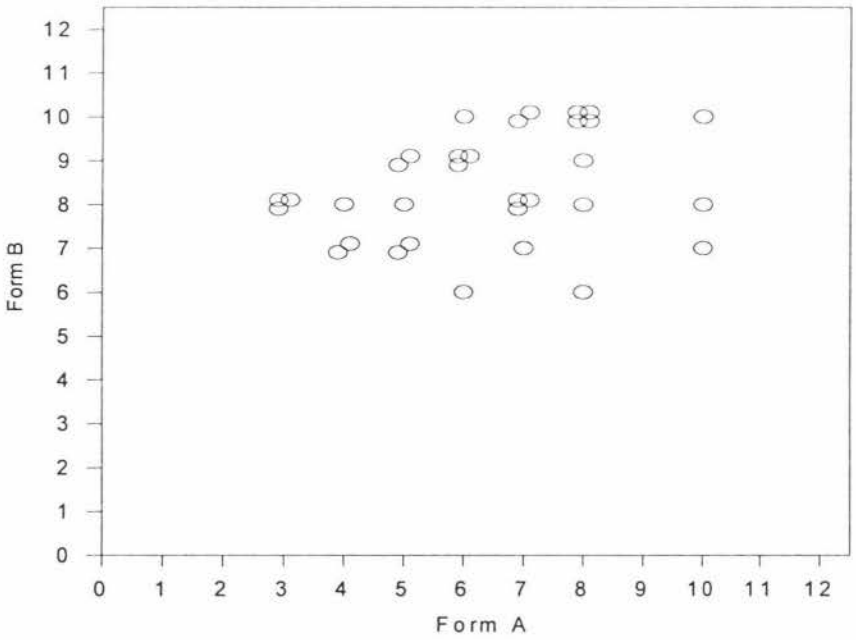


Figure 3

Scatterplot of Form A and Form B total scores for the 32 participants who completed both forms.

Validity

Correlations between the Numerical Business Acumen and NMG2 measures

Correlation coefficients were computed between total scores on each version of the Numerical Business Acumen measure and the comparative measure, the NMG2 (see Table 5). Overall, there were no statistically significant correlations between the total scores on either measure of Numerical

Business Acumen and the NMG2. Although this result is surprising, it may indicate that the underlying components of the Numerical Business Acumen competency are more distinct from the components underlying critical reasoning than originally believed. This corroborates the finding that no significant correlations were found between scores on Question 1 (designed to measure critical reasoning) in either version and the NMG2 scores. Likewise, there were no significant relationships between scores on the ranking questions (designed to measure problem-solving ability) in either form and the NMG2 scores. Lastly, no significant correlations were found between the two questions measuring the business acumen component of Numerical Business Acumen and the NMG2. This result was expected, because it was hypothesised that critical reasoning tests do not adequately measure this component.

As would be expected, significant relationships were noted between the total scores for Form A and scores for both the multiple-choice and the ranking questions for that form. In contrast, the total score for Form B was significantly correlated with the multiple-choice scores, but not the ranking scores. There were moderate correlations between scores for Question 1 in Form A and scores for both Question 1 in Form B (r =.46), and all of the multiple-choice questions in Form B (r =.48). This may indicate that participants found those questions easy to answer.

Table 5
Correlations between the Numerical Business Acumen measures and NMG2 scores

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Multi-choice questions (A)								
2. Ranking questions (A)	.25							
3. Question 1 (A)	.48**	.11						
4. Total score (A)	.92**	.60**	.44**					
5. Multi-choice questions(B)	.29	-.26	.48**	.19				
6. Ranking questions(B)	.11	.30	-.32	.21	-.41*			
7. Question 1 (B)	.19	.00	.46**	.15	.66**	-.32		
8. Total score (B)	.38*	-.07	.29	.27	.78**	.25	.48**	
9. NMG2 total score	-.07	-.09	.21	-.09	.05	-.30	-.26	-.15

** $p < .01$
* $p < .05$

Correlations between measures and self-rankings

Correlations were also computed to examine the relationship between total scores on both the Numerical Business Acumen measure and the NMG2, and the Numerical Business Acumen competency level each participant perceived themselves to be at (see Table 6). No significant correlations were found between total scores for either version of Numerical Business Acumen and participants' self-rankings. However, the total scores for the NMG2 were highly correlated (r =.59) with the self-rankings. This finding is counter-intuitive, but may be explained by the fact that the majority of the behavioural indicators described behaviours demonstrating analytical and numerical ability, rather than business acumen per se.

Table 6
Correlations between self-ranking levels and all total scores

Measure	Self-ranking Level
Numerical Business Acumen measure (Form A)	.01
Numerical Business Acumen measure (Form B)	.01
NMG2 test	.59**

** $p < .01$

Descriptive statistics

Descriptive statistics for both versions of the Numerical Business Acumen measure and the NMG2 are presented in Table 7. 32 (i.e., 70%) of the 46 participants who sat the first version of the Numerical Business Acumen measure also sat the alternate version.

The mean for the alternate version, Form B, of the Numerical Business Acumen measure was much higher than the mean for Form A. The frequency distributions for Form B shown in Figure 5 also shows a negative skew. The frequency distributions for Form A and NMG2 scores are relatively normal in shape, and are shown in Figures 4 and 6.

Table 7

Descriptive statistics for all measures. The statistics given for each measure are the sample size (N), mean (M), median (Mdn), standard deviation (SD), and range of scores.

Measure	*N*	*M*	*Mdn*	*SD*	Range
Test measure					
Numerical Business Acumen (Form A)	46	6.48	7.0	1.88	3 to 11
Alternate form					
Numerical Business Acumen (Form B)	32	8.37	8.0	1.24	6 to 10
Comparative measure					
NMG2	46	20.0	20.0	7.18	5 to 34

Note: The total number of test items for the Numerical Business Acumen measure is 12 and the total number of test items for NMG2 is 35.

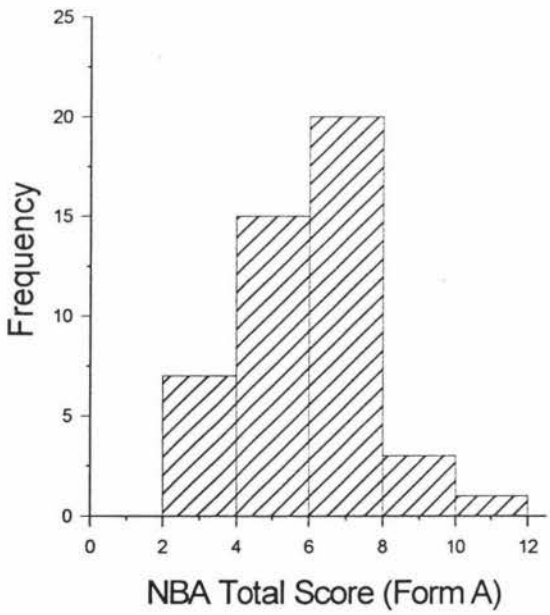


Figure 4
Frequency distribution for total scores on the Numerical Business Acumen measure (Form A)

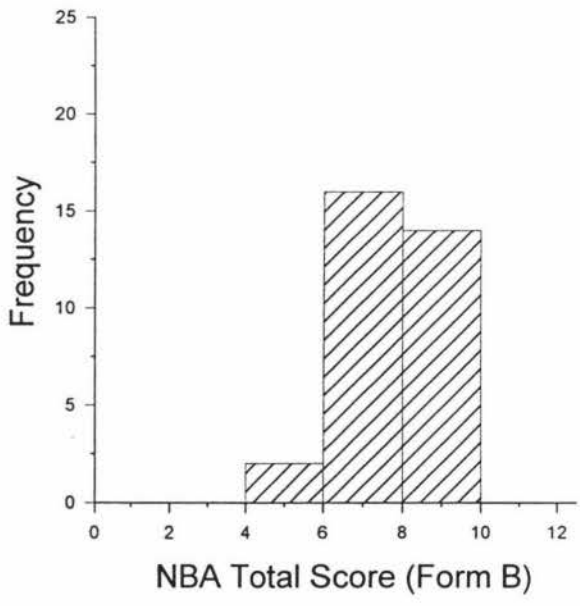


Figure 5
Frequency distribution for total scores on the Numerical Business Acumen measure (Form B)

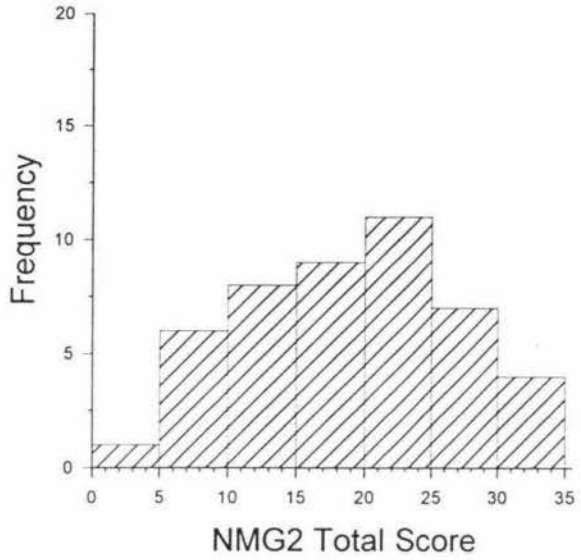


Figure 6
Frequency distribution for total scores on the NMG2 test

Demographic characteristics

Comparisons between industries

Because there were six different industries represented in this sample from the seven participating organisations, an ANOVA was performed to compare outcome measures across organisations. Whereas a significant difference was found between organisations for the NMG2 total scores ($F = 4.6$, $p < .01$), the results were statistically nonsignificant for both versions of the Numerical Business Acumen measure ($F = 0.68$, $p > .05$ for Form A; $F = 0.69$, $p > .05$ for Form B). However, this may in part have been due to the small sample size for some organisations. Therefore, for the sake of completeness, means were compared to see if there were any suggestive differences across the industry groups (see Table 8).

Overall, the means for each organisation for the Numerical Business Acumen measures were relatively similar, with all of the means falling within one standard deviation either side of the overall mean. However, several industries did stand out. The mean for one division of the fast-moving consumer goods industries was higher than most of the other organisations in the Numerical Business Acumen measure, yet substantially lower on the

NMG2 than all of the other organisations. In contrast, the mean for the organisation in the financial services industry was the lowest on the Numerical Business Acumen measure, yet was amongst the highest for the NGM2. The other organisation that stood out was in the retail industry, which had higher means than the overall group on both measures.

Table 8
A comparison of the mean Numerical Business Acumen (NBA) and NMG2 scores between organisations in different industries. The relevant sample sizes are shown in parentheses after the means.

	NBA (Form A)		NBA (Form B)		NMG2	
Organisation	*M*	*SD*	*M*	*SD*	*M*	*SD*
Fast-moving Consumer goods (1)	7.2 (11)	1.9	8.8 (6)	1.3	13.2 (11)	5.1
Fast-moving Consumer goods (2)	6.0 (4)	0.8	9.0 (4)	1.4	24.5 (4)	4.9
Energy	6.3 (3)	1.5	-	-	25.7 (3)	1.5
Airline	6.5 (13)	2.3	7.8 (7)	1.7	19.8 (13)	6.8
Human Resources Consulting	6.2 (9)	1.9	8.4 (9)	1.1	20.3 (9)	6.3
Retail	7.5 (2)	0.7	8.0 (2)	0.0	28.5 (2)	3.5
Financial services	5.3 (4)	1.5	8.0 (4)	1.2	25.5 (4)	6.7
Overall	6.5 (46)	1.9	8.4 (32)	1.2	20.0 (46)	7.2

Note: There were two companies in the fast-moving consumer goods industry which were separate divisions of the same organisation.

Comparisons between job roles

Based on these findings, means were also computed for the various job roles represented in the sample (see Table 9). The job role for each participant was categorised according to the job title given and, where necessary, from the researcher's informal knowledge of the various job roles represented. The participants who had sales roles were the only category to have a higher mean

on the Numerical Business Acumen measure than the overall mean, and also the only job role to have a lower mean than the overall mean for the NMG2. By contrast, the means for both the analyst and accountant job roles were below the overall mean for the Numerical Business Acumen measure, yet well above the overall mean for the NMG2. The fact that participants with a marketing role scored well on the NMG2 was counter-intuitive, because it would be expected that this type of role would require a greater degree of business acumen than critical reasoning skills. However, the number of participants in that category was fairly small and may not therefore be representative of marketing job roles as a general rule.

Table 9

A comparison of the mean Numerical Business Acumen (NBA) and NMG2 scores between different job roles. The relevant sample sizes are shown in parentheses after the means.

	NBA (Form A)		NBA (Form B)		NMG2	
Organisation	*M*	*SD*	*M*	*SD*	*M*	*SD*
Sales	7.2 (11)	1.9	8.8 (6)	1.3	13.2 (11)	5.1
Consultant	6.3 (6)	2.2	8.2 (6)	1.2	21.2 (6)	6.3
Management Graduate	6.4 (12)	2.4	7.8 (7)	1.7	20.6 (12)	6.4
Analyst	5.6 (6)	1.4	8.5 (6)	1.2	23.6 (6)	4.5
Accountant	6.2 (6)	1.5	8.6 (6)	0.8	23.3 (6)	7.8
Marketing	6.5 (4)	1.3	8.0 (1)	0.0	27.0 (4)	2.9
Overall	6.5 (46)	1.9	8.4 (32)	1.2	20.0 (46)	7.2

Comparisons between managerial levels

Because the level of competency in Numerical Business Acumen was hypothesised to vary between management levels, mean scores were computed for the following two categories: Those who were in a managerial role, and those were in a non-managerial role (see Table 10). There were no significant

differences between these means for either version of the Numerical Business Acumen measure, or for the NMG2.

Table 10
A comparison of the mean Numerical Business Acumen (NBA) and NMG2 scores between managerial levels. The relevant sample sizes are shown in parentheses after the means.

	NBA (Form A)		NBA (Form B)		NMG2	
Organisation	*M*	*SD*	*M*	*SD*	*M*	*SD*
Non-managerial role	6.5 (27)	2.1	8.3 (20)	1.3	19.7 (27)	6.9
Managerial role	6.5 (17)	1.6	8.4 (12)	1.2	21.2 (17)	7.7
Overall	6.5	1.9	8.4	1.2	20.0	7.2

Relationships between demographic variables and the other measures

Correlation analyses

Correlations were computed between the demographic variables (e.g., academic qualifications, years of experience) on one hand and the total scores for both versions of the Numerical Business Acumen measure and the NMG2 test on the other. These correlation coefficients are shown in Table 11. There were no significant correlations between the demographic variables and total scores on either of the Numerical Business Acumen measures. However, the total scores for the NMG2 were significantly correlated with academic qualifications (r =.47), self-ranking level (r =.59) and gender (r =-.29, with men receiving higher scores than women on average).

There were several significant, yet predictable inter-correlations between the demographic variables. Age was highly correlated with both years of experience (r =.85) and managerial level (r =.51), and years of experience was also highly correlated with managerial level (r =.52). Interestingly, gender was significantly correlated with both age (r = -.30) and self-ranking level (r = -.45, indicating that the older participants were more likely to be men, and that men seemed to perceive themselves as more highly skilled in the Numerical Business Acumen competency than women.

Table 11

Correlations between demographic variables and total scores for all measures

Variable	1.	2.	3.	4.	5.	6.	7.	8.
1. Academic Qualification								
2. Age	-.04							
3. Years of Experience	-.07	.85**						
4. Self-ranking level	.26	.14	.18					
5. Managerial level	.15	.51**	.52**	.22				
6. Gender	-.05	-.30*	-.28	-.45**	-.16			
7. NBA (Form A) score	-.06	-.13	.05	.00	-.00	-.02		
8. NBA (Form B) score	-.30	-.03	-.02	.00	.03	.22	.27	
9. NMG2 total score	.47**	.21	.26	.59**	.09	-.29**	-.09	-.15

** $p < .01$

* $p < .05$

Regression analyses

Standard multiple linear regression was conducted to examine the how well both the Numerical Business Acumen and the NMG2 scores could be predicted from the demographic information collected about the participants. The following variables were used as predictors: Academic qualifications, age, years of experience, self-ranking level, managerial level and gender.

The proportion of variance in scores for Form A of the Numerical Business Acumen measure accounted for by these predictors was nonsignificant (r =.16)^2. Therefore, the Numerical Business Acumen scores could not be accurately predicted from the selected demographic variables. However, the proportion of variance in scores for the NMG2 test accounted for by the predictors was .63, indicating that 63% of the variance in those scores could be predicted by the demographic variables. Table 12 shows the outcome of this regression analysis. The variables that were statistically significant in terms of their relative importance to the prediction were academic qualifications, self-ranking level and managerial level. Interestingly, the regression coefficient for managerial level was negative, although the correlation between managerial level and NMG2 scores was positive. According to Howell (1992), the occurrence of such a situation indicates that

managerial level is likely to be functioning as a suppressor variable (i.e., it is suppressing variance that is irrelevant to prediction of the dependent variable).

Table 12
Regression coefficients and standardised regression coefficients for the prediction of NMG2 scores.

Variable	Regression Coefficient	Standardised Coefficient
Academic qualifications	2.04*	.44
Age	0.63	.17
Years of experience	1.29	.19
Self-ranking level	3.05*	.52
Managerial level	-4.14 **	-.29
Gender	0.33	.02

** $p < .01$
* $p < .05$

Response accuracy

The number of items attempted by each participant was calculated for all tests. Only 74% of the sample completed all 12 test items in Form A of the Numerical Business Acumen measure, whereas the majority of the sample (88%) completed all 12 items in Form B. Only 15% of the participants completed all 35 items of the NMG2 test, with the majority (72%) attempting no more than 30 items.

Accuracy scores were calculated for the total scores for all tests, as well as for the multiple-choice and ranking questions for Form A. These scores were computed as the proportion of attempted test items each participant answered correctly, and are shown in Table 13. Clearly, respondents were less accurate on Form A than the other measures, with the ranking questions in particular being answered inaccurately. Therefore, although a large proportion of the participants completed all of the test items, only half of those responses were correct. By comparison, although most of the sample did not complete the NMG2, the proportion of attempted items which were correct was substantially higher.

---

[2] Because Form B was primarily used to measure consistency across forms, it was not included in the regression analysis.

Table 13
Proportion correct of items attempted for the Numerical Business Acumen and NMG2 measures

Measure	Accuracy score (*Mean*)
Numerical Business Acumen measure (Form A):	
Multiple-choice questions	.59
Ranking questions	.47
Total score	.57
Numerical Business Acumen measure (Form B):	
Total score	.70
NMG2 total score	.73

In order to address the question of possible speed versus accuracy tradeoffs, the accuracy scores of both Form A of the Numerical Business Acumen measure and NMG2 were plotted. Figure 7 shows the total scores for the Numerical Business Acumen measure correlated with the total number of items attempted. The distribution of these scores suggests that because the majority of participants completed the test, but with a relatively high degree of inaccuracy, that it was either regarded as a speeded test, or that the test items were too difficult and resulted in guessing.

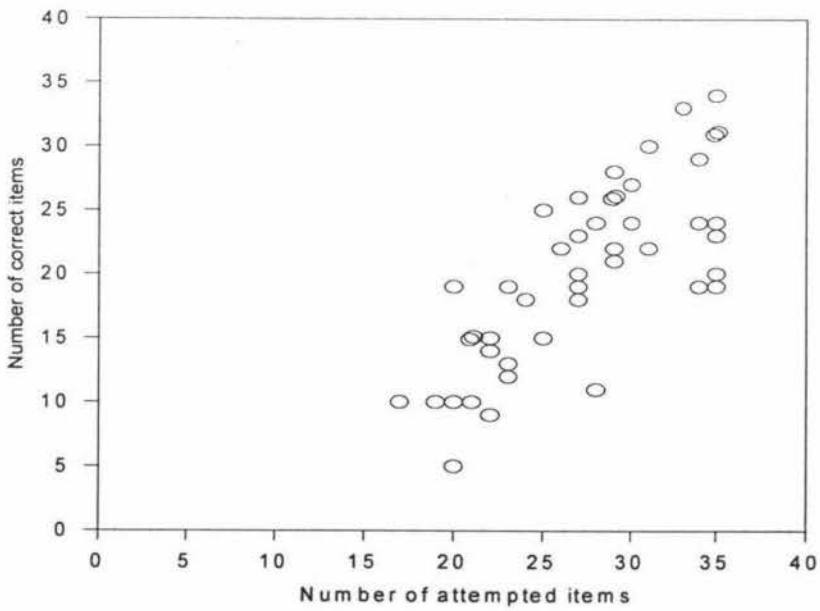


Figure 7
Accuracy scores for Numerical Business Acumen measure (Form A)

Figure 8 shows the total scores for the NMG2 test correlated with the total number of items attempted. Several clusterings were identified: Those participants who attempted relatively few questions but with a high degree of accuracy, those who attempted most or all of the questions but with less accuracy, and those who attempted all of the questions with a high degree of accuracy. These clusterings indicate that, as is often the case with speeded tests, the test-taking strategy for the majority of the participants was primarily focused on finishing the test as opposed to ensuring that they were giving the correct response to each question.

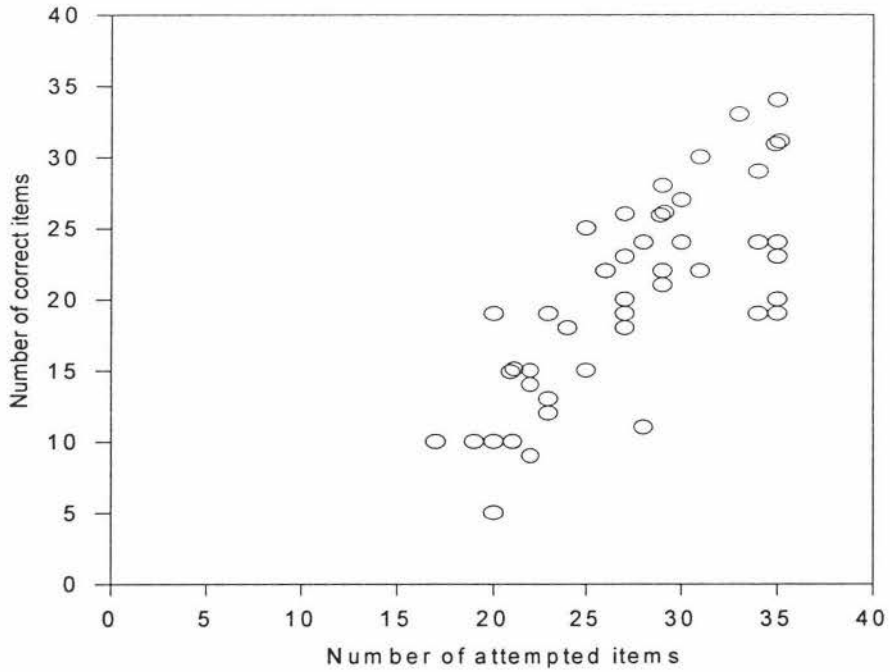


Figure 8
Accuracy scores for NMG2 test

Qualitative comments

The frequency of the three rationales used to explain answers to Question 4 for each of the three scenarios were determined, and these data are shown in Table 14. Each rationale was used fairly evenly, with consideration of the organisation's performance and strategy (39%) being stated slightly more often than consideration of the external environment (30%) and the customers' needs (31%).

Table 14
Frequency with which various rationales were used in answering Question 4. (N=78).

Rationale	*Frequency*	% of responses	% of cases
Organisational performance and strategy	82	38.9	182.2
External environment	63	29.9	140.0
Customer needs/perceptions	66	31.3	146.7
Total	211	100.0	469.8

Note: The percentage of cases added to more than 100 as many subjects reported more than one rationale for their answers. There is a possible total of 405 cases if all participants reported three rationales.

These rationales were also cross-tabulated with the total scores for Form A of the Numerical Business Acumen measure. No significant differences in the type of rationales used corresponded with the number of correct answers. There was a relatively even distribution of all three rationales used by the test participants, with the previously identified focus on organisational performance being reflected in all test scores. However, the high scorers (over 75%) appeared to use all three rationales more equally, perhaps indicating that consideration of all of those areas is an aspect of the Numerical Business Acumen ability. However, the fact that there were only four participants in the high score group limits any possible generalisation.

Discussion

The purpose of this research was to develop a measure of the Numerical Business Acumen competency, and to assess its psychometric properties. As such, this was an initial, and somewhat exploratory, study into this area, which should be taken into account when reviewing the results. The psychometric properties of the Numerical Business Acumen measure are discussed, followed by a discussion of how the results of this study have contributed to the understanding of this construct.

How reliable is the Numerical Business Acumen measure?

Internal consistency

The Numerical Business Acumen measure exhibited lower-than-acceptable levels of reliability. Firstly, the estimate of internal consistency was only .36, indicating that approximately two-thirds of the variance in the Numerical Business Acumen measure could be attributed to errors of inconsistency in content sampling. Thus, the degree to which the individual test items in this measure are inter-correlated is relatively low, presumably because the Numerical Business Acumen construct is being measured inconsistently across those questions.

The low estimate of inter-item consistency could also have been influenced by error variance attributable to the heterogeneity of the construct being measured. Higher inter-item consistency is associated with more homogeneous behaviour domains (Anastasi & Urbina, 1997). This is one possible explanation for why the NMG2, which included items measuring only critical reasoning, showed greater inter-item consistency (r =.91) than the Numerical Business Acumen measure, which consisted of test items designed to measure components of both critical reasoning and business acumen.

Although homogenous test scores allow relatively unambiguous interpretation and higher internal consistency, a single homogenous test would obviously not be an adequate predictor of the heterogeneous construct

underlying Numerical Business Acumen. Most general intelligence tests consist of heterogeneous items designed to measure heterogeneous criteria, but those test items usually consist of several relatively homogeneous tests which each measure a different aspect of intelligence (Anastasi & Urbina, 1997). This method could be used in improving and enhancing the Numerical Business Acumen measure. A combination of homogenous item sets designed to measure numerical critical reasoning and business acumen separately would allow for unambiguous test interpretation while adequately measuring the construct. Of course, this would require a greater number of test items. In turn, this could improve the measure's internal consistency by providing a wider sample of the content domain, and, presumably, a wider range of scores (Cascio, 1998).

Item analysis

The total test scores derived from the Numerical Business Acumen measure provided limited information about the relative components being measured. An item analysis revealed which items of the Numerical Business Acumen measure participants had performed well on. For example, Question 1 in each scenario was answered correctly by the majority of participants (78% to 96%), which was expected as the sequence of questions was in order of difficulty. However, this question was also the only item designed to measure critical reasoning, and, as such, the low difficulty level made it hard to determine whether a correct answer given was due to a high level of critical reasoning or not.

The proportion of the participants who correctly answered the other questions was evenly spread, with 40% to 60% of the participants providing correct answers. However, the majority of participants found the ranking questions difficult in two of the three scenarios, with only 20% to 33% providing correct answers. These ranking questions were also difficult to score, as many permutations were possible, resulting in the probability of someone providing the correct answer in order of one particular sequence

being very low. However, this type of question format was used successfully by Wagner and Sternberg (1986) in their Tacit Knowledge Inventory to measure the ability to assess the context, as well as the content, of a real-life situation. Therefore, this type of question should be retained, subject to further item development.

Overall, the level of item difficulty for both versions of the Numerical Business Acumen measure, and also the NMG2 test, met the requirements for maximum differentiation, as stated by Anastasi and Urbina (1997). There was a moderate spread of difficulty among test items, with an average difficulty level of approximately .50, meaning that, on average, 50% of the test sample answered the item correctly. However, both the critical reasoning and the ranking items received extreme scores. Further development of these test items could result in a more moderate spread, but should still produce a level of difficulty ensuring a wide range of scores.

In contrast, the degree to which the individual test items correctly differentiated among participants in terms of their competency in Numerical Business Acumen was relatively low. As noted previously, the item discrimination score obtained for Form B should be viewed with caution because of likely contamination by practice effects.

Three questions were identified as being poorly correlated with the total score in the Numerical Business Acumen measure. One of these questions contained jargon that could be relatively specific to sales and marketing roles and, although pilot testing indicated that it was not perceived as such, the answers selected for this question indicated that some participants may not have understood the jargon used. The other two questions were identified by the judging panel as containing some ambiguous response options. One of these questions contained a distractor that was in fact correct if participants did not consider the overall situation, and the other question perhaps contained a greater number of ranking options than was necessary, with some options having some overlap in content. Consequently, if these three items were deleted, the overall internal consistency of the measure may be somewhat

improved. Rejecting items that have low correlations with the total score provides a way of purifying or homogenising a measure, because only those items with the highest average inter-correlations are retained. However, this practice only increases validity when the item pool is measuring a single characteristic and when this characteristic is present in the criterion (Anastasi & Urbina, 1997). In this instance, the Numerical Business Acumen measure was designed to assess a combination of behaviours and skills, and the criterion was the total score itself; so purifying the measure by this method could in fact reduce the measure's validity.

An alternative would be to modify and improve the content of those test items to reduce any ambiguity. Consider the following facts: The current Numerical Business Acumen measure, even having deleted these items, would have a reliability estimate of approximately .40. The level of reliability desirable for a measure used to compare one individual to another is a minimum of .90 (Cascio, 1998). Accordingly, the current Numerical Business Acumen measure still has a long development process ahead in order to meet this requirement. In addition, even if these three items were deleted, the maximum item discrimination score found was only .60. By comparison, the item discrimination scores for the NMG2 test ranged from .21 to .77. Therefore, a better option could be to address the ambiguities of the above test items, and invest more time in modifying and improving the current test items. Further test validation would hopefully demonstrate an increase in reliability. The current test item sets, combined with feedback received from test participants, could also be used as a basis for the development of additional items.

Accuracy scores

Although the results of the item analysis indicated that the difficulty of the items used was satisfactory, they also raised the issue of speed versus power. As a result of the pilot testing, a time limit was determined that would allow most participants the opportunity to attempt all of the questions, yet still set an adequate ceiling. In this way, the measure was designed to include

elements of both power and speed, in that the average person would have sufficient time to complete the test yet the questions would be graded in order of difficulty (Cascio, 1998).

Although imposing a time limit does not signify a speed test (Anastasi & Urbina, 1997), the test participants appeared to view the Numerical Business Acumen measure as such. Seventy-four percent of the test participants attempted all of the questions, but only half of the responses were correct. This relatively high degree of inaccuracy could be attributed to the difficulty of test items, resulting in participants guessing their answers. Therefore, the Numerical Business Acumen measure can be characterised as calling on both speed and power.

Item analysis of the scenarios

Each scenario was analysed as a unit to determine if there were certain scenarios that were easier than others, or that seemed to provide a more reliable measure of the Numerical Business Acumen construct. In comparison to the NMG2 test, in which items were ordered in terms of approximate difficulty, the order of the scenarios in the Numerical Business Acumen measure appeared to be the opposite, with the easiest scenario being the last one. The proportion of participants who provided correct answers to all of the questions associated with Scenario 3, in both versions, was significantly higher than the proportion of participants who correctly answered all of the questions in the other two scenarios. In fact, Scenario 3 may have been too easy, as suggested by the resulting discrimination coefficient for Form B being significantly lower (r =.58) than the coefficients for Scenarios 1 and 2 (r =.92 in both cases). In fact, that scenario may well have simply been testing the ability of participants to apply similar cognitive processes to similar situations as much as the construct it was designed to measure.

Feedback received from the participants suggested that the first scenario was often perceived to be more difficult than the others. Therefore, changing the order of the scenarios could increase the confidence of future test-takers,

with Scenario 3 positioned as the first scenario in the set, because it was slightly easier than the others.

Interestingly, the proportion of test participants who answered all of the questions correctly in both Scenarios 1 and 3 doubled in Form B, yet the proportion for Scenario 2 remained much the same in the second form. This finding is somewhat counter-intuitive when considering the effects of practice on those scores, but could be explained by the previously mentioned design fault in the second question. This fault resulted in another response option, in addition to the 'correct' answer, being 'technically' correct. Because only one answer was marked as correct, some of the participants in fact chose the 'technically correct' answer, which was scored as incorrect.

In summary, although the reliability estimate for the current Numerical Business Acumen measure was relatively low, deleting the poorly correlated items may not increase reliability because the Numerical Business Acumen construct is heterogeneous in structure. A more effective method of improving the test design could be to develop further items based on the question format presently used. The result would be a combination of homogenous item sets designed to measure numerical critical reasoning and business acumen separately, which would allow for unambiguous test interpretation while adequately measuring the dual aspects of the construct. This type of compromise has also been suggested by Anastasi and Urbina (1997) as a way of achieving breadth of coverage. By segregating those items of low internal consistency into homogenous sub-sets rather than discarding them, fairly high internal consistency would be attained for those item groups.

Reliability over time and individual performance

The delayed alternate-form method was used to measure both temporal stability and consistency of scores across parallel test forms. The resulting reliability coefficient of .27 indicates that less than a third of the variance in test scores between forms was due to systematic variance in the Numerical Business

Acumen competency. Over two-thirds was due to error variance, attributable to both fluctuations in performance across forms, and fluctuations over time.

It is possible that a strong practice effect occurred amongst participants, with total scores for Form B being on average 23% higher than the total scores for Form A. This may have been due to the parallel form being too similar, because many participants observed that it seemed easier than Form A. Reasons given for this were familiarity with the scenarios, the types of questions asked, and the cognitive processes required to answer the questions. Guidelines given by Anastasi and Urbina (1997) on appropriate construction for alternate forms state that test items should be independently constructed, but still meet the same specifications of the first test. Both tests should have the same number of test items, be identical in format, and cover the same type of content; the range and level of difficulty should be equal; and other aspects of the test such as administration, instructions, time limits should be identical (Anastasi & Urbina, 1997).

In reality, there were practical difficulties in constructing a truly equivalent alternate form to the Numerical Business Acumen measure. As this construct was not well defined, as opposed to say, simple mathematical ability, the measure was more difficult to develop and time constraints did not allow for the development of a totally independent alternate form. Therefore, the scenarios used for Form A also formed the basis for Form B. Changes were made to the facts presented in each scenario, and much of the data within the graphs and tables were also altered, in order to present a scenario involving the same company but in a different situation, and facing different challenges. The response options were designed to be applicable to both versions of the measure, but the correct answer differed across versions. Therefore, the test items were identical in format and in the content covered. The correct answers for the ranking question in each version were deemed to be the same because this item was measuring each participant's ability to analyse and solve problems, which one would expect to be consistent across situations. Interestingly, the correlations between the versions of the Numerical Business

Acumen for the ranking questions was not significant (r =.29), indicating that perhaps that ability does vary across situations.

It may also be that the nature of the behaviours being assessed in the Numerical Business Acumen measure are subject to a large practice effect because participants apply similar problem-solving, analytical and strategic thinking skills to each form and, as such, the use of alternate forms would likely reduce but not eliminate this practice effect. Anastasi and Urbina (1997) suggested that changing the content of test items in a test of problem-solving skills involving the same reasoning process to answer all the questions (such as the Numerical Business Acumen measure) is unlikely to eliminate practice effects being carried over from the first form (Anastasi & Urbina, 1997).

If all the test participants had shown the same improvement over forms, the correlation between their scores would not have been affected as a constant amount would have been added to each score (Anastasi & Urbina, 1997). However, although the average total scores increased by 23% over forms, this increase was inconsistent across participants. Although the majority showed a substantial increase in scores from Form A to Form B, some participants received either the same or lower scores the second time. This may indicate that their test performance was affected by other variables such as motivation, familiarity with the cognitive processes required, previous experience and knowledge in this area, and other factors. Any differences in administration, due to distractions or seating conditions, or differences in the individual taking the test, such as concentration levels, would also have contributed random error.

The error attributed to fluctuations over time may also have been affected by the time interval over which the two forms were administered. This interval, which varied between two and four weeks, appeared to be too short to offset the effects of practice. In general, the time interval between test administrations should be long enough to offset practice effects, which can be anything from one day to six months. (Cascio, 1998).

The observed strong practice effect amongst participants appears to have contaminated the estimate of internal consistency for Form B, which was very high (r =.91). As the design and item format of the parallel form was identical to Form A, one would expect that both forms would product similar correlation coefficients as they are measuring the same construct. In fact, the internal consistency for Form B was a similar level to the coefficient for the NMG2, which measures a relatively homogeneous behaviour domain. It was also established that there was no evidence of participants declining to sit the alternate form as a result of finding the first measure too difficult. In fact, the mean total score for those who sat both Form A and Form B was slightly lower than the mean for those who just sat Form A, indicating that many of the participants who performed well in Form A did not sit Form B. Therefore, it seems clear that participants generally performed better on the alternate form simply because they were familiar with the scenarios and the types of questions asked, which in turn contaminated the inter-item scores - not because the individual test items correlated well with each other per se.

The range of individual differences within the sample was considered for any restriction of variance, which could affect the magnitude of the reliability estimates. Overall, the sample appeared sufficiently heterogenous in nature, covering a wide range of ages, years of experience, and occupational backgrounds. Because it was hypothesised that those in managerial roles would be most likely to demonstrate Numerical Business Acumen in everyday situations, the majority of the sample consisted of managers. This may account for the fact that most of the participants possessed tertiary qualifications, which could, in turn, imply that this sample had a higher-than-average level of cognitive ability. However, the variability in scores indicated that the relatively low reliability coefficients could not be attributed to any restriction in variance caused by cognitive ability.

How valid is the Numerical Business Acumen measure?

Face validity

The Numerical Business Acumen measure did appear to have a certain amount of face validity. Test participants were asked at the end of the first round of administration sessions, during which they sat both the Numerical Business Acumen measure and the NMG2 test, which test they felt was more applicable to real-life situations. The majority of participants perceived the Numerical Business Acumen measure to be more relevant than the NMG2.

Content and construct validity

Although the measurement of both content and construct validity is somewhat dependent on the behaviour domain underlying the Numerical Business Acumen competency being clearly defined, these methods of validation can themselves provide useful sources of evidence to add to the understanding of this construct. Both content and construct validity were built into the design of the measure, by using the knowledge of subject-matter experts to both devise scenarios based on real-life situations, and to serve as judges of the appropriateness and representativeness of the test items.

In order to gather information that would further define the Numerical Business Acumen construct, the participants also provided feedback regarding their perceptions of the Numerical Business Acumen measure. The three scenarios were generally seen as credible and relevant to issues faced in real-life, although several participants observed that, if those scenarios did occur in real-life, they would not have total confidence in their decisions because of the inadequate information provided regarding the context of the situation. Several participants also questioned the relevance of the ranking questions, as well as suggesting that open-ended questions would be a better way of addressing the Numerical Business Acumen construct. As content validity depends more on the relevance of participants' test responses to the behaviour domain being measured, than on the apparent relevance of item content

(Anastasi & Urbina, 1997), the adequacy of the range of test responses can be important.

Many of the participants felt that the Numerical Business Acumen measure and the NMG2 test required different skills. The consensus was the NMG2 test primarily measured numerical ability and critical reasoning skills, whereas the Numerical Business Acumen measure also required the use of a certain amount of business acumen and conceptual thinking.

Criterion-related validity

The degree to which the Numerical Business Acumen measure accurately measures the underlying construct was addressed primarily by means of concurrent criterion validation. Ideally, the criterion for this type of test would be at least current, if not future, job performance. However, the time constraints of the project, and the confidentiality issues that were raised regarding the collection of performance measures meant that this was not possible.

Self-ratings of Numerical Business Acumen

Performance ratings collected from either supervisors or peers have been used extensively in the criterion validation of many tests, and would have been a valuable source of criterion data in the present research. However, most of the Human Resources managers within the organisations taking part in the present research were reluctant to provide access to this type of confidential information. Therefore, self-ratings were included as a means of exploring self-perceptions of Numerical Business Acumen competency. Because the participants understood that the Numerical Business Acumen measure was being used for research purposes only, it was concluded that the problems of self-inflation of abilities reported with self-ratings in previous research (Shore et al., 1992) would be minimised.

Interestingly, there was no relationship between the total scores for either version of the Numerical Business Acumen measure and the self-reported level of proficiency in Numerical Business Acumen. This corresponds

with previous research which indicates that individuals tend to have a significantly different perception of their own job performance from that of other sources (Thornton, 1980). Although self-perceptions are often self-inflating, self-ratings used for developmental purposes tend to be more candid, and therefore less lenient (Shore et al., 1992). This self-perception stems from the pre-existing schemas held by the individual about their attributes. These internal cognitions, in turn, are likely to affect their self-evaluation (Shore et al.). Therefore, it is possible that the test participants were unsure of their ability in the Numerical Business Acumen competency, because it has not previously been used as a performance measure, or possibly even formally distinguished within the organisations as a key skill for some roles.

It is possible that the significant relationship between the NGM2 test scores and self-ranking levels resulted from the fact that the behavioural indicators used to describe the various levels of Numerical Business Acumen were predominantly examples of numerical and critical reasoning. As the descriptions of behaviour were based on previous research (Spencer & Spencer, 1993), there was a scarcity of competencies, or even behaviours, relating to business acumen. This may have been due to both the lack of academic literature on business acumen, and the recent emergence of its importance in the current business environment. Therefore, the additional behavioural indicators of business acumen skills, which were created by the researcher, may have been insufficient in number to provide a balanced and accurate portrayal of the Numerical Business Acumen construct.

This bias in the Numerical Business Acumen levels of proficiency could also account for the lack of relationship between the total scores on the Numerical Business Acumen measure and participants' self-ratings. Because the behavioural indicators of Numerical Business Acumen were not related to the total scores on that measure, it would appear that they were not measuring the same construct - or that one of these tools was not measuring the Numerical Business Acumen construct accurately. The predominance of numerical and critical reasoning behavioural indicators used, together with the

high correlation between self-ratings and the NMG2 scores, indicates that the Numerical Business Acumen scale was perhaps a more accurate measure of numerical critical reasoning than of Numerical Business Acumen. Therefore, one could conclude that the Numerical Business Acumen measure may be measuring a construct separate from numerical critical reasoning.

Numerical critical reasoning and Numerical Business Acumen

This supposition was tested by correlating the total scores of both versions of the Numerical Business Acumen measure with the total scores of the NMG2 test. The lack of any significant relationship (r = -.09; p >.05), may indicate that some of the underlying components of the Numerical Business Acumen competency are more distinct from the components underlying critical reasoning, than originally believed. The theoretical construct underlying the Numerical Business Acumen competency was hypothesised to consist of a combination of numerical critical reasoning skills and conceptual thinking using business acumen. It was also hypothesised that scores on a test which measured a major component of that construct (numerical critical reasoning) would be at least moderately correlated with Numerical Business Acumen scores, thus demonstrating that the two tests were measuring somewhat related constructs.

The results of the present research indicated that the Numerical Business Acumen measure is not assessing a similar construct to numerical critical reasoning, but rather one that appears to be quite separate. This is corroborated by the non-significant relationship between scores on Question 1 in both versions (designed to measure critical reasoning) and scores on the NMG2. Similarly, the results confirmed the hypothesis that no relationship would be found between (a) scores on the ranking questions in both versions (designed to measure the ability to assess the context of a situation) and the NMG2 scores; and (b) scores on the two questions measuring the business acumen component of Numerical Business Acumen and the NMG2 scores; because current critical reasoning tests do not adequately measure those abilities.

These findings correspond with the results of other research which have examined the relationship between tacit knowledge (an important component of business acumen) and cognitive ability tests (of which critical reasoning is commonly a sub-set). The majority of correlations between IQ and tacit knowledge scores were small in terms of effect size. For example, Wagner and Sternberg (1985) found a correlation between verbal reasoning test scores and tacit knowledge scores of .16 (p >.05); Wagner (1987) found a correlation between the same pair of variables of -.12 (p >.05); and Wagner (1994) reported a correlation between scores on an IQ test and tacit knowledge scores of -.14 (p >.05). The explanation offered for this relative lack of a relationship between tacit knowledge and verbal intelligence was that the tacit knowledge measure was tapping skills which are distinct from IQ (Wagner & Sternberg, 1985). In other words, the IQ test was measuring 'academic' intelligence, whereas the tacit knowledge measure was measuring 'practical' intelligence.

This corresponds with Barrette and Durivage's (1997) interpretation of the correlation found between scores on their measure of tacit knowledge and scores on a logical reasoning test of -.13 (p <.05). It was suggested that their tacit knowledge measure was gathering specific information not measured by the logical reasoning test. The authors also noted that although the correlation was not strong enough to infer that tacit knowledge depends on a high level of reasoning, a certain level of cognitive ability appeared to be necessary in the acquisition of tacit knowledge. In other words, 'academic' and 'practical' intelligence are inter-related.

How appropriate was the NMG2 as a comparative measure?

Before considering the implications of these results, it is important to evaluate the suitability of the NMG2 as a comparative measure. This test was used primarily because it was accessible, appeared to possess an acceptable level of reliability and validity, and, most importantly, was the only numerical critical reasoning test commercially available. However, further examination of the accompanying Manual and User's Guide (Saville & Holdsworth, 1991), as

well as comments made by academic reviewers (Impara & Conoley, 1996) revealed that the NMG2 has a number of limitations.

Firstly, the development of the norms appears to have been based on convenience sampling, which calls into question the generalisability of these 'norms' to other managerial positions (Impara & Conoley, 1996). Secondly, it is emphasised in the Manual and User's Guide that the NMG2 is suitable for use in the selection process. However, the reliability coefficients reported were .84 to .87, which could be considered a little below what is recommended (.90 and above) for making important individual decisions (Cascio, 1998). However, this level of reliability would be acceptable if the NMG2 was used as part of a test battery for selection, and its results were considered in conjunction with other selection techniques. Thirdly, and perhaps the most important limitation; despite claiming that numerical critical reasoning is a key criterion associated with successful job performance, no actual evidence of criterion validity was given. An earlier version of the NMG2, the NMG1, produced mixed results as a predictor of job performance. Furthermore, there is little evidence of construct validity, which is surprising, given that critical reasoning is a theoretical construct, and is claimed to be an important component of managerial positions (Impara & Conoley).

Despite these limitations, the NMG2 is a reliable instrument; it was the only test reviewed measuring numerical critical reasoning (as opposed to just critical reasoning); and its poor validation would not have affected this study as measures of job performance were not available to be used as criterion measures. In addition, reviews of the NMG2 (Impara & Conoley, 1996) suggest that this test may actually assess aspects of academic intelligence rather than job performance. Therefore, the level of convergence in performance on both the NMG2 and the Numerical Business Acumen measure could provide a useful example of the level of convergence between 'practical' and 'academic' intelligence.

Theoretical implications

The following theoretical implications can be drawn from the present research, although it is important to note that these implications are tentative due to the questionable psychometric quality of the Numerical Business Acumen measure. Although some of these implications are speculative, they are worthwhile as they may help elucidate some of the important concepts that underlie the Numerical Business Acumen construct.

Numerical Business Acumen was defined earlier as the interpretation of numerical data within varying contexts. Following on from this, it was hypothesised that each of the three identified levels of data interpretation builds on and requires proficiency at the lower level(s). Conversely, the results of the present research suggested that an individual may not have to possess critical reasoning skills in order to demonstrate behaviours of Numerical Business Acumen, because many participants scored highly on the Numerical Business Acumen measure, but not on the critical reasoning test.

Moreover, it could be argued that the lack of any statistical relationship shown between the scores on the Numerical Business Acumen measure and the NMG2 suggests that the Numerical Business Acumen measure may be assessing a construct totally separate from any type of critical reasoning. Therefore, unless any of the underlying aspects of Numerical Business Acumen measure are related to critical reasoning (which will be discussed later), one would not expect any correlation between two tests measuring separate constructs.

This raises the question: If the Numerical Business Acumen measure does not assess critical reasoning, what does it measure? In the initial conceptualisation of Numerical Business Acumen, the construct was hypothesised to consist primarily of numerical critical reasoning and business acumen. Therefore, could it be measuring aspects of business acumen, such as job experience and tacit knowledge? Years of job experience was not significantly correlated with scores on the Numerical Business Acumen measure, which may indicate that the level of competency in Numerical

Business Acumen does not automatically increase with each year. However, only the quantity, as opposed to the quality, of job experience was measured. Previous research has shown that qualitative aspects of experience, such as the level of management attained, appear to be more important to job performance than the quantity of experience based on years served (Barrette & Durivage, 1997; Schmidt et al., 1986; Wagner & Sternberg, 1985). In the present research, years of experience was highly correlated with managerial level, which could indicate that some qualitative aspects of job experience were measured indirectly. However, although the overall quality of job experience may be high, only some of those experiences may result in the acquisition of tacit knowledge.

It was hypothesised that Numerical Business Acumen would measure a combination of aspects of 'practical intelligence' as well as 'academic intelligence'. However, the lack of a relationship between scores on the Numerical Business Acumen measure and academic qualifications may indicate that only aspects of practical intelligence are being assessed. Also, the NMG2 scores were highly correlated with academic qualifications, indicating that this test may be primarily assessing skills learnt in academic settings. This supports Saville and Holdsworth's own research, with NMG2 scores correlating more highly with university graduates' scores on ability tests than with managers' scores (Impara & Conoley, 1996). Therefore, the lack of any relationship between the Numerical Business Acumen measure and the NMG2 may suggest that the Numerical Business Acumen measure is primarily assessing aspects of business acumen - knowledge which is gained in real-life settings as opposed to academic settings.

In an effort to explore further the cognitive processes underlying Numerical Business Acumen, the rationales employed by participants to explain why a particular strategy was chosen in response to a given situation were considered. Overall, the frequency of the rationales used suggested that many of the participants were aware of the importance of interpreting data within the organisational context. Those scoring highly on the Numerical

Business Acumen measure tended to use all three rationales equally (admittedly, the sample size here was small). This may indicate that superior performance in Numerical Business Acumen depends not only on the ability to interpret data within organisational circumstances, but also on the ability to consider the impact of what's happening within the external environment and customer needs and trends - in other words, the ability to consider 'the bigger picture'.

In summary, the Numerical Business Acumen measure seems to primarily assess an individual's ability to consider the context of a situation before making decisions. This construct appears to be quite separate from the theoretical construct of critical reasoning. Therefore, this measure may be assessing an element of practical intelligence, as opposed to academic intelligence. These findings correspond with other research showing that these two types of intelligence are distinct from each other (Barrette & Durivage, 1997; Neisser, 1976; Scribner, 1986; Wagner & Sternberg, 1985).

However, the fact that critical reasoning was not being assessed by this measure should be viewed in light of the fact that an academic view of critical reasoning formed the construct of the comparative measure, the NMG2. A recurrent theme has emerged in the academic literature on critical reasoning: Critical reasoning can not be separated from the context in which the situation occurs (Brookfield, 1987; De Bono, 1985; Norris & Ennis, 1990; Schon, 1983). It is possible that critical reasoning skills were being assessed by the Numerical Business Acumen measure, but that they were a different type of critical reasoning from that measured in the NMG2. Participants who scored highly on the Numerical Business Acumen measure may have been using a type of critical reasoning, but one that developed more as a result of what they have learnt on the job than through formal schooling.

Sternberg and Wagner (1989) suggested that the thinking skills used in practical intelligence are not necessarily qualitatively different in nature from those used in academic intelligence. Rather, they could be the same skills applied to different information within a different domain, which is stored and

represented differently. Accordingly, the participants who scored highly on the Numerical Business Acumen measure but not on the NMG2 test may be prime examples of 'thinking while doing'. Thus, they may be demonstrating Scribner's (1986) concept of practical thinking, or analysing and testing their understanding of a situation while they are taking action, rather than before taking action, as with traditional critical reasoning methods. This corresponds to De Bono's (1985) concept of 'action thinking', and Schon's (1983) concept of reflection-in-action. That is, the participants may be displaying an ability to reason critically, but not in a way that is recognised by the critical reasoning tests currently available.

Is Numerical Business Acumen generic?

Comparisons between the mean Numerical Business Acumen scores for the six different industries were made to see if there were any differences across the industry groups. Overall, these means were relatively similar, suggesting that the Numerical Business Acumen competency may be relatively generic. However, this does not necessarily mean that those demonstrating this competency within different industries would display identical behaviours. Rather, it is likely that this competency would manifest itself in different forms of behaviours within different job roles, and within different organisations. Therefore, although the behavioural indicators may be universally applicable across industries, the extent to which each behaviour is displayed would depend on the context and content of each situation, the organisational culture and the unique challenges confronted in each industry.

Bearing in mind that the results of this study have shown that the Numerical Business Acumen measure is not measuring critical reasoning to any significant degree, one could infer that the universality of the Numerical Business Acumen competency may be due to the underlying business acumen component (if that is what the Numerical Business Acumen measure is actually measuring). This inference concurs with Wallace and Hunt's (1996) assertion that managers use both their business acumen and their previous experience to

make sense of new issues or environments, allowing them to perform competently under any altered context or changed circumstances. Although a certain amount of business acumen is industry-specific, much of it is generic because it consists of core personality traits, and is therefore portable across situations.

Is Numerical Business Acumen related to managerial level?

It was hypothesised that Numerical Business Acumen would be related to managerial level, because the complexity of strategic thinking skills and business acumen required as one moves up the management hierarchy also increases. This was not proven from the comparisons conducted between those in managerial roles and non-managerial roles, with no differences between the overall means for both groups. This could be explained by the fact that changes in traditional organisational structures have meant that non-managerial roles are often taking on more responsibility and autonomy (McLagan, 1997). Thus, the Numerical Business Acumen competency, as with other competencies, may no longer be a skill only used and demonstrated by those in managerial roles.

Practical implications

Notwithstanding the exploratory nature of this study, these findings have useful implications for future research in this area. The results have contributed to the theoretical understanding of the Numerical Business Acumen construct, although the underlying components are still not clearly defined. It is clear that the Numerical Business Acumen measure is assessing constructs which are quite distinct from the type of critical reasoning skills assessed by typical critical reasoning tests. Yet it remains to be established whether another type of critical reasoning, which can be defined as 'critical reasoning in context', is an important component of this construct. In addition, the construct that is being assessed has tentatively been defined as an aspect of

business acumen, but still requires further investigation to determine precisely what characteristics, knowledge, and skills contribute most.

What other possible components should be considered?

Bearing in mind these developments in the theoretical construct, the current Numerical Business Acumen measure clearly did not assess this construct adequately. It is reasonable to argue that this initial version possesses a relative degree of content validity, because the scenario design and questions were based on examples given by subject-matter experts, and were also assessed for relevance by another group of experts in that area. However, the content was valid only in terms of how the Numerical Business Acumen construct was understood at that stage.

Further research needs to investigate what other behaviours or sub-competencies the theoretical construct may subsume. It has been argued that a major component of Numerical Business Acumen may still be critical reasoning, but a type that is learnt through experience and within the context of a given situation. The frequent use of the three identified rationales suggests that the sub-competencies of organisational awareness and external awareness should also be assessed for relevance. Together, these traits contribute to an individual's ability to 'see the big picture' and to consider the context of a situation before making an appropriate decision. Therefore, this ability seems to be an important component of business acumen, and traditional skills of both analytical and conceptual thinking should also be considered, in order to further elucidate the underlying cognitive processes.

In order to effectively assess these developments in the theoretical construct, a new measure could be required. One option could be to develop a totally different measure from the current one, such as an actual simulation of a given situation where a high level of Numerical Business Acumen would be required. However, scenario-based measures have been proven to be a valid means of replicating real-life situations, and have been used successfully to measure tacit knowledge, as well as practical competence (Barrette & Durivage,

1997; Cascio, 1998; Wagner & Sternberg, 1985). In addition, although the psychometric properties of the current measure require much improvement, the research undertaken so far in understanding the Numerical Business Acumen construct has provided a valuable base for further development. Therefore, the most sensible options would be to either (a) iterate the current measure in an attempt to improve its psychometric properties, or (b) modify the current measure in order to include the newly identified components of the underlying construct. These options are discussed in detail below.

Limitations of the current measure and recommendations for future research

At present, both the Numerical Business Acumen construct and the measure used in this research are somewhat ambiguous. The underlying construct must be better understood before the measure can be refined. One way of achieving this is by attempting to improve the psychometric properties of the current measure. Hence, the following changes should be made in an effort to further assess the psychometric properties of this measure. The relatively poor internal consistency of the current measure could be attributed to the following aspects of the design of the measure: The inadequacy of the question formats used; and the heterogeneity of the construct being measured. The poor stability of scores over time was attributed to the over-similarity of the alternate form which led to a substantial practice effect. Given that it has been established that as the Numerical Business Acumen construct is heterogeneous in nature, the deletion of the three questions identified as poor test items would tend to homogenise the measure, and would therefore not necessarily increase reliability. Furthermore, any increase in reliability would be small, and still far from adequate in achieving a desired level of reliability (a coefficient of at least .80, as recommended by Cascio, 1998).

Therefore, a better option would be to further modify and improve the current items. Firstly, the type of question formats should be retained as multiple-choice questions have been shown to effectively measure both numerical and mental abilities as well as open-ended questions (Anastasi &

Urbina, 1997; Haladyna, 1997), and ranking options have been used in an effective measure of tacit knowledge (Wagner & Sternberg, 1985). However, the quality of the questions used and the associated response options could be improved.

Guidelines for the construction of test items evaluating critical reasoning suggest that multiple-choice questions consist of one best answer and a maximum of four distractors (Haladyna, 1997; Norris & Ennis, 1990). The present scoring procedure, as well as feedback received from participants, indicated that some of the questions did not have one answer that was clearly the 'best' option. Therefore, a careful analysis of each test item would be required, with poor response options deleted and any new, more relevant options added.

In addition to addressing any ambiguity within the questions, the scenarios would also need to be reviewed for ambiguity. Feedback received from both the judges, and those involved in the pilot testing, revealed that the information given in the scenarios was also seen as somewhat ambiguous. The suggestion was made that the questions may not effectively measure business acumen, because it was difficult to make decisions based on the information available. Although this effect was intended, it is possible that an individual possessing a high level of business acumen would refuse to make a decision based solely on the given information. Therefore, further information would need to be included in the scenarios.

The ranking questions appeared to have poor face validity and, in addition, were difficult to score. However, the format of Wagner and Sternberg's (1985) Tacit Knowledge Inventory consisted solely of ranking options associated with a work-related situation, and has proven reliability and validity. Tacit knowledge is still deemed to be an important component of the Numerical Business Acumen construct, and this type of question format has the potential to measure the ability to assess the context of a situation. Therefore, a ranking format should be retained in the current measure but modified to replicate the Tacit Knowledge Inventory more closely .

The current test format could also be improved by including the rationale questions in the scoring system. Norris and Ennis (1990) claim that this type of open-ended request for justification of a particular response is an effective way of taking into account variations in background experience and beliefs. As the three distinct rationales were employed by participants across all scenarios, these questions could increase the internal consistency of the measure. Naturally, issues of objective scoring are raised, but these could be addressed by the measure being scored by two different raters, with all possible correct answers determined by subject-matter experts.

Overall, these different question formats were designed to measure two aspects of Numerical Business Acumen. However, only one question in each scenario was designed to measure numerical critical reasoning, and three were designed to measure different aspects of business acumen. A larger number of test items, divided into homogeneous subsets which measure these two aspects separately, would allow for more unambiguous test interpretation and still adequately measure the construct. It has been shown that accurate measurement of heterogeneous criteria can consist of heterogeneous items, if they are separated into several homogeneous sub-sections, each measuring a different aspect of that construct (Anastasi & Urbina, 1997). However, this must be offset against a desire to keep the items to a reasonable number, in order to ensure that the test is not overly long.

The over-similarity of the alternate form used in the present research resulted in contamination of any estimates of reliability over time and across forms. Truly equivalent alternate forms are difficult to construct (Anastasi & Urbina, 1997; Cascio, 1998). This may be especially true for this measure due to the amount of detail that is required for each scenario. Therefore it would cost both time and money to create a more viable alternate form. A more practical alternative might be to assess the measure's test-retest reliability instead. An adequate time interval between administrations of the measure could be allowed to minimise practice effects. Given that some of the participants in this study sat the alternate form one month after the first version and still exhibited

familiarity with the measure, a longer time period would be required, perhaps in the region of six months, as suggested by Cascio (1998). In addition, the measure could be administered for a third time, with the same time interval between the second and third administration sessions as the first and second, in order to comprehensively assess stability of scores over time. However, it would be difficult to allow for any experience or knowledge gained by the participants which could improve their performance during the time intervals between administration sessions. Individual variation in job experience could be measured as a predictor variable, because it has been shown that tacit knowledge does not necessarily increase automatically with quantity of job experience (Sternberg & Wagner, 1985).

Even if further development ensures a more reliable measure, further validation is required in order to establish exactly what the Numerical Business Acumen measure is assessing. The use of the NMG2 as a comparative measure was the main source of criterion validation. The validity of the Numerical Business Acumen measure could benefit from comparisons with other numerical and critical reasoning measures. It has been ascertained that the NMG2 has high reliability, but there is little evidence of its construct or criterion-related validity. Therefore, it could be measuring numerical ability, rather than critical reasoning. Although the NMG2 appears to be the only commercially available test measuring numerical critical reasoning, a test battery consisting of a proven critical reasoning test (such as the Watson-Glaser Critical Thinking Appraisal) and a numerical ability test could be used as a comparative measure.

If one considers that Numerical Business Acumen is a competency, and that a competency is defined by the underlying characteristics that contribute to superior performance, it should be ascertained which of those characteristics, if any, contribute to job performance. Performance ratings would be measured concurrently with the first administration of the Numerical Business Acumen measure, and further measures of performance collected with each subsequent administration of the measure, in order to measure both the concurrent and

predictive validity. In the present study, ethical considerations ruled out the use of performance ratings as a criterion measure. The use of performance ratings in future research would also need to address the same ethical considerations. An additional issue related to the use of performance ratings is how to compare performance across organisations, as there are no absolute measures of job performance. An individual's performance is usually assessed within the context of each organisation's unique standards, objectives, and goals (Boyatzis, 1982).

These suggestions for improving the psychometric robustness of the current measure would also be applicable to the second option: Modification of the current measure to also assess newly identified possible components of Numerical Business Acumen. The measurement of the additional behaviours of organisational and external awareness, and aspects of action thinking (analytical and conceptual) would be added to the current measure and assessed homogeneously by way of separately grouped test items.

Because the major modification recommended to the current measure consists of measuring the circumstances surrounding a real-life situation, these test items would have to be developed with care. This could be done by asking a wide range of subject-matter experts to think of situations that would involve these characteristics, and then identify possible differences in reactions to those situations, and behaviour outcomes which would reflect different manifestations of these characteristics (Mumford, Constanza, Connelly, & Johnson, 1996). The difficulty of measuring context and the individual experience brought to each situation, and being able to measure this in an objective way, may account for the lack of critical reasoning tests available which measure context (despite several researchers stressing the need for this type of test). However, as managerial roles continue to broaden and change, so will the need for Numerical Business Acumen to be portable across situations, and context-specific knowledge will become less important.

Finally, any further assessment of the reliability and validity of the Numerical Business Acumen measure should also be carried out on a larger

sample. Although the sample used in this study represented a relatively broad cross-section of industries and job roles, the actual size was still relatively small (N = 46), and the selection of those organisations was non-random. Therefore, any generalisations from the findings of this study beyond that of the current sample would be inappropriate.

Uses in Human Resources Management

In its current form, the Numerical Business Acumen measure should only be used for research purposes. In order to be useful in a practical sense, the measure would have to able to produce reliable scores. Thus, use even as a developmental tool should be treated with caution, because it is currently impossible to attribute any increases in scores to an individual's progress within that competency.

Comparisons between job roles suggests that this measure may be able to differentiate between these groups. However, further measure development and validation, such as the modifications suggested above, is required in order for the measure to be a reliable predictor of job performance.

Conclusion

The main aims of this research project were:

- To develop a measure of the Numerical Business Acumen competency; and
- To assess its psychometric properties.

The development of the Numerical Business Acumen measure was an initial step in creating a reliable and valid means of assessing this competency. This aim was an ambitious one, in that the underlying construct was not clearly defined, and there was no evidence of previous research measuring a managerial characteristic that combined traditional analytical skills with the ill-defined concept of business acumen.

The results of this empirical study were:

- The internal consistency of the measure was relatively poor. This was primarily attributed to an insufficient number of test items being used to measure what is essentially a heterogeneous construct.
- Some evidence of discriminant validity was shown because scores on the Numerical Business Acumen measure did not correlate at all with scores on the comparative measure used, the NMG2. This finding indicates that a construct distinct from numerical critical reasoning is being assessed by the Numerical Business Acumen measure.

The results of this research have contributed to the understanding of the components contributing to the Numerical Business Acumen construct. The theoretical and practical implications of these findings are that:

- It has been established that although this construct appears to be distinct from critical reasoning skills based on academic learning, another type of thinking may still form part of Numerical Business Acumen. This type of reasoning is hypothesised to evolve through learning by experience, and takes place within the context of that situation.

- The major component of Numerical Business Acumen appears to be the ability to consider the context of a situation before making decisions and, as such, is an aspect of practical, as opposed to, academic intelligence.
- Suggested improvements to the measure consisted of further modification of the existing questions, and the inclusion of additional questions to measure other components identified as contributing to the Numerical Business Acumen competency.
- The current Numerical Business Acumen measure should only be used for research purposes, as extensive item analysis and further validation is required to improve the psychometric robustness of the measure.

In conclusion, the underlying construct of the Numerical Business Acumen competency is complicated in structure. This may demonstrate its importance to managerial performance, because most managerial tasks involve using several skills simultaneously, or sequentially. In addition, identifying the importance of considering context when making decisions indicates that this measure could add potential value in the assessment process of managers in the future.

MASSEY
UNIVERSITY

A L B A N Y

DEPARTMENT OF PSYCHOLOGY

The development of a measure for Numerical Business Acumen

INFORMATION SHEET FOR HUMAN RESOURCE MANAGERS

My name is Jocelyn Anso and I am a graduate student at Massey University's Albany Campus undertaking my Masters thesis research project. My supervisor is Dr Philip Voss, Lecturer in Psychology, and Anthony Mitchell at Sheffield Consulting Group is acting as a consultant.

The purpose of the research project is to develop a measure of the managerial competency Numerical Business Acumen, and to assess its reliability and validity. Numerical Business Acumen can be described as the knowledge, skills and attributes required when interpreting numerical data in relation to its impact on the wider organisational and external context.

I would like to invite you to consider the participation of your organisation in my research project on Numerical Business Acumen. As the initial contact person for your organisation, it is understood by the researcher that your consent to participate is dependent upon consultation with appropriate others in the organisation.

Once you have identified members of your organisation who have expressed interest in participating in this research project, you will be given Information Sheets and Consent Forms to pass on to those prospective participants. The Consent Forms are to completed at their leisure, to allow them time to consider the Information Sheet.

Once prospective participants have consented to participate, they will be given a series of pencil-and-paper measures to complete. These measures all relate to Numerical Business Acumen, and should take approximately an hour to complete. The researcher will be present at all times whilst the measures are being completed. They will be asked to repeat another version of the Numerical Business Acumen measure again two weeks after the first session, which should take approximately 30 minutes to complete. I will contact you to organise times for these sessions.

Private Bag 102 904, North Shore MSC, Auckland, New Zealand
Telephone +64-9-443 9693 Facsimile +64-9-443 9732

Participation is voluntary, and prospective participants are to be advised that their involvement is independent of any requirements of their position within your organisation. As the initial contact person, you have the right to decline to participate, and to withdraw from the study at any time. You also have the right to refuse to answer any particular questions.

The confidentiality and anonymity of participants is assured as they will be identified only by a code, which will be chosen by themselves. The results obtained will be confidential to the researcher. Neither the organisation nor participants will be identified in any publications resulting from the research. A summary of the research findings will be available from the researcher, when the research project is completed.

Should you have any queries about the content or nature of this test measure, please feel free to contact either the researcher directly by e-mail at j.anso@xtra.co.nz or Dr Voss on 443 9663 or by e-mail at P.J.Voss@massey.ac.nz.

MASSEY
UNIVERSITY

A L B A N Y

DEPARTMENT OF PSYCHOLOGY

The development of a measure for Numerical Business Acumen

INFORMATION SHEET

My name is Jocelyn Anso and I am a graduate student at Massey University's Albany Campus undertaking my Masters thesis research project. My supervisor is Dr Philip Voss, Lecturer in Psychology, and Anthony Mitchell at Sheffield Consulting Group is acting as a consultant.

The purpose of the research project is to develop a measure of the managerial competency Numerical Business Acumen, and to assess its reliability and validity. Numerical Business Acumen can be described as the knowledge, skills and attributes required when interpreting numerical data in relation to its impact on the wider organisational and external context.

I would like to invite you to participate in my research project on Numerical Business Acumen. Should you choose to participate, you will be given a series of pencil-and-paper measures to complete. These measures all relate to Numerical Business Acumen, and should take approximately an hour to complete. The researcher will be present at all times whilst the measures are being completed. You will be asked to repeat another version of the Numerical Business Acumen measure again two weeks after the first session, which should take approximately 30 minutes to complete. I will contact you to organise times for these sessions.

Participation is voluntary, and your involvement is independent of any requirements of your position within your organisation. You have the right to decline to participate, and to withdraw from the study at any time. You also have the right to refuse to answer any particular questions.

Your confidentiality and anonymity are assured as you will be identified only by a code, which will be chosen by yourself. The results obtained will be confidential to the researcher. You will not be identified in any publications resulting from the research. A summary of the research findings will be available from the researcher, when the research project is completed.

Should you have any queries about the content or nature of this test measure, please feel free to contact either the researcher directly by e-mail at j.anso@xtra.co.nz or Dr Voss on 443 9663 or by e-mail at P.J.Voss@massey.ac.nz.

Private Bag 102 904, North Shore MSC, Auckland, New Zealand
Telephone +64-9-443 9693 Facsimile +64-9-443 9732

MASSEY
UNIVERSITY

A L B A N Y

DEPARTMENT OF PSYCHOLOGY

The development of a measure for Numerical Business Acumen

CONSENT FORM

I have read the Information Sheet and have had details of the research project explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I understand that my participation in this research project is voluntary, and that I have the right to withdraw from the research at any time and the right to decline to answer any particular questions.

I agree to provide information to the researcher based on the understanding that this information will remain confidential and will be used only for this research and publications arising from this research project.

I agree to participate in this research project under the conditions set out in the Information Sheet.

Signed: .....

Name: .....

Date: .....

Private Bag 102 904, North Shore MSC, Auckland, New Zealand
Telephone +64-9-443 9693 Facsimile +64-9-443 9732

NUMERICAL BUSINESS ACUMEN MEASURE

Administrator Guide

A. Pre-test Instructions

Thank you for taking part in this research project which is a major component of my Masters degree.

I would like to remind you that participation in this research is voluntary, and your involvement is independent of any requirements of your position within your organisation. You have the right to refuse to participate, and to withdraw from the study at any time. You also have the right to refuse to answer any particular questions.

Your confidentiality and anonymity are assured as you are identified only by a code, which will be chosen by yourself. The results obtained will be confidential to the researcher. Management within your organisation will receive a summary of the aggregated research findings which will not contain any information that could identify an individual.

A summary of the research findings will be available from the researcher, when the research project is completed. Feedback regarding individual results may also be obtained by contacting the researcher.

Please take the time now to read the Consent Form and, if you are willing to participate, sign and date the form in the spaces provided. Please keep the Information Sheet so that if you have any questions you may contact me.

This session will consist of two pencil-and-paper tests. In both these tests, you will be using various statistical data and accompanying information to answer questions designed to assess your ability to analyse and interpret numerical data.

The first test consists of a scenario booklet containing three work-related scenarios, each accompanied by various tables and figures, and a question booklet containing four associated questions for each scenario (giving a total of 12 questions).

You will have 35 minutes in which to complete all of the questions in the measure. Complete as many of the scenarios as possible in the time allocated.

Each scenario has a stated primary objective: this is to be kept in mind when answering the associated questions. As with real work situations, some of the information provided in each scenario may be either incomplete, or irrelevant. If you are not sure of an answer, please make the best choice possible from the information available.

Enclosed in the Question booklet on Page 1 are several general questions regarding your background. Page 2 contains a self-rating scale of the Numerical Business Acumen competency. Please answer these questions as they will add to the understanding of the results of the tests.

There are two types of questions used: a multi-choice format where there is one correct answer, and a ranking format where you will be asked to rank several options. You will also be asked to explain the rationale for your choice of answer in some questions.

Do not write anything in the scenario booklet as it will be used again. Please indicate your answer by filling in the appropriate space in the question booklet. Please fully erase any answers you wish to change.

You may use a calculator for this test. Rough paper is also provided for your working out.

The second measure consists of 35 multi-choice questions, each with one correct answer. (Refer now to Saville and Holdsworth's administrator guide for appropriate instructions)

Please select a code that is meaningful to you e.g. mother's maiden name, license plate number, someone's birthday and write it on the top right-hand corner of both answer booklets. Please ensure that you select a code that you will remember, as you will need to use the same code in two weeks time.

NUMERICAL BUSINESS ACUMEN MEASURE

SCENARIO BOOKLET

Form A

NUMERICAL BUSINESS ACUMEN MEASURE

Instructions

In this measure, you will be using various statistical data and accompanying information to answer questions designed to assess your ability to analyse and interpret numerical data with a view to gaining optimum commercial benefit.

The measure consists of a scenario booklet containing three work-related scenarios: each accompanied by various tables and figures, and a question booklet containing four associated questions for each scenario (giving a total of 12 questions).

You will have 35 minutes in which to complete all of the questions in the measure.

Each scenario has a stated primary objective: this is to be kept in mind when answering the associated questions. As with real work situations, some of the information provided in each scenario may be either incomplete, or irrelevant. If you are not sure of an answer, please make the best choice possible from the information available.

Do not write anything in this scenario booklet as it will be used again. Please indicate your answer by filling in the appropriate space in the question booklet.

You may use a calculator for this test. Rough paper is also provided for your working out.

Scenario 1

You are Sales Manager in the Spirits division of a large liquor group for the Auckland region. Your share of the premium vodka market has dropped by 5 percentage points in the last year (1997). Several promotional launches of your brand 'Wodka' have occurred throughout the year. Any sales data available are shown below.

Primary Objective

Your primary objective is to assess the effectiveness of the promotions used so far in increasing your market share.

Figure 1: Total Sales for 1997

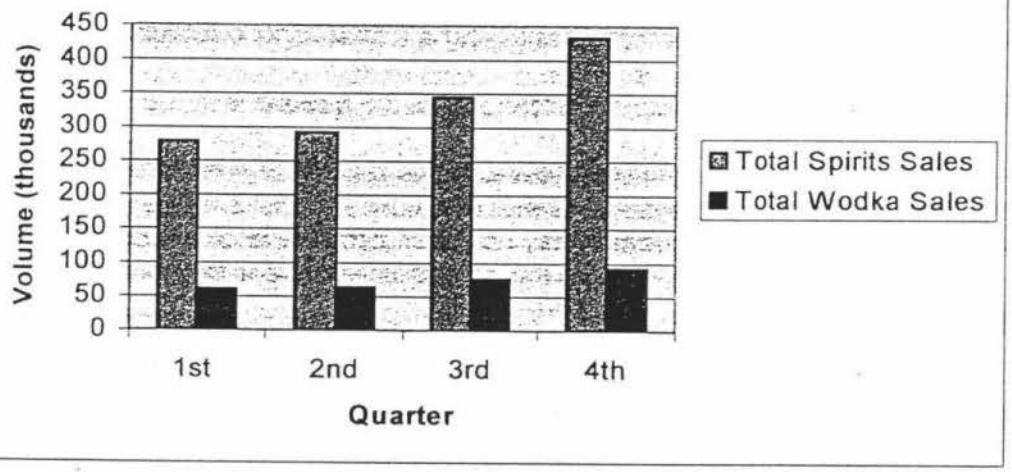


Table 1: Total Sales for 1996

1996		
Quarter	**Total Sales – Spirits**	**Total Sales-Wodka**
1st	275 000	55 000
2nd	268 000	53 000
3rd	315 000	58 000
4th	403 000	93 000

Figure 2: Retail Prices for 1997

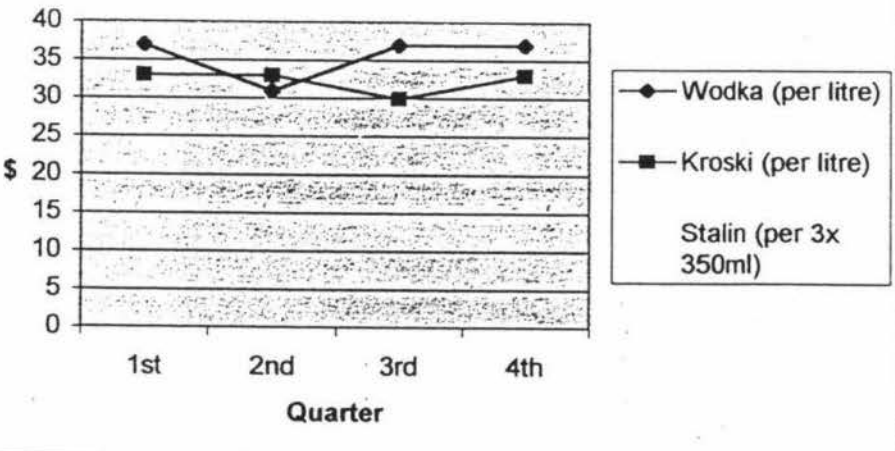
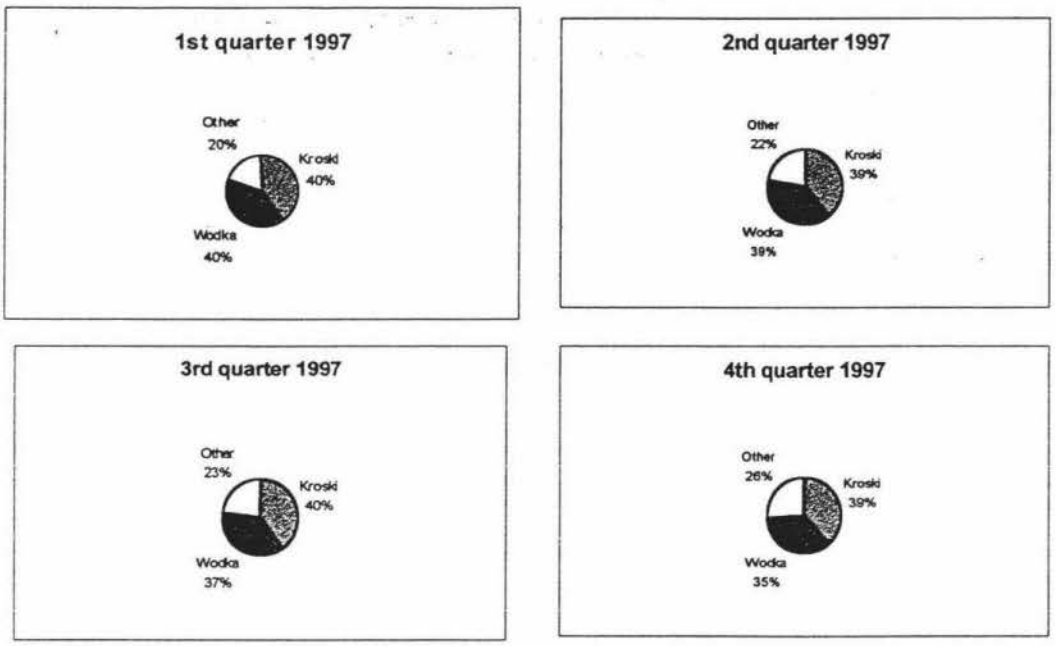


Figure 3: Market Share for 1997



Facts about your company:

- In 1997, there was a Wodka promotion every quarter:
  - A free six-pack was offered with every purchase of Wodka in the first quarter.
  - A 'limited time' price reduction campaign took place in the second quarter.
  - Promotional activity in the third quarter involved a prize draw for a trip to Moscow with every purchase.
  - The usual Christmas campaign took place in the fourth quarter.
- In 1996, there were two promotions of Wodka. The first was in the second quarter, with a 'limited time' price reduction campaign. The second promotion, in the fourth quarter, was the same Christmas advertising campaign that is used every year.
- Promotional costs to the company have been negligible.
- Throughout 1996/97 Wodka promotions were supported by magazine and billboard advertising using the slogan 'Wodka – it brings out the Tsar in you'. This profile of Wodka was consistent with market research results which portray Wodka as a luxurious indulgence.
- You have a team of 5 sales reps: 2 servicing bars and other licensed drinking establishments, and 3 servicing the wholesale retailers.

Facts about your competitors:

- Your strongest competitor, Borin, launched a 'special' price reduction campaign for their premium vodka brand 'Kroski' in the third quarter of 1997.
- There have been rumours that Borin are about to launch subliminal advertising in several popular bars in central Auckland.
- Borin has a team of 6 sales reps: 3 servicing the bars and licensed establishments, and 3 servicing the wholesale sector.
- Other independent companies have been heavily promoting 'alcho-pop' ready mixed drinks. These bottled drinks are now becoming a strongly growing force.
- The pre-mixed vodka and lemon squash drink 'Stalin' is being heavily promoted in bars in the central Auckland area.

Scenario 2

You are the managing partner for a small financial consultancy, Leith Financial; based in Tauranga. Your core business is providing financial advice for clients through one-to-one consultations.

Primary Objective

Your primary objective for the coming financial year is to increase the amount of revenue earnt, while minimising costs. This will be measured by the number of new clients, and also the average revenue gained per employee.

You have obtained your main competitors' profit and loss statements from past company reports (and other sources) to compare with Leith Financial's data. This information is shown in Table 1.

Table 1: Profit and Loss figures

Company	Year ending 31 March 1996	Year ending 31 March 1997	Year ending 31 March 1998
Leith Financial			
Revenue	300 000	400 000	450 000
Expenses	200 000	350 000	420 000
Trust Services			
Revenue	200 000	230 000	220 000
Expenses	140 000	150 000	160 000
Money Minders			
Revenue	250 000	300 000	350 000
Expenses	150 000	180 000	300 000

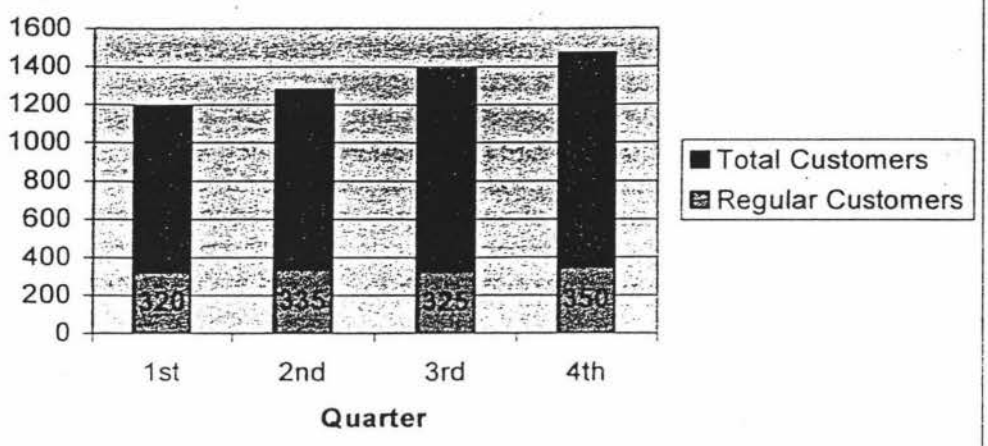
Facts about Leith Financial:

- The company was set up five years ago and is still establishing itself in the marketplace.
- Your clientele is a mix of one-off and regular consultations.
- The primary revenue for the company is obtained from consultation fees.
- The major costs to the company are staff-related (wages, travel etc.)
- The five people you employ as consultants vary widely in their customer service skills and experience within the industry.
- Employees who are new to the financial industry charge a lower hourly fee. This is increased once they achieve the necessary industry knowledge – on average it takes two years to achieve this.
- You have been targeting a niche market not currently serviced by the major financial institutions: providing investment advice to the 25-40 age group.
- Your main advertising strategy has been inserts in the local newspaper, and distribution of flyers at the local tertiary institutions and in the central business area of Tauranga.

Table 2: Employee Revenue Details (Leith Financial)

Employee	Average Fee per hour	Average Monthly Revenue	Years of Industry Experience
Hector	$80	$8000	12
Kerry	$50	$6000	1
Liz	$80	$6400	5
Andrew	$80	$6000	10
Tina	$80	$7200	3

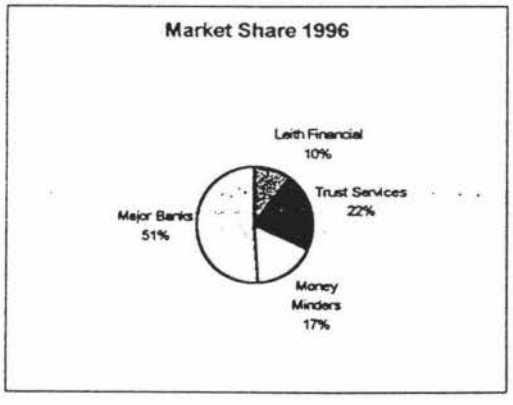
Figure 1: Total customers 1998 (Leith Financial)

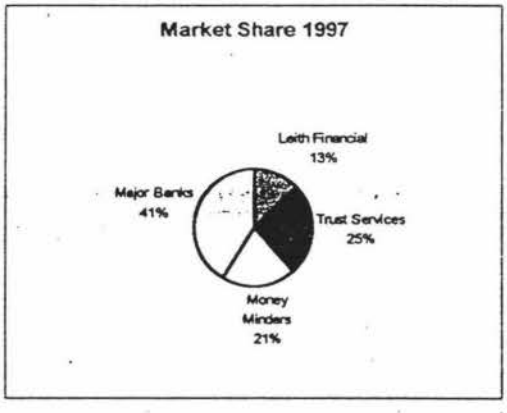


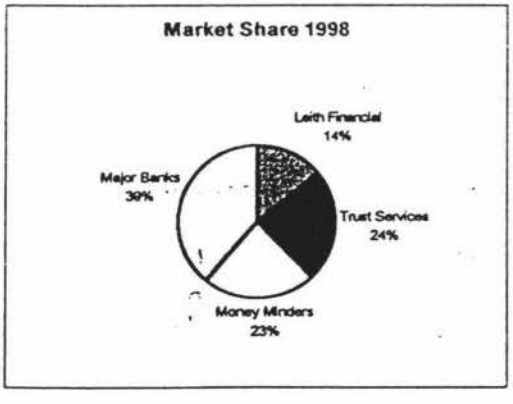
Facts about the competition:

- Free money management seminars are currently being offered by Trust Services. These have apparently been very popular, although you have heard rumours that the hard-sell approach adopted at the conclusion of the seminar has not been successful in winning new customers.
- It has been rumoured that Money Minders may have set up a deal with one of the major banking institutions to receive referrals of all those setting up graduate banking packages.
- The majority of people in the 25-40 age group still consult their bank first if they require financial advice.

Figure 2: Market Share







Scenario 3

You are the Business Development Manager for a home furnishings store called Dwellings. You have several stores located within the Auckland region, which all vary in size and layout; the range of goods available and the service provided. The quality and price of the goods are standard across the stores.
It has been recently noted that your Newmarket store has a significantly higher customer complaint rate than your other stores. A market research survey investigating customer satisfaction was carried out last month (February 1998). The survey questioned the customers' degree of satisfaction with various aspects of all your stores, and asked them to rate how important those aspects were to their continued patronage. Both are measured on a scale of 1 to 5 (1 = Poor/Unimportant, 5 = Excellent/Very important).

Primary Objective
Your primary objective is to analyse Dwellings' current level of customer service orientation and assess its impact on customer loyalty.

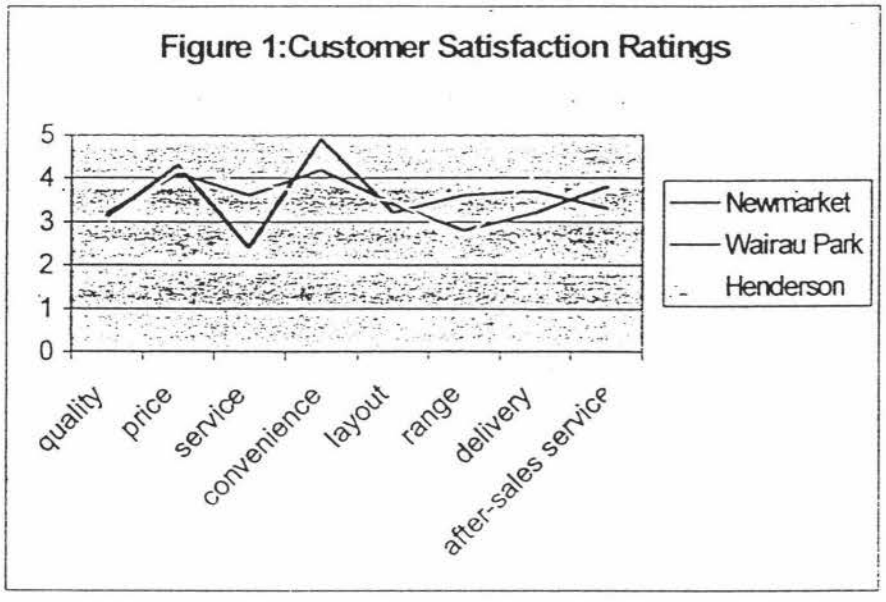


Table 1: Customer satisfaction ratings: Newmarket store (n=250)

Area	Satisfaction (1-5)	Importance (1-5)
Quality	3.1	3.2
Price	4.3	3.4
Service	2.4	4.2
Location/Convenience	4.9	2.0
Layout of store	3.2	2.5
Range	3.6	3.0
Delivery	3.7	1.5
After-sales service	3.3	3.8

- A representative sample of the qualitative comments regarding the Newmarket store are listed below:

'unhelpful'	'lack of knowledge'	'didn't listen'
'no smiles'	'unresponsive'	'bunch of kids'
'desk unattended'	'no gift wrapping'	'had to hunt for someone to help me'

Facts about Dwellings:

- A customer satisfaction survey undertaken in July 1997 indicated that customers rated Dwellings' strongest feature as customer service, followed by price, convenience and then quality.
- Your advertising strategy for the past six months (August 1997 – January 1998) has been 'The best products at the best price'.
- Extra temporary staff were hired in all stores to cope with the Christmas rush.
- Staff turnover is high, with a noticeable increase since a major competitor went into receivership. There have been rumours since then that the home furnishings retail industry is dying.
- Customer loyalty is low, with customers shopping around for the best perceived price.

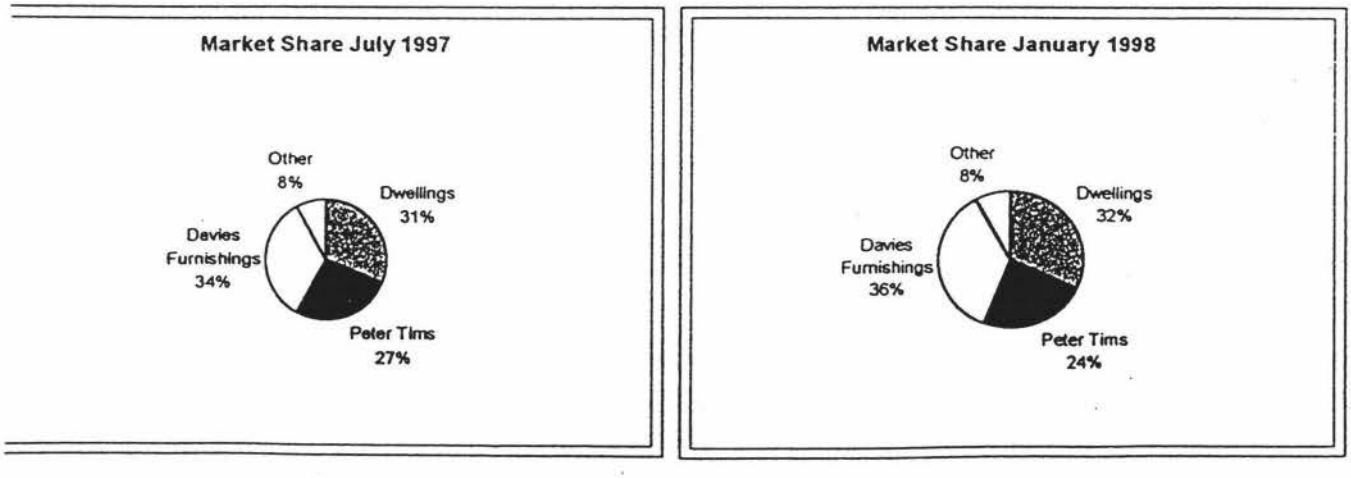
Facts about the Competition:

- Customer satisfaction ratings for your two major competitors, Peter Tims and Davies Furnishings, were also measured in the same survey. These are shown in Table 2.
- Peter Tims has two stores: one in Mount Wellington and one at Wairau Park. Its main advertising strategy for the last six months has been 'We go the extra mile'.
- Davies Furnishings has two super-stores: one in Greenlane and one in Henderson. Its main advertising strategy for the last six months has been 'Customer satisfaction –guaranteed'.

Table 2: Customer satisfaction: Competitors (n=250)

	Peter Tims		Davies Furnishings	
Area	Satisfaction (1-5)	Importance (1-5)	Satisfaction (1-5)	Importance (1-5)
Quality	2.6	3.2	2.4	3.2
Price	4.1	3.4	3.7	3.4
Service	3.9	4.2	4.3	4.2
Location/Convenience	4.4	2.0	3.5	2.0
Store Layout	2.9	2.5	3.0	2.5
Range	2.4	3.1	2.6	3.0
Delivery	2.9	1.5	3.8	1.5
After-sales service	3.8	3.8	4.5	3.8

Figure 2: Market Share



NUMERICAL BUSINESS ACUMEN MEASURE

QUESTION BOOKLET

Numerical Business Acumen is the ability to analyse and interpret data and make sound business judgements as a result. The table below lists the key behaviours for six levels of Numerical Business Acumen.

Using this table as a reference point, please circle on the scale below the level you feel best describes your competency level in Numerical Business Acumen at present:

Level 1 Level 2 Level 3 Level 4 Level 5 Level 6

Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
• Little understanding of numerical principles and limited numerical skills • Only uses 'rules of thumb', common sense and past experience to identify problems • Sees only essential similarities between current and past situations • Responds to situations without analysing impact of actions	• Uses basic numeracy skills • Observes basic discrepancies, trends, and interrelationships in data • Sees crucial similarities and differences between current and past situations • Recognises the implications of patterns and trends shown in data to overall organisational performance	• Understands and uses basic statistical analysis and established methodologies • Can see multiple relationships among several parts of the situation • Recognises several likely causes of events, and consequences of actions between parts of the data • Distinguishes between relevant and irrelevant data • Applies past knowledge to analyse current situations • Determines current problems or opportunities based on the data, with a basic understanding of possible impact on organisational performance	• Uses statistical analysis to manipulate and interpret data • Makes long chains of causal connections, by identifying cause and effect and seeing multiple relationships between parts of the problem • Identifies key issues in a complex situation • Considers the advantages, risks, and implications of any decisions made on the organisation, in evaluating and choosing alternative solutions • Recommends possible courses of action to maximise possible opportunities and minimise potential problems, with the objective of increasing profit.	• Uses statistical tools to manipulate data to reveal trends or predictions not evident without analysis • Sees multiple relationships among several parts of a problem, recognises several causes of events and several consequences of actions • Identifies useful relationships among complex data from unrelated areas, and pulls together unrelated pieces into a comprehensive whole • Extracts larger meaning from data to recommend possible courses of actions to opportunities/problems working within current organisational objectives and long-term strategy • Identifies key opportunities/ problems that could impact on the organisation in the near future	• Uses complex and multi-variate numerical analysis to manipulate and interpret data • Can deduce correlation and causation; generates and tests multiple hypotheses for a situation; identifies useful relationships among complex data from unrelated areas • Identifies several alternative solutions to key opportunities/ problems, developed to correspond with political / competitive environments • Uses data to compare organisational performance to that of competitors; and to recommend ways of increasing competitive advantage • Uses industry knowledge and skills to recommend possible courses of actions working within organisational objectives and long-term strategy

Please complete the following questions by placing a tick in the boxes that relate to you:

What is your current position title?

How many years of business experience (in any industry) have you had?

0 – 5	❒
6 – 10	❒
11 – 15	❒
16 – 20	❒
21 – 25	❒
26 +	❒

What is your highest academic qualification?

MBA	❒
Post-graduate degree (PhD, Masters, Diploma)	❒
Bachelors degree	❒
Polytechnic qualification	❒
Higher School Certificate	❒
University Entrance (or equivalent)	❒
School Certificate (or equivalent)	❒
None	❒

What type of industry are you currently employed in?

Airline	❒
Energy	❒
Fast moving consumer goods	❒
Financial services	❒
Human Resource Consulting	❒
Retail	❒

What age group are you?

20 – 25	❒
26 – 30	❒
31 – 35	❒
36 – 40	❒
41 – 45	❒
46 – 50	❒
51 – 55	❒
56 – 60	❒
61 – 65	❒

What is your gender?

Male	❒
Female	❒

Scenario 1 Questions:

1. How would you best describe Wodka?
a. Low price; price sensitive.
b. High price; price sensitive.
c. Low price; price non-sensitive.
d. High price; price non-sensitive.
e. Can't say from the information given.

2. Which one of the following statements best describes the relationship between total Wodka sales and the promotions undertaken in the last two years?
a. The promotions used in each quarter have had little effect on Wodka sales overall.
b. The 'special price' campaigns in 1996 and 1997 appear to correlate with an increase in Wodka sales for those quarters.
c. The high-value promotion (the trip to Moscow) appears to correlate with an increase in Wodka sales for that quarter.
d. The Christmas campaign seems to be the only one in both 1996 and 1997 that caused an increase in Wodka sales.
e. The increase in promotional launches used in 1997 correlates with an increase in overall Wodka sales.

Keeping in mind that your primary objective is to assess the effectiveness of previous promotions in increasing market share:

3. Rank the following steps in order of execution (from first step to last step) when considering what advertising promotions to use in the coming year:
___ a. Obtain market research regarding your market positioning, and that of your competitors.
___ b. Determine what types of promotions are effective in increasing your market share.
___ c. Determine what increase in market share you are aiming for.
___ d. Devise an action plan for your sales reps to prioritise key customers.
___ e. Ascertain details of the stereotypical Wodka drinker.
___ f. Assess effectiveness of promotions used in 1996 and 1997.

4. Based on the perceived effectiveness of previous promotions, what type of promotional activity should be used for the coming year?
a. Continue to have similar promotions to 1997 launched every quarter at wholesale retailers only.
b. Activity should focus on high-value promotions, such as a prize draw for an Audi (a high-prestige European car), offered in up-market bars.
c. Offer more give-aways in wholesale outlets.
d. Increase your sales team by two reps who will focus on relationship-building with the bars in the central Auckland area.

Please explain the rationale for your choice in Question 4:

Scenario 2 Questions:

1. What appears to be the general trend over time shown in Table 1?
a. Revenue is going down, but expenses as a percentage of revenue are going up.
b. Revenue is going down, and expenses as a percentage of revenue are also going down.
c. Revenue is going up, but expenses as a percentage of revenue are going down.
d. Revenue is going up, and expenses as a percentage of revenue are also going up.

2. How would you best describe the segmentation of Leith Financial's customer base?
a. The ratio of regular customers to total customers has stayed constant over 1998.
b. The greatest increase in total customers in any one quarter was due to an increase in new or one-off customers.
c. The ratio of regular customers to total customers is increasing.
d. The ratio of new/one-off customers to total customers is increasing.

Keeping in mind that your primary objective is to maximise revenue earned at minimum cost:

3. Rank in order of execution (from the first to last step) the following steps to be considered in setting revenue targets for each employee:

___ a. Determine employee performance statistics such as average fee per hour (and any likelihood of change); average monthly revenue; average number of clients.
___ b. Establish incentives for achieving new targets.
___ c. Estimate potential impact of future advertising campaigns.
___ d. Evaluate customer satisfaction ratings for each employee, such as complaints about service and retention figures.
___ e. Consider market research results, and any external factors.

4. Which of the following strategies would be most effective in increasing the number of new customers (who will become regular customers as opposed to one-off customers):

a. Offer money management seminars for $60 per person, with a free financial consultation for each customer.
b. Conduct a survey to ascertain what potential clients are looking for in a financial consultancy and follow up the people surveyed, offering a reduced-price initial consultation.
c. Approach a major bank with a proposal that they outsource their financial advice service to Leith Financial.
d. Offer a customised service for each customer which will involve 6 consultations for optimum benefit, with further consultations at reduced rates.
e. Provide a 'planning for the future' package specifically aimed at couples – this will require several consultations both individually and together.

Please explain the rationale for your choice in Question 4:

Scenario 3 Questions:

1. Which predominant area in the Newmarket store needs to be improved to meet customers' needs?

a. Quality.
b. Range of goods.
c. Service.
d. Location and convenience.
e. Price.

2. What appears to be Dwellings' current market perception overall?

a. High quality, low prices and great service.
b. Poor quality, low prices and average service.
c. High quality, low prices and poor service.
d. Average quality, low prices and average service.
e. Can't say from the information given.

Keeping in mind that your primary objective is to assess the impact of Dwellings' customer service on customer loyalty:

3. Rank the following steps in order of execution (from first step to last step) in implementing a programme promoting superior customer service:

___ a. Consider overall market research results and determine current levels of satisfaction via customer feedback surveys.
___ b. Determine what customers perceive as important in customer service.
___ c. Determine the relationship between current levels of customer satisfaction and customer loyalty.
___ d. Provide incentives to motivate employees to provide superior service.
___ e. Analyse the performance of individual stores.

4. What would your advertising strategy be for <u>all</u> your stores for the next six months?

a. Promote Dwellings' superior customer service, with an emphasis on employees' product knowledge and their readiness to help, using the slogan 'our place is your place'.
b. Use promotions focusing on quality and value.
c. Offer several 'interest-free terms' promotions on all purchases over $500.
d. Use promotions focusing on consistently high quality and value, and improved customer service.
e. Launch a promotion emphasising the locations of your stores – 'they're just down the road!'

Please explain the rationale for your choice in Question 4:

NUMERICAL BUSINESS ACUMEN MEASURE

SCENARIO BOOKLET

Form B

NUMERICAL BUSINESS ACUMEN MEASURE

Instructions

In this measure, you will be using various statistical data and accompanying information to answer questions designed to assess your ability to analyse and interpret numerical data with a view to gaining optimum commercial benefit.

The measure consists of a scenario booklet containing three work-related scenarios: each accompanied by various tables and figures, and a question booklet containing four associated questions for each scenario (giving a total of 12 questions).

You will have 35 minutes in which to complete all of the questions in the measure.

Each scenario has a stated primary objective: this is to be kept in mind when answering the associated questions. As with real work situations, some of the information provided in each scenario may be either incomplete, or irrelevant. If you are not sure of an answer, please make the best choice possible from the information available.

Do not write anything in this scenario booklet as it will be used again. Please indicate your answer by filling in the appropriate space in the question booklet.

You may use a calculator for this test. Rough paper is also provided for your working out.

Scenario 1

You are Sales Manager in the Spirits division of a large liquor group for the Auckland region. Your share of the premium vodka market has dropped by 3 percentage points in the last year (1997). Several promotional launches of your brand 'Wodka' have occurred throughout the year. Any sales data available are shown below.

Primary Objective

Your primary objective is to assess the effectiveness of the promotions used so far in increasing your market share.

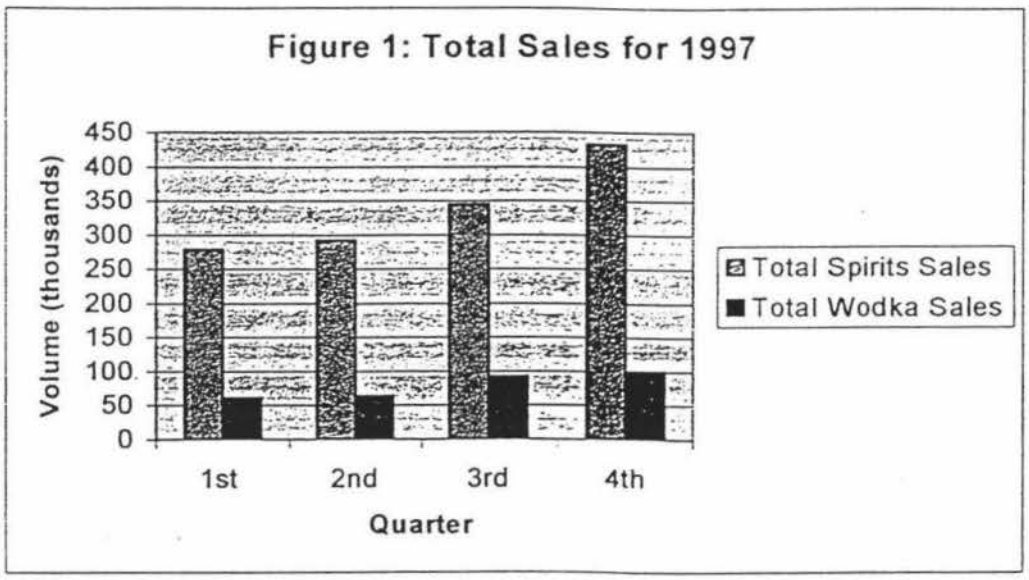


Table 1: Total Sales for 1996

1996		
Quarter	Total Sales – Spirits	Total Sales-Wodka
1st	275 000	55 000
2nd	278 000	56 000
3rd	325 000	61 000
4th	423 000	97 000

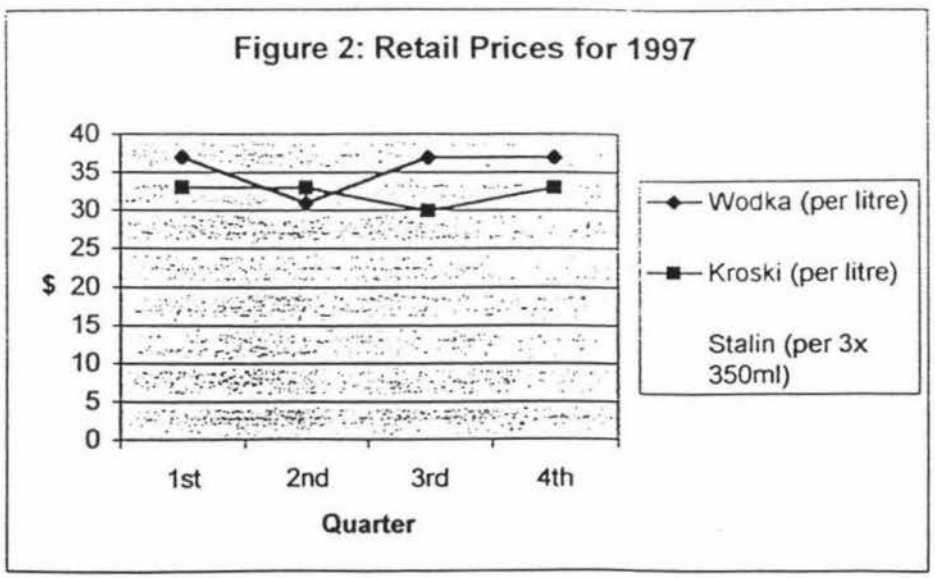
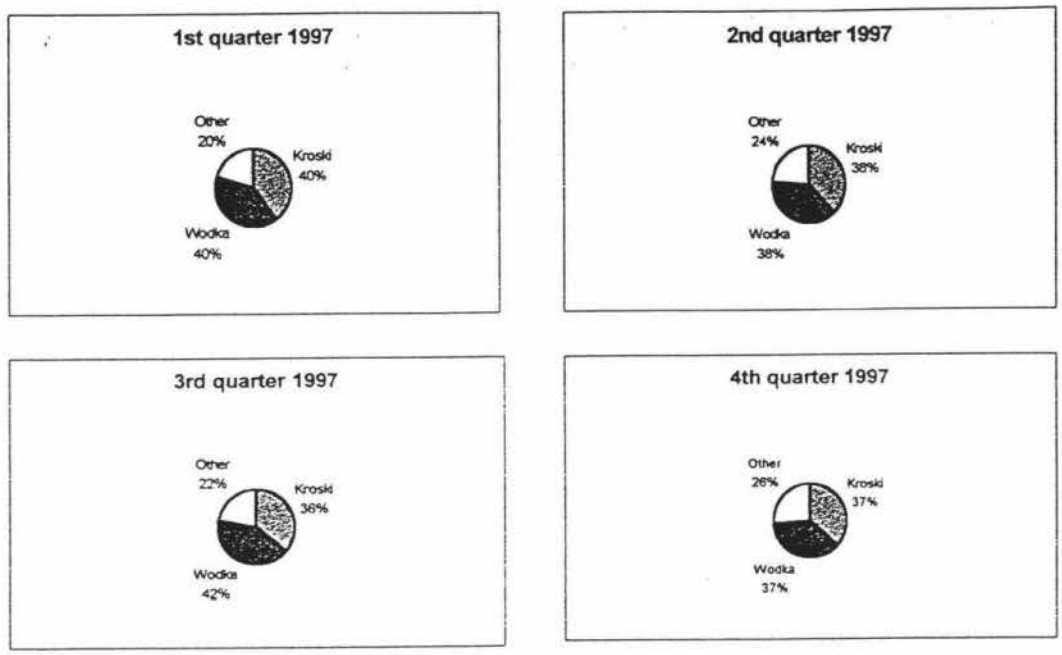


Figure 3: Market Share for 1997

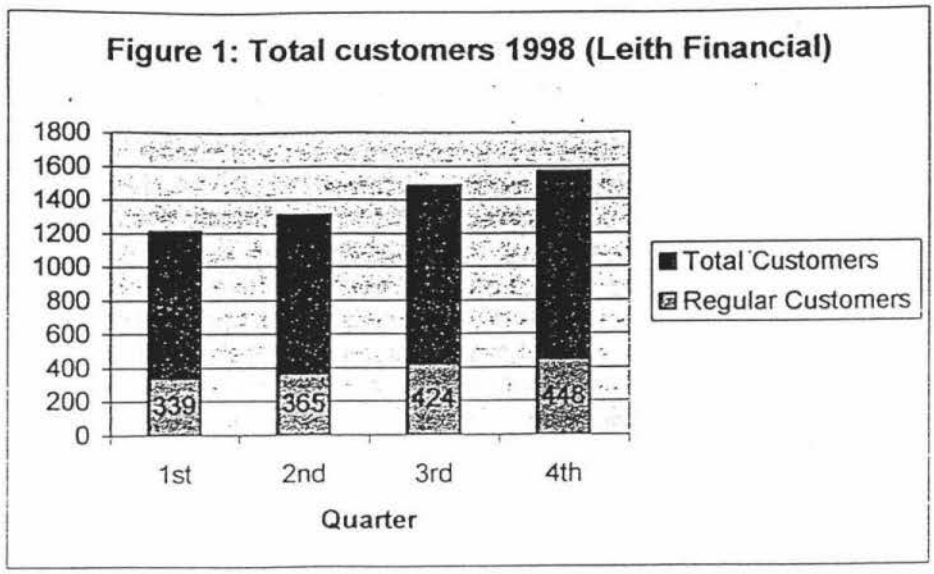


Facts about your company:

- In 1997, there was a Wodka promotion every quarter:
  - A free six-pack was offered with every purchase of Wodka in the first quarter.
  - A 'limited time' price reduction campaign took place in the second quarter.
  - Promotional activity in the third quarter involved a prize draw for a trip to Moscow with every purchase.
  - The usual Christmas campaign took place in the fourth quarter.
- In 1996, there were two promotions of Wodka. The first was in the second quarter, with a 'limited time' price reduction campaign. The second promotion, in the fourth quarter, was the same Christmas advertising campaign that is used every year.
- Promotional costs to the company have been negligible.
- Throughout 1996/97 Wodka promotions were supported by magazine and billboard advertising using the slogan 'Wodka – it brings out the Tsar in you'. This profile of Wodka was consistent with market research results which portray Wodka as a luxurious indulgence.
- You have a team of 6 sales reps: 3 servicing bars and other licensed drinking establishments, and 3 servicing the wholesale retailers.

Facts about your competitors:

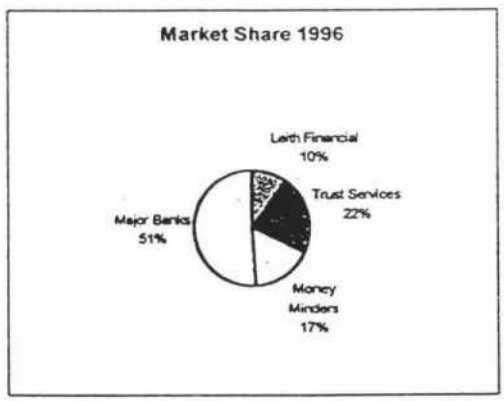
- Your strongest competitor, Borin, launched a 'special' price reduction campaign for their premium vodka brand 'Kroski' in the third quarter of 1997.
- There have been rumours that Borin are about to launch subliminal advertising in several popular bars in central Auckland.
- Borin has a team of 6 sales reps: 3 servicing the bars and licensed establishments, and 3 servicing the wholesale sector.
- Other independent companies have been heavily promoting 'alcho-pop' ready mixed drinks. These bottled drinks are now becoming a strongly growing force.
- The pre-mixed vodka and lemon squash drink 'Stalin' is being heavily promoted in bars in the central Auckland area.

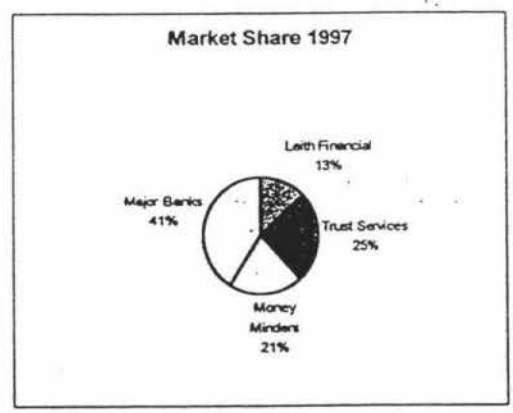


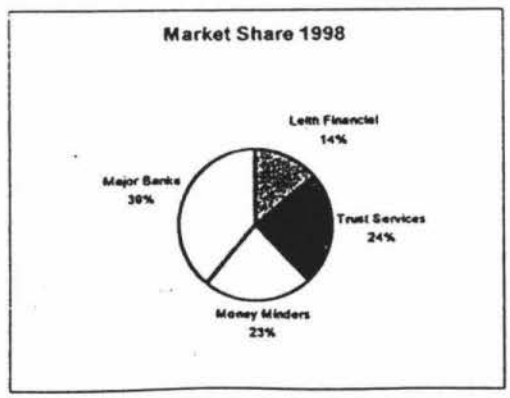
Facts about the competition:

- Free money management seminars are currently being offered by Trust Services. These have apparently been very popular, although you have heard rumours that the hard-sell approach adopted at the conclusion of the seminar has not been successful in winning new customers.
- It has been rumoured that Money Minders may have set up a deal with one of the major banking institutions to receive referrals of all those setting up graduate banking packages.
- The majority of people in the 25-40 age group still consult their bank first if they require financial advice.

Figure 2: Market Share







Scenario 2

You are the managing partner for a small financial consultancy, Leith Financial, based in Tauranga. Your core business is providing financial advice for clients through one-to-one consultations.

Primary Objective

Your primary objective for the coming financial year is to increase the amount of revenue earnt, while minimising costs. This will be measured by the number of new clients, and also the average revenue gained per employee.

You have obtained your main competitors' profit and loss statements from past company reports (and other sources) to compare with Leith Financial's data. This information is shown in Table 1.

Table 1: Profit and Loss figures

Company	Year ending 31 March 1996	Year ending 31 March 1997	Year ending 31 March 1998
Leith Financial			
Revenue	300 000	295 000	297 000
Expenses	200 000	250 000	255 000
Trust Services			
Revenue	200 000	190 000	185 000
Expenses	140 000	150 000	160 000
Money Minders			
Revenue	250 000	230 000	225 000
Expenses	150 000	160 000	180 000

Facts about Leith Financial:

- The company was set up five years ago and is still establishing itself in the marketplace.
- Your clientele is a mix of one-off and regular consultations.
- The primary revenue for the company is obtained from consultation fees.
- The major costs to the company are staff-related (wages, travel etc.)
- The four people you employ as consultants vary widely in their customer service skills and experience within the industry.
- Employees who are new to the financial industry charge a lower hourly fee. This is increased once they achieve the necessary industry knowledge – on average it takes two years to achieve this.
- You have been targeting a niche market not currently serviced by the major financial institutions: providing investment advice to the 25-40 age group.
- Your main advertising strategy has been the distribution of flyers in the central business area of Tauranga. However, the majority of your new customers are gained via referrals from regular customers (who are often rewarded with a free consultation at the consultant's discretion).

Table 2: Employee Revenue Details (Leith Financial)

Employee	Average Fee per hour	Average Monthly Revenue	Years of Industry Experience
Hector	$80	$8000	12
Kerry	$50	$6000	1
Andrew	$80	$6000	10
Tina	$80	$7200	3

Scenario 3

You are the Business Development Manager for a home furnishings store called Dwellings. You have several stores located within the Auckland region, which all vary in size and layout, the range of goods available and the service provided. The quality and price of the goods are standard across the stores.
It has been recently noted that your Newmarket store has a significantly higher customer complaint rate than your other stores. A market research survey investigating customer satisfaction was carried out last month (February 1998). The survey questioned the customers' degree of satisfaction with various aspects of all your stores, and asked them to rate how important those aspects were to their continued patronage. Both are measured on a scale of 1 to 5 (1 = Poor / Unimportant, 5 = Excellent / Very important).

Primary Objective
Your primary objective is to analyse Dwellings' current level of customer service orientation and assess its impact on customer loyalty.

Figure 1:Customer Satisfaction Ratings

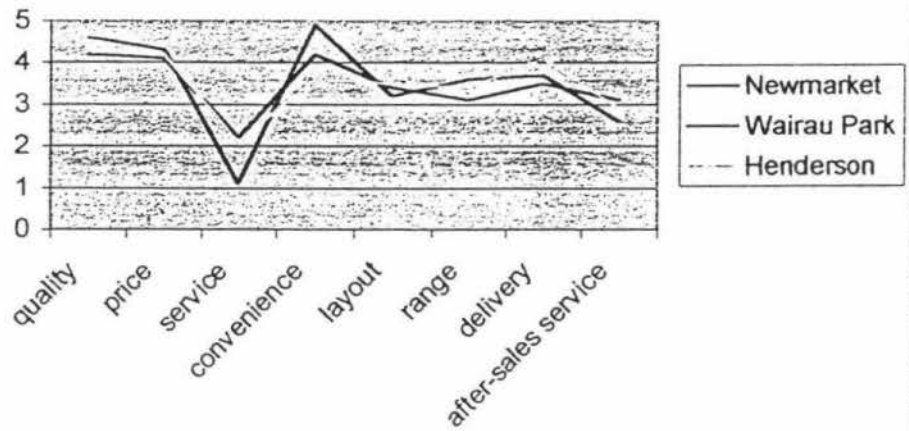


Table 1: Customer satisfaction ratings: Newmarket store (n=250)

Area	Satisfaction (1-5)	Importance (1-5)
Quality	4.6	4.2
Price	4.3	4.3
Service	1.1	4.5
Location/Convenience	4.9	2.0
Layout of store	3.2	2.5
Range	3.6	3.0
Delivery	3.7	1.5
After-sales service	2.6	3.8

- A representative sample of the qualitative comments regarding the Newmarket store are listed below:

'unhelpful'	'lack of knowledge'	'didn't listen'
'no smiles'	'unresponsive'	'bunch of kids'
'desk unattended'	'no gift wrapping'	'had to hunt for someone to help me'

Facts about Dwellings:

- A customer satisfaction survey undertaken in July 1997 indicated that customers rated Dwellings' strongest feature as customer service, followed by price, convenience and then quality.
- Your advertising strategy for the past six months (August 1997 – January 1998) has been 'The best products at the best price'.
- Extra temporary staff were hired in all stores to cope with the Christmas rush.
- Staff turnover is high, with a noticeable increase since a major competitor went into receivership. There have been rumours since then that the home furnishings retail industry is dying.
- Customer loyalty is low, with customers shopping around for the best perceived price. However, 78% of those customers surveyed also stated that receiving good customer service would make them return to a store, even if they had not previously frequented it.

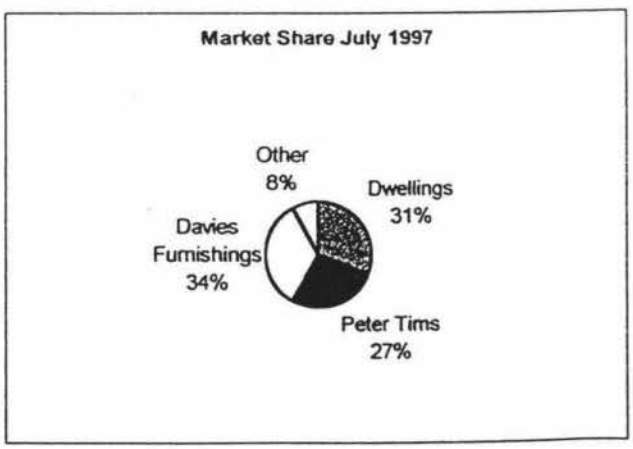
Facts about the Competition:

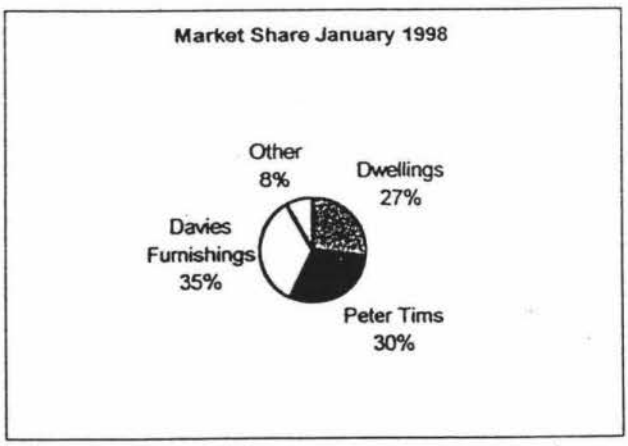
- Customer satisfaction ratings for your two major competitors, Peter Tims and Davies Furnishings, were also measured in the same survey. These are shown in Table 2.
- Peter Tims has two stores: one in Mount Wellington and one at Wairau Park. Its main advertising strategy for the last six months has been 'We go the extra mile'.
- Davies Furnishings has two super-stores: one in Greenlane and one in Henderson. Its main advertising strategy for the last six months has been 'Customer satisfaction – guaranteed'.

Table 2: Customer satisfaction: Competitors (n=250)

	Peter Tims		Davies Furnishings	
Area	Satisfaction (1-5)	Importance (1-5)	Satisfaction (1-5)	Importance (1-5)
Quality	2.6	4.2	2.4	4.2
Price	4.1	4.3	3.7	4.3
Service	4.9	4.5	4.3	4.5
Location/Convenience	4.4	2.0	3.5	2.0
Store Layout	2.9	2.5	3.0	2.5
Range	2.4	3.1	2.6	3.0
Delivery	2.9	1.5	3.8	1.5
After-sales service	4.2	3.8	4.5	3.8

Figure 2: Market Share





References

Agor, W.H. (1984). *Intuitive management: Integrating left and right brain management skills.* Englewood Cliffs, NJ: Prentice-Hall.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed). Upper Saddle River, NJ: Prentice-Hall.

Anso, J.W. (1997). *Competency profiles: An analysis of managerial competencies across industries.* (Unpublished report). Auckland, New Zealand: Massey University.

Ballantyne, I., & Povah, N. (1995). *Assessment and development centres.* Hampshire, England: Gower.

Ballé, M. (1994). *Managing systems thinking: Making dynamics work for you in business decision making.* London: McGraw-Hill.

Barrett, J., & Durivage, A. (1997). Qualite de l'experience et intelligence pratique: La validite croissante des connaissances tacites a la prediction du rendement au travail. *Psychologie Appliquee: Revue Internationale, 46,* 89-104.

Behling, O., & Eckel, N.L. (1991). Making sense out of intuition. *Academy of Management Executive, 5,* 46-54.

Bennett, R.H., Wheatley, W.J., Maddox, E.N., & Anthony, W.P. (1994). The mind's eye and the practice of management: Envisioning the ambiguous. *Management Decision, 32,* 21-29.

Bethell-Fox, C.E. (1992). Competency-based recruitment and selection. In A. Mitrani, M. Dalziel, & D. Fitt (Eds.), *Competency based human resource management: Value-driven strategies for recruitment, development and reward* (*pp.*68–84). London: Kogan Page.

Boyatzis, R. E. (1982). *The competent manager: A model for effective performance.* New York: Wiley.

Brockman, E.N., & Simmonds, P.G. (1997). Strategic decision making: The influence of CEO experience and use of tacit knowledge. *Journal of Managerial Issues, 9,* 454-467.

Brookfield, S. (1987). *Developing critical thinkers.* San Francisco: Jossey-Bass.

Burgoyne, J., & Stuart, R. (1978). *Management development: Context and strategies.* Westmead, England: Gower.

Cascio,W.F. (1998). *Applied psychology in human resource management* (5th ed.) Upper Saddle River, NJ: Prentice-Hall.

Cederblom, J., & Paulsen, D.W. (1986). *Critical reasoning* (2nd ed.). Belmont, CA: Wadsworth.

Conoley, J.C., & Kramer, J.J. (Eds.) (1989). *The 10th mental measurements yearbook.* Lincoln, Nebraska: Buros Institute of Mental Measurement of University of Nebraska-Lincoln.

De Bono, E. (1985). *Six thinking hats.* London: Penguin.

Elkin, G. (1995). The development of managerial competence and leadership. In P. Boxall (Ed.), *The challenge of human resource management: Directions and debates in New Zealand* (*pp.* 250-267). Auckland, New Zealand: Longman Paul.

Filipczak, B. (1994). The training manager in the 90's. *Training, 31,* 31-35.

Garrison, D.R. (1992). Critical thinking and self-directed learning in adult education: An analysis of responsibility and control issues. *Adult Education Quarterly, 42,* 136-148.

Haladyna, T.M. (1997). *Writing test items to evaluate higher order thinking.* Needham Heights, MA: Allyn and Bacon.

Hirsh, H.R., Schmidt, F.L., & Hunter, J.E. (1986). Estimation of employment validities by less experienced judges. *Personnel Psychology, 39,* 337– 344.

Holland, D., & Williams, J.B. (1994). Sharpening your executive competencies. *Healthcare Executive, 9,* 13-17.

Howell, D.C. (1992). *Statistical methods for psychology* (3rd ed.). Belmont, CA: Duxbury Press.

Hunter, J.E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29,* 340-362.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Impara, J.C., & Conoley, J.C. (Eds.). (1996). *The 11th mental measurements yearbook.* Lincoln, Nebraska: Buros Institute of Mental Measurement of University of Nebraska-Lincoln.

Johnsson, J. (1991). How to tap board members' business acumen. *Hospitals, 65,* 64.

Kaplan, R.M., & Saccuzzo, D.P. (1997). *Psychological testing: principles, applications, and issues.* (4th ed.). Pacific Grove, CA: Brooks/Cole.

Kerr, M. (1994). Skill New Zealand: Strategy for a bright new future. *People and Performance, 2,* 35-37.

Keyser, D.J. & Sweetland, R.C. (Eds.) (1992). *Test critiques: Vol 9.* Austin, Texas: Pro-ed.

Klemp, G.O. Jr., & McClelland, D.C. (1986). What characterizes intelligent functioning among senior managers? In R.J. Sternberg & R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (*pp.* 31-50). New York: Cambridge University Press.

Lawler, E.E. (1994). From job-based to competency-based organizations. *Journal of Organizational Behavior, 15,* 3-15.

Lillibridge, J.R., & Williams, K.J. (1992). Another look at personality and managerial potential: Application of the five-factor model. In K. Kelly (Ed.), *Issues, theory, and research in industrial / organisational psychology* (*pp.*91-115). Amsterdam: Elsevier Science.

Limerick, D.C. (1990). Managers of meaning: From Bob Geldof's Band Aid to Australian CEOs. *Organizational Dynamics, 18,* 22-33.

Little, J.F. (1980). *Critical thinking and decision making.* Ontario, Canada: Butterworth and Co.

Longenecker, C.O., Simonetti, J.L., & Mulias, M. (1996). Survival skills for the IS professional. *Information Systems Management, 13,* 26-31.

Maccoby, M. (1990). Forces for increased organizational competition: Implications for leadership and organizational competence. In D.L. Fishman & C. Cherniss (Eds.), *The human side of corporate competitiveness* (*pp.* 35-50). Newbury Park, CA: Sage Publications.

McDaniel, M.A., Schmidt, F.L., & Hunter, J.E. (1988). Job experience correlates of job performance. *Journal of Applied Psychology, 73,* 327-330.

McKenzie, L. (1992). Critical thinking in health care supervision. *Health Care Supervisor, 10,* 4, 1-11.

McLagan, P. (1997). Competencies: The next generation. *Training and Development, 51,* 40-47.

Miller, L.R. (1997). Better ways to think and communicate. *Association Management, 49,* 71-73.

Mitrani, A., Dalziel, M., & Fitt, D. (1992). *Competency based human resource management.* London: Kogan Page.

Mo, Y.H. (1996). Orientating values with eastern ways. *Personnel Management, 2,* 28-30.

Myers, C., & Davids, K. (1992). Knowing and doing: Tacit skill at work. *Personnel Management, 24,* 2, 45-47.

Myers, C., & Davids, K. (1993). Tacit skill and performance at work. *Applied Psychology: An International Review, 42,* 117-137.

Neisser, U. (1976). General, academic, and artificial intelligence. In L. Resnick (Ed.), *The nature of intelligence (pp.* 135-144). Hillsdale, NJ: Earlbaum.

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

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5,* 1, 14-37.

Norris, S.P., & Ennis, R.H. (1990). *Evaluating critical thinking.* Cheltenham, Australia: Hawker Brownlow Education.

O'Connor, G., & Smallman, C. (1995). The hybrid manager: A review. *Management Decision, 33,* 19-28.

Pascarella, P. (1997). The secret of turning thinking into action. *Management Review, 86,* 38-39.

Polanyi, M. (1962). *Personal knowledge: Towards a post-critical philosophy.* London: Routledge & Kegan Paul.

Pomeroy, L. (1995). Contributors to performance: competence and other things. *People and Performance, 3,* 13-16.

Ree, M.J., Earles, J.A., & Teachout, M.S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79,* 518-524.

Russell, C. J. (1990). Selecting top corporate leaders: An example of biographical information. *Journal of Management, 16,* 73-86.

Saint-Onge, H. (1996). Tacit knowledge: The key to the strategic alignment of intellectual capital. *Planning Review, 24,* 2, 10-14.

Saville & Holdsworth Limited (1991). *SHL management and graduate item bank: Manual and user's guide.* Surrey, UK: SHL.

Schmidt, F.L., Hunter, J.E., Croll, P.R., & McKenzie, R.C. (1983). Estimation of employment test validities by expert judgement. *Journal of Applied Psychology, 68,* 590-601.

Schmidt, F.L., Hunter, J.E., & Outerbridge, A.N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71,* 432-439.

Schon, D.A. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic books.

Scribner, S. (1986). Thinking in action: Some characteristics of practical thought. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (*pp.* 13-30). New York: Cambridge University Press.

Sen, J., & Das, J.P. (1991). Planning competence and managerial excellence: A research framework. In R.F. Mulcahy, R.H. Short, & J. Andrews (Eds.), *Enhancing learning and thinking* (*pp.* 209-226). New York: Praeger.

Shore, T.H., Shore, L.M., & Thornton III, G.C. (1992). Construct validity of self and peer evaluations of performance dimensions in an assessment center. *Journal of Applied Psychology, 77,* 1, 42-54.

Smith, A. (1994). Is business ignoring its customers? *Management-Auckland, 41,* 80-81.

Spencer, L.M. Jr., & Spencer, S.M. (1993). *Competence at work: Models for superior performance.* New York: John Wiley & Sons.

Sternberg, R.J. (1994). PRSVL: an integrative framework for understanding mind in context. In R.J. Sternberg & R.K. Wagner (Eds.), *Mind in context: interactionist perspectives on human intelligence* (*pp.* 218-232). New York: Cambridge University Press.

Sternberg, R.J., & Wagner, R.K. (1989). The fate of the trait: A reply to Cantor and Kihlstrom. In R.S. Wyer Jr. & T.K. Srull (Eds.), *Advances in social cognition, Volume II: Social intelligence and cognitive assessments of personality,* (*pp.* 175-185). Hillsdale, NJ: Erlbaum.

Sternberg, R.J., & Wagner, R.K. (1989a). Individual differences in practical knowledge and its acquisition. In P.L. Ackerman, R.J. Sternberg & R. Glaser (Eds.), *Learning and individual differences: Advances in theory and research* (*pp.* 255-278). W.H. Freeman: U.S.A.

Sweetland, R.C., & Keyser, D.J. (Eds.) (1991). *Tests: A comprehensive reference for assessments in psychology, education, and business.* (3rd ed.). Austin, Texas: Pro-ed.

Thomson, A. (1996). *Critical reasoning: A practical introduction.* London: Routledge.

Thornton, G.C. (1980). Psychometric properties of self-appraisals of job performance. *Personnel Psychology, 33,* 263-271.

Thornton, G.C., & Byham, W.C. (1982). *Assessment centres and managerial performance.* New York: Academic press.

Thornton, G.C., & Cleveland, J.N. (1990). Developing managerial talent through simulation. *American Psychologist, 45,* 190-199.

Vasilash, G.S. (1997). Clueless: How American companies aren't optimizing their brainpower. *Production, 109,* 68-70.

Wagner, R.K. (1987). Tacit knowledge in everyday intelligent behavior. *Journal of Personality and Social Psychology, 52,* 1236-1247.

Wagner, R.K. (1991). Managerial problem solving. In R.J. Sternberg, & P.A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (*pp.*159-183). Hillside, NJ: Lawrence Erlbaum.

Wagner, R.K. (1994). Context counts: the case of cognitive ability testing for job selection. In R.J. Sternberg & R.K. Wagner (Eds.), *Mind in context: interactionist perspectives on human intelligence* (*pp.* 133-150). New York: Cambridge University Press.

Wagner, R.K., & Sternberg, R.J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49,* 436-458.

Wagner, R.K., & Sternberg, R.J. (1986). Tacit knowledge and intelligence in the everyday world. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world* (*pp.* 51-83). New York: Cambridge University Press.

Wagner, R.K., & Sternberg, R.J. (1987). Tacit knowledge in managerial success. *Journal of Business and Psychology, 1,* 301-312.

Wallace, J., & Hunt, J. (1996). An analysis of managerial competencies across hierarchical levels and industry sectors: A contemporary Australian perspective. *Journal of the Australian and New Zealand Academy of Management, 2,* 36-47.

Watts, G.W. (1993). Exercises for bringing the head and heart together. *Training & Development, 47,* 17-19.

Wilkinson, L.J., & Zwanenberg, N. (1994). Development of a person specification system for managerial jobs. *Personnel Review, 23,* 25-36.

Wills, S. (1993). MCI and the competency movement: The case so far. *Journal of European Industrial Training, 17,* 9-11.

Woodruffe, C. (1993a). *Assessment centres: Identifying and developing competence* (2nd ed). London: Institute of Personnel Management.

Woodruffe, C. (1993b). What is meant by a competency? *Leadership & Organization Development journal, 14,* 29-36.

Yukl, G. (1994). *Leadership in organizations* (3rd ed.) Englewood Cliffs, NJ: Prentice Hall.

Zingheim, P.K., Ledford, G.E. Jr., & Schuster, J.R. (1996). Competencies and competency models: Does one size fit all? *ACA Journal,* 56-64.

Zuckerman, S. (1995). Brilliant mind, but is he a leader? *United States Banker, 105,* 2, 24-32.