

國立台灣大學商學研究所博士班入學考試試卷（102 學年度）
科目 策略管理（含基本概念和文獻評析） 第 1 頁/共 1 頁

（考試時間 3 小時）

請由下列兩篇文獻中挑選一篇，回答下列問題（切勿兩篇皆作答）：

1. 請說明該篇論文所牽涉到的策略管理基本概念(20%)
2. 寫出 300 字以內的中文摘要，以及 150 字以內的英文摘要(20%)
3. 說明該篇論文的：(a)研究問題、(b)理論架構、(c)研究方法、
(d)研究發現(40%)
4. 請條列說明論文中的貢獻及缺點，並提出可能的改進之道。(20%)

文獻另附

請考生註明文獻一或文獻二

試題請啟回

KNOWLEDGE STRUCTURES OF PROSPECTORS, ANALYZERS, AND DEFENDERS: CONTENT, STRUCTURE, STABILITY, AND PERFORMANCE

BORIS KABANOFF* and SHANE BROWN

Queensland University of Technology, School of Management, Brisbane, Queensland, Australia

INTRODUCTION

Cognitive psychology rests on two key assumptions: (1) people's responses to stimuli are mediated by information processing; and (2) the capacity of the information-processing system is limited, generating the need for a mechanism of selection, attention being such a mechanism (Moors and De Houwer, 2006). Organizational theorists have drawn on these ideas to propose a tripartite model of managerial cognition involving attention, interpretation, and action to explain how top managers function in environments where they face a barrage of ill-defined events and trends that cannot all be attended to (Abrahamson and Hambrick,

1997; Daft and Weick, 1984; Walsh, 1995). At the heart of this information-processing system lies the knowledge structure concept that Walsh (1995: 281) defined as 'a mental template that individuals impose on an informational environment to give it form and meaning.' Knowledge structures order an information environment in a way that enables subsequent interpretation and action, are built on past experience, and represent organized knowledge about a given concept or type of stimulus. Porac and Thomas (2002: 178) described them as the 'essential lens' and 'fundamental assumption' for understanding how strategists make sense of their environment. Walsh (1995: 280) concluded that 'This [knowledge structure] issue has captured the imagination of managerial and organizational researchers' but observed also that 'their inquiry has been eclectic in focus and method.' Walsh (1995) viewed such eclecticism as understandable in a new, growing field, but there is

growing consensus (e.g., Hodgkinson, 2001; Porac and Thomas, 2002; Walsh, 1995) that we now need to move beyond eclecticism. These authors argue that we require two things: (1) a stronger theoretical base for understanding and studying knowledge structures; and (2) methodologies that have a less qualitative, researcher-driven orientation, and which permit large sample, longitudinal studies of knowledge structures that include organizational outcomes such as firm performance. The present study takes up these two challenges by describing a theory-based model of strategic knowledge structures and a method for describing their structure and content over time.

A MODEL OF STRATEGIC KNOWLEDGE STRUCTURES

Core assumptions

Consistent with existing models, we assume that top managers operate in an information environment (internal and external to the organization) too rich to be fully attended to; therefore they employ an attention process. Our focus is the knowledge structures of organizations' top-team or dominant coalition who share a common understanding of their firm's overall strategy. While individual members need not have identical knowledge, nor equal influence on its development, we assume that it is broadly shared and reflects a form of consensus among this group. Figure 1 shows that attention both influences how strategic knowledge structures develop and how they shape subsequent strategic interpretations and actions. The issues that managers attend to influence their learning over time and shape their knowledge structures, which, in turn, influence strategic choices by affecting what managers subsequently focus on (Ocasio, 1997). In the initial stages, as executives explore and learn about their environment, their attention can be described as 'mindful' in that it involves active attention, but as the knowledge structure develops, the process becomes relatively more automatic¹ (Moors and De Houwer, 2006) or, using Walsh's term (1995: 305) 'mindless.'

¹ The term 'relatively' is significant since, as Moors and De Houwer (2006) note, cognitive psychologists no longer subscribe to the view that attention or its opposite automatically is an all-or-none phenomenon; therefore knowledge structures should be seen as involving a relatively smaller role for attention rather than no role.

During the development process, managerial attention is influenced by a number of factors, including the cognitive and other psychological characteristics of top-team members (Hambrick and Mason, 1984), top-team interaction (Miller, Burke, and Glick, 1998), and organizational structures and policies (Ocasio, 1997), as well as by exogenous variables, specifically the prevalence of different types of information in the environment (internal and external). For example, some environments have a higher density of signals suggesting that a mature, low-growth, and technologically relatively stable market exists, and that the firm has capabilities enabling it to compete by being highly efficient at what it already does (these are signified as ●-type signals in Figure 1 and are shown as relatively more dense on the right side in Figure 1). Others firms encounter signals suggesting a dynamic, growing marketplace where competitors are likely to introduce new products, and where innovation capacity is the key to competing successfully (▲-type signals in Figure 1 and denser on the left). Therefore top teams from different organizations, as a result of individual and environmental differences, attend to different features of their environment, which in turn guide their strategic choices. In Figure 1, for example, the top team in firm 1 initially attends mainly, but not solely, to ▲-type signals, while the reverse happens in firm 3. This leads to our first hypothesis, which, while rather obvious, is important to shaping the overall model:

Hypothesis 1: Top teams from different organizations attend to a variety of strategically relevant elements in the internal and external organizational environment.

A generic strategy perspective

While the variety of elements that top teams can attend to is virtually infinite, the elements that top teams actually come to focus upon are restricted by both the inherent limitations of human information processing and the kinds of strategic choices that are perceived to be viable in the setting. From a generic strategy perspective, as a result of organizational and environmental characteristics that are critical to competition regardless of industry, there are a limited number of strategic configurations that are viable in any industry (Doty, Glick, and Huber, 1993; Ketchen, Thomas, and Snow, 1993).

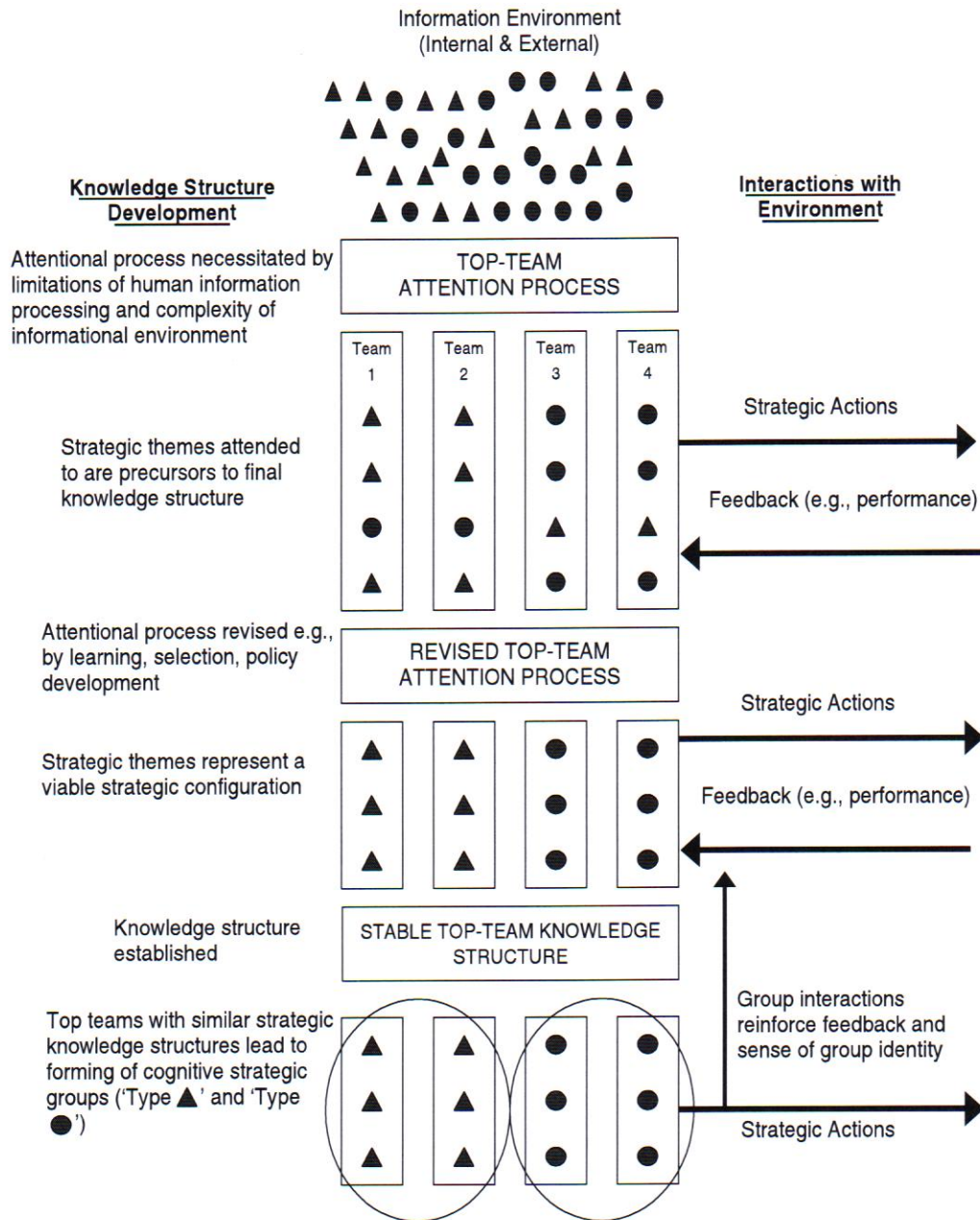


Figure 1. A model of strategic knowledge structure development

Based on Ketchen *et al.* (1993: 1278) we define strategic configurations as 'commonly occurring clusters of strategic elements.' This is represented in Figure 1, which shows that as the top teams' initial strategic choices interact with and receive feedback from the environment (for example, in terms of firm performance) their attention process is modified to focus on elements that are consistent with a particular strategic configuration drawn from a limited set of feasible, strategic configurations. (In this simplified model, there are only two types of configurations: either all ●-type or ▲-type elements).

For several reasons we draw upon Miles and Snow's (1978) typology to describe the feasible set of strategic configurations. Hambrick (2003: 116) commented on the staying power of Miles and Snow's typology, observing that 'of the several strategy classification systems introduced over the past 25 years, it has been the most enduring, the most scrutinized, and the most used' (see also DeSarbo *et al.*, 2005). Miles and Snow's framework has the advantages of being both parsimonious and rich. Its parsimony stems from its proposal that, whichever strategy a firm follows, the managers of that business have to deal with

issues and decisions across three main domains: entrepreneurial (product-market decisions), engineering (production and delivery), and administrative (structures, roles, and policies). Its richness lies in the form of quite detailed descriptions of how the attributes of the main strategic types differ across these three domains. Therefore there is a relatively well-specified, limited set of strategic elements that managers following different generic strategies should focus their attention on, and these elements are patterned or configured in a limited number of ways. Thus, prospectors' knowledge structures differ from defenders' not simply because they attend to a broader, more dynamic domain externally, but also because they combine this with a focus on process flexibility internally, while defenders combine a low level of external scanning with an internal focus on efficiency (see Table 2). Therefore, what managers attend to varies not only in terms of specific, strategic elements (or *content*; Walsh, 1995) but also in how these are structured.

Another aspect of Miles and Snow's model helps explain why, eventually, all of the elements that top teams focus on are of a particular type (in our model either ▲ or ●). The adaptive cycle hypothesis argues that, whatever type of strategy an organization follows, it is required to find different, mutually aligned solutions across the three domains previously mentioned. A business that makes an entrepreneurial decision of a prospector type then encounters issues and problems in the engineering and administrative structures domains, which it needs to address and resolve in ways that are aligned with being a prospector in the product-market domain. Over a series of adaptive cycles, managers and their firms ideally find different, mutually aligned solutions to these three types of problems; and both firms and managers become more focused on dealing with certain types of problems, and less focused on others. In Figure 1 we see this in the nature of each top team's revised attention process that includes only ▲ or ● elements. Our second hypothesis reflects this:

Hypothesis 2: The strategic elements top-teams attend to form higher order structures or configurations that broadly align with the major strategic types identified by Miles and Snow (1978).

To this point we have concentrated on how individual top teams develop their strategic knowledge structures in interaction with their environment. Next we consider how interaction with other top teams with a similar strategic outlook helps the development of cognitive strategic groups in an industry.

Strategic groups

Porter (1980: 129) defined a strategic group as 'the group of firms in an industry following the same or similar strategy along the strategic dimensions.' Strategic groups are an intermediate concept, located between whole industries and single firms, and help explain patterns of competition within industries. Two key features of strategic groups are that: (1) organizational members of different groups have attributes that are not readily imitable by members of other groups, creating mobility barriers; and that (2) restrictions on mobility enable some groups in a favorable 'niche' in their industry to consistently outperform others. Despite some stern criticisms (e.g., Barney and Hoskisson, 1990) the strategic groups concept continues to attract considerable interest (e.g., Ketchen *et al.*, 1993).

Managerial cognition researchers adopted the strategic group concept and a substantial literature developed concerned with identifying 'cognitive communities' or cognitive strategic groups (Porac, Mishina, and Pollock, 2002; Reger and Huff, 1993) that have been described as the 'cognitive analogue to conventional strategic groups' (Hodgkinson, 2001: 71; Nath and Gruca, 1997; Osborne, Stubbart, and Ramaprasad, 2001). Porac, Thomas, and Badden-Fuller (1989) described cognitive strategic groups as groups of companies whose top managers hold similar or shared mental models of strategy within their industry. The similarity to Porter's definition is obvious.

Cognitive strategic groups help us understand how top teams' strategic knowledge structures are shaped and maintained over time. As described in Porac *et al.*'s seminal study:

... the mental models and strategic choices of key decision-makers intertwine to create a stable set of transactions in the marketplace ... [so that] the mental models form a critical link between group-level and firm-level dynamics ... [and] such beliefs are reinforced by a mutual enactment process in which the technical choices constrain the flow of

information back to decision-makers ... to what has already been determined by existing beliefs. (Porac *et al.*, 1989: 412)

This self-fulfilling quality of knowledge structures has been identified as both their major benefit and potential major cost (Walsh, 1995).

Strategic groups and cognitive strategic groups theories make a similar prediction that top teams' knowledge structures tend to be stable over time, though the reasons differ in an interesting way. Mobility barriers emphasize that managers tend to focus on the same strategic issues and choices over time because organizational and environmental attributes constrain managerial options. Cognitive theorists argue that knowledge structures are the ultimate source of mobility barriers (Porac *et al.*, 1995): it is managers' beliefs about their firm, group, and industry that form the barrier. Consistent with both perspectives, we hypothesize:

Hypothesis 3: Top-teams' strategic knowledge structures and firms' membership of cognitive strategic groups tend to be relatively stable over time.

Performance differences

Strategic groups and cognitive strategic groups theorists both agree in another respect: group structures are linked to performance differences. However, evidence for this core prediction is lacking. Lewis and Thomas (1990: 385) observed that 'the empirical evidence linking performance differences with strategic groups is not extensive and is conflicting.' Evidence of performance differences between cognitive strategic groups is notable mainly for its paucity; as Hodgkinson (2001: 75–77) noted, the 'modal' design for empirical work in this area has been a cross-sectional study, involving single representatives from a relatively small number of organizations within a single industry, and having no measures of organizational performance.

While the context for most previous research into strategic groups, both conventional and cognitive, has been a single industry, we examine a multi-industry sample; thus our hypotheses recognize that different cognitive groups are likely to outperform in different industry contexts, that is, strategy and industry interactively determine firm performance (Hambrick, 1983; Misangyi *et al.*, 2006):

Hypothesis 4: Cognitive strategic groups differ in their financial performance, and these differences vary by industry (i.e., group and industry interactively effect performance).

Given the exploratory nature of our research, this hypothesis is a general one. However, since we aim to describe strategic knowledge structures constructed around a well-studied typology (Hambrick, 2003), we also propose several more specific hypotheses about performance differences. We base these primarily upon Hambrick (1983).

Hambrick (1983) used the PIMS dataset (Anderson and Paine, 1978) to classify firms into two main strategies—prospectors and defenders—according to the percentage of sales they derived from new products. Analyzers were used as the reference strategy and were defined as those with moderate scores on this variable. The study had three measures of performance: return on investment (ROI), cash flow on investment (CFOI), and market share change (MSC). Hambrick (1983: 18) found that 'in general the "superior" strategy was neither of the two extreme strategies. It was the base case strategy—that is, the analyzer.' Therefore, given that we are able to identify groups that are a reasonable match with defenders, analyzers, and prospectors, we propose the following hypothesis:

Hypothesis 4a: In general, analyzer-type cognitive strategic groups outperform prospector and defender-types on Return on Assets (ROA) and Return on Equity (ROE).

Hambrick (1983: 19) found that while 'being a prospector hurt ROI and CFOI in general ... being a prospector helped MSC in innovative industries.' That is, in innovative industries prospectors grew their market share more than other types. Since we do not have measures of industry innovation or market share change available for this study, we are unable to compare our results directly with Hambrick's. However, if we assume that the price/earnings ratio (PER), which reflects investor expectations about future firm performance, in part at least reflects investors' expectations about future growth in market share, we can advance the following hypothesis:

Hypothesis 4b: In some industries, cognitive groups of the prospector type will outperform

defender and analyzer types in terms of their PER.

MEASURING TOP-TEAM ATTENTION

While there have been many different approaches to describing and measuring knowledge structures (Huff and Jenkins, 2002), the approach we employ involves describing managers' attentional focus. Following D'Aveni and MacMillan (1990: 640) we define attention as 'allocating information-processing capacity (receiving, cognitive processing, disseminating) to environmental stimuli over time,' which is close in meaning to the everyday phrase 'paying attention to.' In the next section we explain how we describe what top teams pay attention to by analyzing the content of firms' annual reports.

Analyzing attentional focus in annual reports

By analyzing the words, word categories, and broader themes people use in speech, text, or other forms of communication, the Whorf-Sapir hypothesis argues, we can describe the relative importance of various cognitive categories in the content and structure of people's 'world view.' Content analysis is a long-established means for doing this (Weber, 1990).

Content analysis allows us to measure a number of aspects of communicators' cognition including: *what* they attend to (the words in text are grouped into categories of related words called themes), *how much* they attend to different themes (higher frequency reflects greater cognitive centrality), *changes in centrality* (changes in frequency of references), and *mental connection between themes* (the juxtaposition or co-occurrence of themes). The method has many advantages for the study of strategic cognition including its unobtrusive nature, rigor, and transparency, its utilization of archival data that is available over time, and the access it offers to the cognitions of people who are difficult to access (Duriau, Reger, and Pfarer, 2007). In the context of the present study the method allows us to deal with many of the limitations of the 'modal' design described by Hodgkinson (2001).

A possible criticism of the approach centers not on content analysis *per se* but the annual report text analyzed, and suggests that since annual reports are

prepared by communication departments, consultants, or PR practitioners, they may not reflect the cognitions of senior managers; and, even if senior managers influence the content, it is more likely to reflect their attempts at impression management than their actual cognition. The evidence for and against these criticisms has been considered in detail by Abrahamson and Hambrick (1997: 519), who concluded that, on balance, an information-processing interpretation of annual report content was more appropriate than an impression-management interpretation when we are dealing with non-evaluative comments; however, evaluative comments are more strongly influenced by impression management. A considerable number of studies using both text-based and non-text-based measures of managerial and organizational behaviors have found theoretically meaningful and statistically significant associations between them (see reviews by Duriau *et al.*, 2007; Morris, 1994). This leads us to conclude that annual report text can be used for studying managerial cognition, particularly when the focus is on non-evaluative, descriptive themes, as in our case.

METHOD

Developing a machine-learning approach to text analysis

Computer-aided text analysis (CATA) describes a family of text analytic approaches that vary in the complexity of concepts they seek to investigate, as well as the relative amount of automation involved. The complexity of the thematic categories involved in CATA varies widely between applications, from single words (Osborne *et al.*, 2001) to higher-order compound concepts such as attention to the external environment (Levy, 2005). At the minimal automation end of the scale, software can be used to select broadly relevant portions of text that is then coded more precisely by human raters (e.g., Palmer, Kabanoff, and Dunford, 1997), while concept discovery (text-mining) systems such as Leximancer (Smith and Humphreys, 2006) are at the almost fully automated end of the scale.

Our method is a 'machine-learning' (Sebastiani, 2002) approach to text categorization. Machine learning (ML) refers to the process by which a text classifier is created by 'learning' a set of pre-classified texts. The machine-learning approach has advantages over traditional dictionary

approaches; in particular it does not require the extremely time-intensive development of coding schema, rules, or word lists that include all or most of the synonyms for any broad theme, which can approach that of manual coding. The ML approach allows the identification of themes of interest from words that tend to co-occur with that topic or theme. It uses a mathematical basis for determining words associated with a theme of interest, rather than requiring a human coder to identify and design a rule for identifying each association. ML has inherent disambiguation² properties such as found in the General Inquirer (Stone *et al.*, 1966), since classification depends on a combination of associative probabilities between words that belong to a theme of interest (the 'category') and other words that tend to occur in the same context, but without explicitly specifying a set of disambiguation rules. We use the terms 'theme' and 'category' interchangeably because both occur commonly in text analysis in this context, though in the present case text is not, strictly speaking, allocated to one and only one category of meaning, as we now explain.

Our classifier uses the naïve Bayesian method to train a classifier of strategic themes in annual reports, and subsequently score reports for their content. The framework central to the approach draws upon a Perl module family (Williams, 2003), which provides methods for all aspects of the ML task, including document feature selection, naïve Bayesian training, and classification. The ML approach relies on example category sets (collections of text representing true members of categories or themes), which the algorithm uses to 'train' a classifier. The naïve Bayesian method uses the probabilities of association between words and themes to accomplish this training. The comparative frequencies of the words associated with

different themes, and the texts as a whole, determine how these weightings are calculated: words often seen in conjunction with a particular theme will generate a high likelihood of presence for the theme. The most likely category 'score' for a section of text is the category (cat_k) for which the following formula is maximized:

$$P(w_1|cat_k) \times P(w_2|cat_k) \dots P(w_n|cat_k) \times P(cat_k)$$

where $P(w_i|cat_k)$ is the probability that word 'i' is present for that category (calculated from the training set), and $P(cat_k)$ is the base probability of category membership. Since the naïve Bayesian method is probabilistic, these outcomes can be used to order each text segment (sentences in this case) in terms of the probability that the category or categories of interest are present. This allows us to use aggregated category likelihoods as the unadjusted measure of theme presence in a report, rather than applying an arbitrary cutoff to each sentence's results. More details about this method follow the description of our text data and firm performance measures.

Sample

Text from 5,000 electronically available annual reports from 2006 Australian companies listed on the Australian Stock Exchange (ASX) across 12 years (Connect 4 Annual Report Collection, 1992–2003) forms the dataset for this research. The majority of companies had multiple reports, with 74 percent having two or more, while the mean number of reports per company was 3.65. Connect 4 includes between 120 (in the early years) and 600 per year of the largest publicly listed companies in Australia, which make up between 55 and 75 percent of total market capitalization of the ASX (see Table 1 for sample size by year and industry sector). The top-level Global Industry Classification Standard (GICS) includes 10 sectors: Consumer Discretionary, Consumer Staples, Energy, Financial, Health, Industrial, Information Technology, Materials, Telecommunications, and Utilities; we excluded the last two owing to the small numbers of firms. To ensure a focus on the cognitions of top-level management, only sections of annual reports with titles identifying the source of the content as a top-level executive were analyzed. These sections typically contain a message to shareholders from the CEO

² Disambiguation, a term that comes from the field of computational linguistics, involves specifying a set of rules that allow software to distinguish between homographs—that is, words that have more than one meaning. For example, does the word 'mine' mean first person possessive or the process of extracting ore from the ground? Disambiguation rules typically specify words or phrases that tend to co-occur with the target word in a particular context in order to let the software decide in which context the word is being used. For example, the words 'silver,' 'gold,' and 'coal' may be markers for the second meaning of the word 'mine' in the example just given. In the present case the 'training' of the text classifier, in essence, involves the algorithm developing a set of probabilistic disambiguation rules based on the co-occurrences of words in the training set of sentences.

Table 1. Number of sampled companies by year and industry sector (GICS code)

Industry sector	Year												
	'92	'93	'94	'95	'96	'97	'98	'99	'00	'01	'02	'03	Total
Consumer discretionary	9	24	25	44	48	54	55	59	61	62	64	52	557
Consumer staples	10	17	16	25	22	22	27	33	30	32	28	19	281
Energy	3	11	11	17	19	23	26	15	17	18	18	9	187
Financials	12	42	49	87	94	100	110	99	80	81	88	68	910
Health	2	6	7	10	12	14	16	17	19	30	35	26	194
Industrials	8	29	23	48	55	53	56	54	47	44	54	36	507
Information technology	0	7	7	14	13	11	10	18	30	32	27	23	192
Materials	20	50	62	95	102	83	82	76	69	71	70	55	835
Total	64	186	200	340	365	360	382	371	353	370	384	288	3,663

or managing director and, by and large, equate to the president's letter common to U.S. firms' annual reports. Reports were first filtered through the Wordnet lexical reference system (Fellbaum, 1998) to remove most company, location, and personnel names to reduce the effects of idiosyncratic terminology on the content analysis process.

Measures of financial performance

Given that we were investigating a multi-industry sample, we employed three widely applicable measures of financial performance: ROE, ROA, and PER. ROE is calculated by dividing net profit before abnormals by shareholders' equity. In cases where shareholders' equity is less than or equal to zero, the value of ROE is set to null. ROE is an important indicator of company performance as it provides information on how well managers are employing funds invested by the shareholders to generate returns. ROA is a fiscal year's earnings divided by total assets, that is, shareholders' equity as well as other borrowings. PER is a valuation measure that divides the company's share price by its pre-abnormals earnings per share. PER is calculated as the closing share price on the last day of the company's financial year divided by the pre-abnormals earnings per share. The first two measures are viewed as accounting-based indicators of the efficiency with which a firm applies its resources, while the third is a market-based measure that is generally viewed as reflecting investors' expectations and judgments about the quality of a firm's future performance. The financial data were obtained from *Aspect Fin Analysis*, which provides detailed financial information for companies listed on the ASX.

DEVELOPING THE AUTOMATIC TEXT CLASSIFIER TO IDENTIFY STRATEGIC THEMES

We begin the process of identifying the main content themes in strategic discourse in annual reports by focusing on the two strategic types that are the most different and distinctive in Miles and Snow's (1978) typology: prospectors and defenders. We identify a set of content themes that represent as closely as possible the key descriptors used by Miles and Snow for each of the three main sets of strategic issues: entrepreneurial, engineering, and administrative. We are not able to replicate exactly all of these themes since some features are not discussed in annual reports, and we also identify other themes that reflect important strategic discourse as we explore the text, so that we eventually develop a coding scheme comprising 21 themes (Table 2). Each of the themes is broadly defined; examples of each theme from the text are identified; and we then set out to identify a set of sentences that can be used to 'train' the classifier to recognize these themes in other sentences, as we now explain.

To compile the training set, two coders (one of them was one of the authors, the other was a postgraduate student) independently developed sets of 2,500 sentences from annual reports that they evaluated for the presence or absence of the 21 themes. Software was developed to facilitate this process. The first step in selecting sentences to code is to do simple word searches for key theme terms, such as 'cost' and 'cut,' or 'reduce' and 'overhead' for the cost-cutting theme. As more sentences are identified, the list of words for selecting other sentences related to a theme is expanded. This procedure is carried out by each coder for

Table 2. Defender and prospector characteristics and associated strategic themes

	Strategic themes intended to capture Miles and Snow descriptions (with examples) ^a
Defenders	
<i>Entrepreneurial</i>	<i>Core business/prudence</i> : believe the past year has been one consolidation cautious growth be followed up similar result
Narrow, stable domain	<i>Cost reduction</i> : extensive restructuring entire workforce was largest single factor in cutting operating costs
Market competition (Service, pricing)	<i>Productivity</i> : in expectation bank will maintain emphasis on efficiency productivity improvements ...
Narrow scanning	<i>Collaboration/alliances</i> : strategic alliance in February year demonstrated company's ability in forging ...
Cautious growth	<i>Sale and closure</i> : financial year seven entities within group have been either wound up or divested
<i>Engineering</i>	<i>Restructuring</i> : banks branch networks were restructured year branch management regional structure
Cost minimization	<i>Self-comparison (financial)</i> : record operating profit numerical million an increase numerical compared previous year
Productivity, efficiency	
<i>Administrative</i>	
Finance pre-eminent	
Stable elite	
Intensive, bottom-line, structured planning	
Centralized	
Self-comparisons	
Prospectors	
<i>Entrepreneurial</i>	<i>Market/economic conditions</i> : management shortcomings board senior management might have seen responded trend earlier
Broad, changeable domain	<i>New markets</i> : additionally entry new overseas markets including Chile has assisted growth strategy
Product and market development	<i>Competition</i> : increased market share in most market segments company competes in
Industry leaders, drivers	<i>Product focus</i> : due improved product sourcing introduction new products existing range ...
<i>Engineering</i>	<i>Customer service</i> : group continued commitment quality customer service ...
People are indispensable	<i>Research and development</i> : have created product development division contains all research projects
Process flexibility	
<i>Administrative</i>	<i>Marketing</i> : marketing department will continue develop products on global scale

each of the themes. In effect, this part of the process parallels that normally used to develop content categories incorporating lists of words referring to themes (i.e., content dictionaries); however, the goal here is to identify sentences using a broad range of terms that refer to particular themes so that these sentences can then be used to train the software to recognize these themes in other sentences. The sentences selected are classified according to the broad coding scheme for the 21 themes.

As sentences are added to the training set, the classifier is updated, and feedback is provided

to the coders through the generation of potential theme matches based on the current state of the classifier. Sentences are permitted to contain multiple themes where appropriate. The coders are able to determine how accurately the classifier is performing theme judgments, and expand the number of sentences referring to a theme to improve the classifier accordingly. The coders then compare the performance of their trained classifiers in scoring another set of 2,000 sentences and any areas of disagreement are discussed and resolved. The training sets are then rescored by the same coders to ensure agreement. Finally, the two scored training sets

Table 2. (Continued)

	Strategic themes intended to capture Miles and Snow descriptions (with examples) ^a
Marketing and research pre-eminent Large, transitory elite Broad, exploratory planning Decentralized Peer comparisons Emergent strategic themes	<p><i>Peer comparison:</i> world leader in development diagnostics novel drug therapies treat some worlds ...</p> <p><i>Infrastructure/capabilities:</i> combined elements both building infrastructure business commence ongoing ...</p> <p><i>Governance/management:</i> on behalf board sincerely thank CEO dedicated management team</p> <p><i>Shareholder focus:</i> directors are mindful shareholders investment have therefore outlined profit driven course</p> <p><i>Communities:</i> exciting developments in way manage role in community as responsible corporate citizen</p> <p><i>Employee focus:</i> staff training development programs mentioned last year have continued be developed</p> <p><i>Acquisitions:</i> in addition acquisition opportunities augment existing business or open up new ...</p>

^a These examples read as they appear following the removal of common words by Wordnet.

are combined to train and generate the final classifier.

We use the final classifier to score all relevant annual report sections. The software assigns a probability to each sentence indicating the likelihood that each theme is present in that sentence based on the co-occurrences of words in the sentence. The classifier estimates the probability for each theme based on the probabilities of such co-occurrences in the training set of sentences. These probabilities are aggregated across sentences to calculate a raw estimate of the prevalence of each theme in each annual report. The aggregated probabilities for each report are then adjusted by the number of sentences in the report to provide measures of relative theme presence or 'density' in each report (cf. Abrahamson and Hambrick, 1997). For example, if a report has 30 sentences and the probability for theme A's presence is 0.1 in 10 of the sentences, 0.5 in another 10, and 1.0 in the rest, the prevalence or 'density' of theme A in the report is $((0.1 \times 10) + (0.5 \times 10) + (1.0 \times 10))/30 = 0.53$. Consequently a report receives a high/low score for a theme only when it contains a higher/lower density of that theme relative to other reports; a report in which theme prevalence is close to the average receives a score

close to the standardized mean (i.e., 0.0). Content that is common to annual reports and arguably reflects 'conventional' or impression management content does not therefore contribute to differentiating between reports.

RESULTS

Appendix 1 shows means and standard deviations of the 21 themes prior to standardization across reports, but after adjusting for report length, as well as correlations, including the three financial measures. Two things are notable in Appendix 1. As expected, all themes occur in the text; however, their frequencies vary widely, ranging from Self-Comparison (Financial), which occurs in 0.26 (or about 26%) of sentences, to Communities, found in less than 0.01 (or 1%) of sentences. There are also many small and some moderate correlations (between 0.2 and 0.3) between themes, indicating there may be an underlying structure, which issue we turn to next.

Identifying higher-order structure among strategic themes

At this point we conclude there is evidence that managers attend in annual reports to strategic

themes that Miles and Snow (1978) used to describe main characteristics of their different strategic types, consistent with our first, broad hypothesis. Next we examine whether these themes form higher-order structures that are related to the strategic configurations described by Miles and Snow (1978). We use principal components analysis (SPSS factor—principal components method) to reduce the 21 theme scores for each annual report to a smaller number of higher-order indicators of patterns of strategic cognition. We employ an exploratory principal components approach in preference to confirmatory factor analysis because, while a number of our content themes are based on theory, it is not possible on an *a priori* basis to identify what the factor structure will look like, remembering that few, if any, analyses of this type have been done previously. The present context is quite different from the more typical confirmatory factor analysis problem involving, for example, the

analysis of a set of attitude items thought to be measuring an *a priori* defined construct. We also exclude any reports that do not have at least 25 sentences with a score on one or more themes, leaving 3,663 annual reports from 1,038 unique firms. The principal components analysis is performed at the report level, with reports from any particular organization across multiple years treated as separate observations. We do not aggregate multiple reports from the same firm as that would assume there was continuity of themes and relations between themes across reports from the same firm, which is the subject of our third hypothesis. Seven components with eigenvalues greater than one are retained. Eigenvalues and component loadings from a Varimax rotation of the solution (Table 3) reveal that a reasonably high proportion of variance is captured (55%), and that there is a relatively simple structure with few items having cross-loadings greater than 30. Split-half solutions

Table 3. Principal component analysis of strategic themes

Theme	Innovation/ Expansion	Customer Service vs. Capacity Building	Core Focus	Operational Efficiency	Corporate Social Responsibility vs. Financial Performance	Competitive Context	Restructure
Marketing	0.65						
Product Focus	0.64	0.35					
R&D	0.61						
Peer Comparisons	0.58						
New Markets	0.48					0.46	
Collaboration/ Alliances		−0.76					
Infrastructure/ Capabilities		−0.72					
Customer Service		0.49					
Core Business/ Prudence			0.78				
Shareholders			0.77				
Productivity				0.67			
Cost Reduction				0.66			
Acquisitions				−0.57			
Employees					0.70		
Governance/ Management					0.61		
Communities					0.55		
Self-Comparison	−0.39				−0.48		
Market/Economic Conditions						0.83	
Competition	0.44	0.38				0.47	
Restructuring							0.76
Divestments							0.70
Eigenvalue	3.03	2.17	1.79	1.58	1.31	1.18	1.04
% Total variance	13.80	9.86	8.16	7.18	5.96	5.36	4.72
Cumulative %	13.80	23.66	31.82	39.00	44.96	50.32	55.04

similarly yield seven components and comparable structure.

Validating factor interpretation: an overview

While the factor structure is meaningful and stable, there are reasons to explore its validity further. One reason is that the ML approach is somewhat less transparent than dictionary-based methods where the words and phrases making up each category are explicitly specified. A detailed description of the validation process is set forth in Appendix 2. In brief, it involves four academics reading and rating strategic content in paragraphs from annual reports that are identified as loading on one of the seven factors. On the basis of these results, we are able to conclude that raters can distinguish the presence of the strategic factors in text without any prior knowledge of the nature of the factors, and they can accurately and reliably recognize them when the factors are defined for them.

The seven dimensions of strategic knowledge structures

Based on the theoretical origins of the coding scheme, and the pattern of loadings, the factors are labeled respectively:

- *Innovation/Expansion*: strong emphasis on Marketing, Product Focus, R&D, Peer Comparisons, and New Markets;
- *Customer Service vs. Capacity Building*: a bidirectional factor in which positive scores indicate a stronger Customer Service focus, while negative scores indicate strong Collaboration/Alliances and Infrastructure/Capabilities emphases;³
- *Core Focus*: higher scores indicate an emphasis on Core Business/Prudence, and Shareholders;
- *Operational Efficiency*: the positive score indicates a Productivity and Cost Reduction focus, while a negative score indicates a concern with Acquisitions;
- *Corporate Social Responsibility vs. Financial Performance*: a positive score indicates more

attention to Employees, Governance/Management, and Communities, while a negative indicates more focus on Self-Comparison (Financial);

- *Competitive Context*: indicates concern with Market/Economic Conditions, and Competition; and
- *Restructuring*: indicates a focus on Restructuring and Divestments.

These seven factors form the main dimensions of managerial knowledge structures at an annual report level. Clearly, a number of the dimensions align quite well with Miles and Snow (1978), as we had hoped. For example, Innovation/Expansion seems to capture many prospector attributes, bringing together themes of R&D, marketing, new markets and products, and an external, expansionary focus; Operational Efficiency represents a defender-like element; Core Focus arguably captures some defender- and analyzer-like qualities; while Customer Service may capture an analyzer element. It is not surprising that Miles and Snow's typology, which is concerned with viable, long-term strategies, does not consider restructuring, but restructuring is understandably part of the strategic discourse in annual reports. Overall there is good support for Hypotheses 1 and 2—top teams attend to a variety of strategic issues both internal and external to the organization, and this variety can be encapsulated in terms of a limited set of broader, strategic factors that can be interpreted as representing core features of different generic strategies.

While there is evidence of significant industry differences in some of the factors, we did not adjust factor scores by industry because: (a) it seems plausible that some industries are more concerned than others with certain strategic themes; and (b) while strategies are generic, there is no reason to suppose they occur equally in all industries. Furthermore, the differences we observe are intuitively meaningful. For example, Innovation/Expansion is more prevalent in both Health and IT sectors, Operational Efficiency factor is higher in Materials, while Customer Service is lower in Materials and Energy sectors. None of these differences seem likely to overwhelm within-industry differences.

Identifying cognitive strategic groups

We next employ cluster analysis to see if we can identify cognitive strategic groups with strategic

³ This type of factor structure, also seen in several other instances, indicates that some content themes tend to 'displace' one another, or at least tend not to appear together in the same annual report. This could be the result of a number of factors including industry differences and strategically incompatible themes.)

profiles resembling the Miles and Snow types. This can also be viewed as an analysis of second-order structure among the strategic dimensions, and is thus also relevant to Hypothesis 2. This analysis is performed at the individual report level, so organizations are potentially assigned to different clusters in different years.

The Mclust package (Fraley and Raftery, 2002) in the R statistical environment, a model-based clustering algorithm that determines the optimal number of clusters in a set of cases, and the best distributional model, suggests a six-cluster solution for model VVV (ellipsoidal, varying volume, shape, and orientation) to be optimal (Bayesian Information Criterion (BIC) = 69,563.89). Ten cluster analyses on randomly selected subsamples

of half the cases are also performed, with eight of ten solutions supporting a six-cluster structure. Clusters are given interpretive labels on the basis of average factor scores by cluster, and to reflect their theoretical origins (Figure 2):

- *Restructurers*: highest score of any group on Restructuring, with moderate scores on Corporate Social Responsibility (therefore low on Financial Performance), and Operational Efficiency;
- *Productivity Defenders*: highest score of any group on Capacity Building (therefore low on Customer Service), moderately high on Operational Efficiency, and moderately low on Innovation/Expansion;

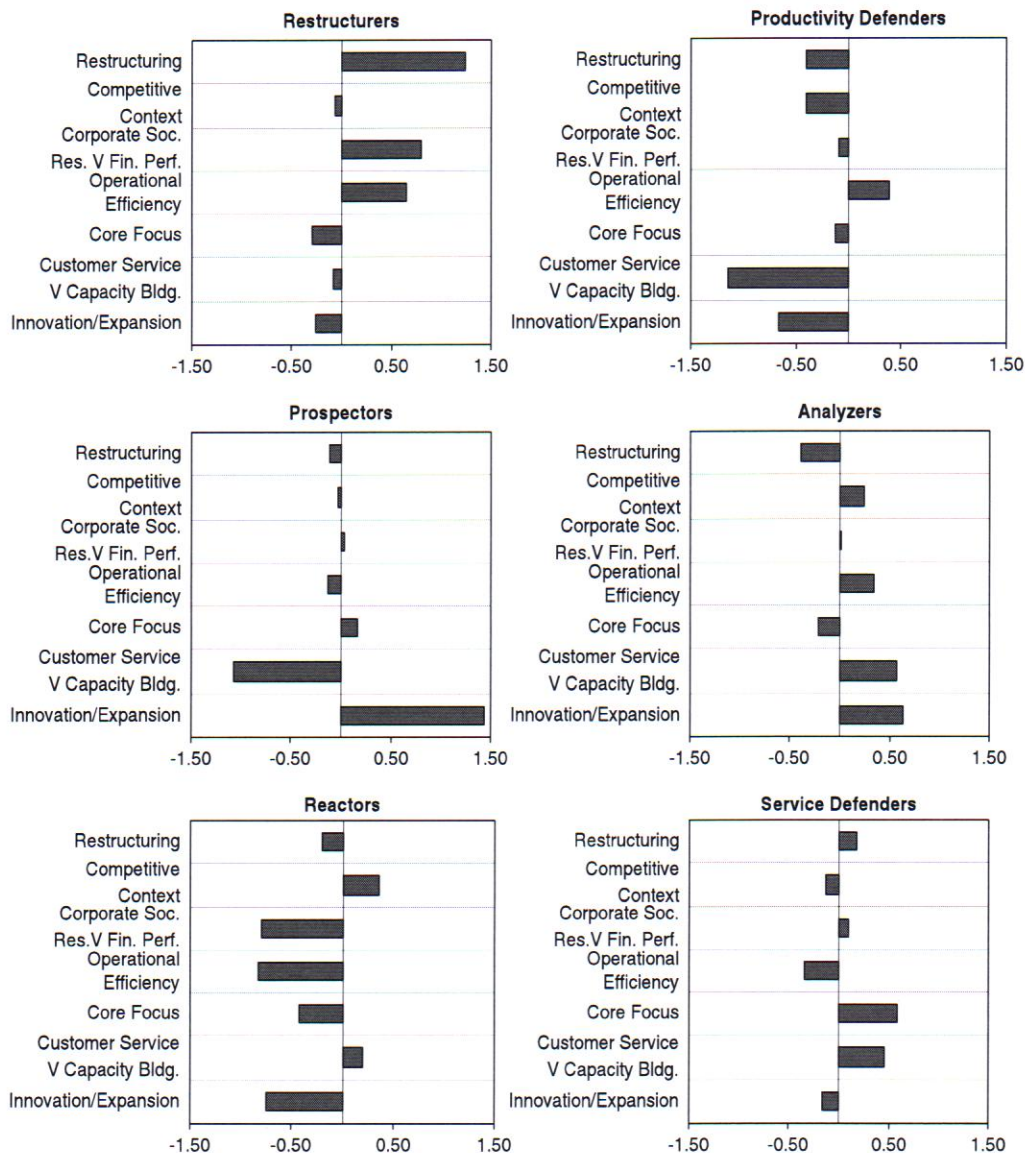


Figure 2. Mean scores on strategic factors of six cognitive strategic groups

- *Prospectors*: highest score of any group on Innovation/Expansion, and also high on Capacity Building (therefore lower on Customer Service);
- *Analyzers*: moderately high scores on Innovation/Expansion, Customer Service (rather than Capacity Building), and some concern with Operational Efficiency, but little Restructuring (it is the combination of innovation, efficiency, and customer elements that provides the rationale for our interpretation of this group);
- *Reactors*: moderately low on Operational Efficiency, Innovation/Expansion, and Corporate Social Responsibility (therefore higher on attention to Financial Performance) and moderately high on Competitive Context (the lack of attention to any elements apart from their own financial performance and external conditions provided the rationale for this interpretation);
- *Service Defenders*: while having a relatively 'flat' profile, the moderately strong focus on Customer Service (rather than Capacity Building) and Core Focus/Shareholder Value by this group suggested this interpretation. A number of previous studies have distinguished between different types of defenders (e.g., Olson, Slater, and Hult, 2005). Overall, these results reinforce the view that there are meaningful, underlying structures to managerial attention and that these patterns can be usefully described in terms of the generic types outlined by Miles and Snow (1978). We emphasize that this is unlikely to be the only way of representing managers' strategic knowledge structures but there is clear evidence that Miles and Snow's typology captures some meaningful differences in the patterns of strategic elements that different top teams focus on.

Stability of strategic knowledge structures

Hypothesis 3 stated that top teams' strategic knowledge structures and firms' membership of cognitive strategic groups tend to be relatively stable over time. Based on this hypothesis, the proportion of firms that are members of a particular strategic group at time one (t_1) and remain members of the same group at time two (t_2) is expected to be significantly different from the proportion that would be expected if allocation was a random process. To test this, it is necessary to demonstrate that strategic group membership across reports for a given organization is not random, but rather

that any organization is likely to be assigned to the same group across time. This does not imply absolute stability, but does imply that group membership does not change randomly.

Determining this is not as simple a problem as it may first appear. Our organizational sample is by definition opportunistic—reports were available for a varying number and composition of large organizations each year (see Table 1). We therefore adopted two complementary approaches to assessing the stability of firms' group membership: one focusing on relatively short-term stability, the other on long-term stability. Both approaches to assessing group stability involved first segmenting the 12-year sampling window into three 4-year blocks (1992–95, 1996–99, 2000–03). To assess short-term stability, cluster membership within each 4-year period was investigated. Longer-term stability was assessed across the three blocks of 4 years. The choice of a 4-year window was based on pragmatic considerations and data constraints—4 years seemed a reasonable time over which to assess short-term stability; it also provided a reasonable number of observations for each period, remembering that the mean number of reports per company was 3.65. All organizations present for at least 3 years of any one of the 4-year blocks were judged to have sufficient data to enable a determination to be made of stability or lack thereof in that period. If an organization had data for all 4 years of a block, three reports were chosen at random so that the number of observations per organization was identical. If an organization had complete data for more than one 4-year block, only one of the blocks, chosen at random, was used in the analysis, meaning that each organization would only feature once in the analysis to preserve the independence of observations.

A measure of stability was defined as the number of different clusters an organization was assigned to within a 4-year block, with '1' signifying stability, '2' some stability, and '3' instability. A total of 460 individual organizations had sufficient data to be included in the short-term stability analysis. Based on overall membership frequency for each cluster, expected frequencies for each of the stability levels were calculated assuming random allocation to a cluster, and compared to the observed frequencies of stability. Figure 3 shows that the proportion of stable firms (31.6%) is clearly much higher than the level that would

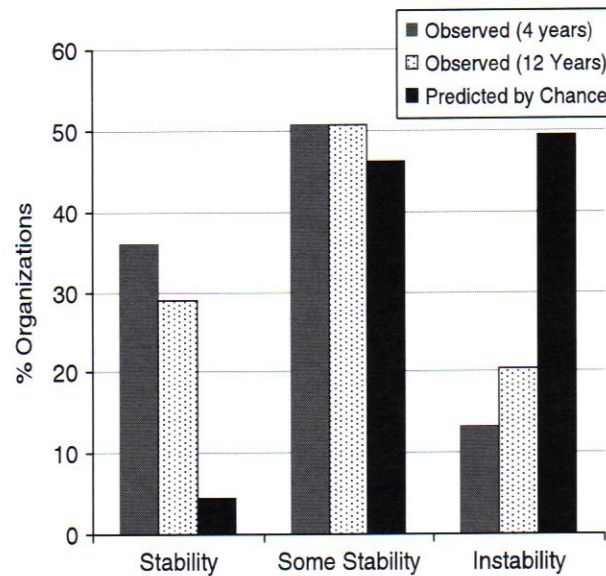


Figure 3. Observed and expected by chance levels of stability of group membership across 4 and 12 years

be expected by chance (4.4%), while the proportion classed as unstable (13.3%) is far lower than would be expected by chance (49.4%). Clearly, organizations' cluster membership is relatively stable across time, at least within 4-year periods. To put this differently, there is evidence that, on average, top teams tend to have a similar pattern of strategic focus in annual reports over a 4-year period.

To assess longer-term stability, organizations with at least one cluster membership (i.e., annual report) in each of the 4-year blocks were selected. If an organization was present in more than 1 year in any block, one of the years was selected at random—the previous results justify treating the two memberships as being the same. Again, expected frequencies for cluster membership were calculated based on the base membership frequencies of each cluster. While the number of organizations is considerably smaller than in the first analysis ($N = 152$), Figure 3 shows that the overall pattern is similar—stability in the longer term also far exceeds what would be expected by chance. Overall, both results indicate that managerial attention patterns are relatively stable over both the short and medium to long term, supporting Hypothesis 3.

Strategic groups and financial performance

We now consider whether firms whose top teams share different, stable knowledge structures differ in their financial performance. In order to test this,

we had once again to overcome the challenge of a dataset that had relatively few cases for which complete data were available; in effect it was a repeated-measures design with a substantial number of missing observations for both cases (firms) and variables (financial data). While this partly reflects limitations of our available datasets, it also reflects that firms' long-term survival is the exception rather than the rule (Stubbart and Knight, 2006). SPSS Linear Mixed Models, a robust technique capable of handling repeated measures data with high levels of 'missingness,' was chosen for the analysis.

Industry and strategic group were employed as between-group fixed effects, while organization was the subject variable, and report year the repeated-measures covariate. Lack of complete cases prevented us from treating time as a separate effect; however, we did not view this as wholly detrimental since the design was already quite complex and we did not articulate a set of hypotheses in relation to time periods. Table 4 shows the overall fixed effects results for ROE, ROA, and PER. In terms of Hypothesis 4, the main results in Table 4 are the ones indicating that there is a significant effect for cluster membership for all three financial measures and also a significant interaction between industry and cluster for all three financial measures, indicating that relative performance of strategic groups differs across industries. The industry effect is not surprising and not integral to this study.

Table 4. Industry and strategic groups' effects on three measures of financial performance

Source	ROA		ROE		PER	
	d.f.	F	d.f.	F	d.f.	F
<i>Fixed factor effects</i>						
Industry	7,166.3	3.3**	7,163.1	2.8**	7,138.8	5.8***
Cluster	5,155.6	12.7***	5,166.5	8.1***	5,154.0	9.4***
Industry \times Cluster	28,138.3	3.9***	28,134.2	2.5***	28,142.9	1.8**
<i>Simple effects (cluster within industry)</i>						
Consumer discretion	5,147.5	2.1†	5,134.7	2.6*	5,109.0	2.2†
Consumer staples	3,764.5	0.3	3,886.9	0.6	3,130.6	1.2
Energy	5,915.3	1.0	5,917.0	0.3	5,133.1	4.6***
Financials	5,141.9	3.7**	5,121.9	12.0***	5,126.3	1.2
Health	3,106.8	43.1***	3,111.7	20.5***	3,113.5	11.9***
Industrials	5,120.6	7.7***	5,116.4	5.0***	5,135.3	0.9
Information technology	2,135.9	0.8	2,138.6	0.6	2,605.8	1.4
Materials	5,681.2	2.0†	5,689.0	1.9†	5,140.1	2.6*

† $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Tests of the simple effects of cluster within each of the eight industry groups reveal significant effects on all three financial measures (Table 4). There are significant differences in ROA and ROE between clusters within the Financial, Health, and Industrial sectors, while Consumer Discretionary shows a significant effect for ROE and trends for both ROA ($p = 0.060$) and PER ($p = 0.055$), with Materials also having trends for ROA ($p = 0.098$) and ROE ($p = 0.082$). There are significant differences between groups in PER in the Energy, Health, and Materials sectors. Thus, there are significant between-group differences on at least one financial measure in six of the eight industries: of 24 significance tests, 10 are clearly significant (at least $p < 0.05$) and a further four have a significant trend ($p < 0.10$). Two industries with no evidence of group differences are IT and Consumer Staples, which both have relatively small numbers (mean number of firms per year is 16 for IT and 24 for Consumer Staples).

Pairwise comparisons between clusters within industry groups identify the actual performance differences in industries. We considered only the four industries where there were consistently significant differences. The Bonferroni test was used to adjust for the large number of individual comparisons. While we do not report all these in detail, a number are worth noting in respect of Hypotheses 4a and 4b. There was good support for Hypothesis 4a particularly in relation to prospectors—firms with a prospector focus performed worse than analyzers in terms of ROE, ROA, or

both in all four industries. The picture for defenders was more mixed and, in general, analyzers did well relative to other types in terms of profitability. Consistent with Hypothesis 4b, prospectors outperformed other groups in Health and Consumer Discretionary sectors in terms of their PER. The Health sector has the highest score on the Innovation/Expansion dimension and Consumer Discretionary also scores quite high, so our results seem consistent with Hambrick's (1983) finding about prospectors' performance in more innovative markets.

DISCUSSION

We find that the content of managerial attention in annual reports can be described in terms of the strategic issues or themes that Miles and Snow (1978) used to describe their strategic types. These themes form higher-order dimensions of strategic cognition that identify cognitive strategic groups broadly similar in their strategic orientation to the Miles and Snow (1978) types. These groups are relatively stable over time and differ in financial performance. The performance differences are quite compelling since the performance measures are conceptually and methodologically independent of our content-based measures of strategic cognition. The fact that we are able to identify stable patterns of strategic attention in annual reports that fit with Miles and Snow's (1978) typology offers further impressive support for their model

using a very different data source and methodology. It also supports the validity of the knowledge structure construct and our measurement by demonstrating a link between the content and structure of managerial attention patterns and a well-established strategic typology.

The performance differences parallel Hambrick's (1983) findings quite closely, which is impressive given the many differences between the studies, including time period, context, and methods. One of the core, still contested, assumptions of Miles and Snow's (1978) framework is that the main strategic types are equally viable across all environments and, by implication, across time (Zajac and Shortell, 1989). This was questioned by Hambrick (1983), and more recently by DeSarbo *et al.* (2005), who endorse the more typical view that different environments favor different strategic types. While our results do not support the original assumption in its strict form, they nevertheless provide rather striking support for the view that broad patterns of performance differences between types are quite persistent over place and time. Several interesting and important questions arise. Are the different types equally likely to occur over time and place, and if not, what factors influence the prevalence of different types? Hambrick (1983) suggested that prospecting may be a short-term strategy, with successful prospectors exploiting their newly established competitive advantage by becoming analyzers or defenders. Zajac and Shortell's (1989) study (now almost 20 years old), which was carried out in a single industry and considered a 2-year time frame only, is one of the few to have examined this issue. The present study suggests ways of studying these issues in the future.

While integral to effective strategizing, knowledge structures are also implicated in the development of strategic blind spots, rigidities, and blunders (Walsh, 1995: 292–293). There are three main explanations for such dysfunctional outcomes. One centers on the accuracy or veridicality of knowledge structures (Walsh, 1995: 303). By definition, knowledge structures represent only a proportion of the total informational environment, and if the reduced informational set is missing elements critical in that environment, negative consequences result. Another explanation focuses on the rigidifying or inertial character of knowledge structures, which can make managers either less likely to perceive and/or to feel it is necessary to respond to

changes in the environment. Gilbert's (2005) distinction between two sources of inertia—resource inertia based on factors external to the organization, and routine inertia primarily based on managers' internal, 'dominant logic'—is a recent, significant contribution in this tradition. A related, yet distinct, explanation is offered in Miller's (1993: 116) notion of the 'architecture of simplicity.' This argues that, over time, the goals, strategies, cultures, and processes of successful organizations become more pure or 'simple' in that they come to focus on an increasingly narrow set of strategic issues, organizational capabilities, values, and processes. This leads to poor outcomes both by increasing inertia in the face of environmental change, and by rendering the organization less capable of dealing with its existing environment. As Miller put it:

Their rich, strategic character will devolve into bland and truncated caricature ... [unable] to embrace and adapt to the complex currents of their settings. (Miller, 1993: 118)

The notion of managers' knowledge structures simplifying or configuring is compatible with Miles and Snow's (1978) adaptive cycle but suggests there is a point at which an ever 'tighter' attentional focus can harm the organization—the adaptive cycle becomes a 'maladaptive cycle' (Walsh, 1995: 303). Our identification of a set of broad and relatively comprehensive themes for describing strategic cognition creates the opportunity for examining several questions related to these explanations for how and why knowledge structures can lead to negative outcomes: is there a tendency for top teams' focus to become narrower over time; is this tendency strengthened by organizational success as Miller (1993) suggests, and, does this increase the risk of organizational decline? A narrowing attentional focus can be indexed by the proportion of total attentional references accounted for by one or more strategic factors over time.

We need to understand what factors influence top teams to develop different kinds of knowledge structures. Hambrick and Mason's (1984) widely studied upper echelons theory proposes that top managers' background, experience, and demographic characteristics are an important influence on the psychological and cognitive 'givens' that shape their strategic decisions. Research testing

this theory has typically used demographic measures as proxies for top-team members' cognitive characteristics, and, as Sparrow observed, it has made the

remarkable discovery... statistically significant predictions can be made from the demographic characteristics of 0.1% of an organization's members. (Sparrow, 1994: 176)

However, the use of demographic proxies has also resulted in an Achilles' heel that Lawrence articulated:

Typically [researchers] have hypothesized various subjective concepts that explain significant association between demographic predictors and outcomes. ... [but they] usually leave the concepts unmeasured and the hypotheses untested. As a result, subjective concepts and their relationships within research models have become the 'black box' of organizational demography (Lawrence, 1997: 2)

Our approach can begin to illuminate the contents of the 'black box' by enabling us to explore the links between top teams' characteristics and strategic cognition instead of inferring cognition from organizational outcomes on a *post hoc* basis.

While future research should compare our findings with those obtained by directly measuring top teams' perceptions of firm strategy and industry structure, an interesting research opportunity lies in exploring when and why cognitive and conventional groups diverge, rather than demonstrating their equivalence. Porac *et al.* described this as follows:

Rather than being an exogenous force acting on managerial minds, market structure is an endogenous product of managerial minds ... While markets are arenas for economic transactions, at their core are routinized thought patterns and interlocked networks of managerial attention. Explanations for market structure must account for these networks by showing how stable cognitive orderings control the flow of managerial attention across organizational fields. At this level of analysis environmental theories of competition fall silent. (Porac *et al.*, 1995: 224–225; emphasis in the original)

Levenhagen *et al.* suggested that industry change can be viewed as the result of entrepreneurs as cognitive agents who first perceive ambiguous and uncertain knowledge spaces and who then discover

or create frame-making or frame-breaking ideas (Levenhagen, Porac, and Thomas, 1993: 77).

Change in the content and structure of managerial cognition may be a harbinger of change in industries and provide the opportunity to study the role of these 'entrepreneurial cognitive agents.'

In conclusion, we believe our study advances research in the field of strategic cognition in two ways: first, by outlining a theory-based model of the content and structure of strategic knowledge structures; and second, by describing a method capable of providing valid, indirect measurement of this content and structure over time. Hopefully, our study will assist researchers in their ongoing quest 'to articulate the intangible and hard to measure contents of the strategic mind' (Porac and Thomas, 2002: 178).

ACKNOWLEDGEMENTS

This research was supported by an Australian Research Council Discovery Grant (grant no. A00106398): chief investigators Boris Kabanoff and Andrew Neal. We thank our colleagues Boaz Bernstein, Jack Keegan, Kavoo Mohannak, and Stephane Tywoniak for their contributions. We also greatly appreciate the insightful comments by Mark Griffin, Sharon Parker, and Per Davidsson.

REFERENCES

- Abrahamson E, Hambrick DC. 1997. Attentional homogeneity in industries: the effect of discretion. *Journal of Organizational Behavior* 18: 513–532.
- Anderson CR, Paine FT. 1978. PIMS: a re-examination. *Academy of Management Review* 3: 602–612.
- Barney JB, Hoskisson RE. 1990. Strategic groups: untested assertions and research proposals. *Managerial and Decision Economics* 11(3): 187–198.
- Daft RL, Weick KE. 1984. Toward a model of organization as interpretive systems. *Academy of Management Review* 9: 284–295.
- D'Aveni RA, MacMillan IC. 1990. Crisis and the content of managerial communications: a study of the focus of attention of top managers in surviving and failing firms. *Administrative Science Quarterly* 34: 634–657.
- DeSarbo WS, Di Benedetto CA, Song M, Sinha A. 2005. Revisiting the Miles and Snow strategic framework: uncovering interrelationships between strategic types, capabilities, environmental uncertainty, and firm performance. *Strategic Management Journal* 26(1): 47–74.

- Doty DH, Glick WH, Huber GP. 1993. Fit, equifinality, and organizational effectiveness: a test of two configurational theories. *Academy of Management Journal* 36: 1196–1250.
- Duriau VJ, Reger RK, Pfarrer MD. 2007. A content analysis of the content analysis literature in organization studies: research themes, data sources, and methodological refinements. *Organizational Research Methods* 10: 5–34.
- Fellbaum C (ed.). 1998. *Wordnet*. MIT Press: Cambridge, MA.
- Fraley C, Raftery AE. 2002. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* 97: 611–631.
- Gilbert CG. 2005. Unbundling the structure of inertia: resource versus routine rigidity. *Academy of Management Journal* 48: 741–763.
- Hambrick DC. 1983. Some tests of the effectiveness and functional attributes of Miles and Snow's strategic types. *Academy of Management Journal* 26: 5–26.
- Hambrick DC. 2003. On the staying power of defenders, analyzers, and prospectors. *Academy of Management Executive* 17: 115–118.
- Hambrick DC, Mason PA. 1984. Upper echelons: the organization as a reflection of its top managers. *Academy of Management Review* 9: 193–206.
- Hodgkinson GP. 2001. The psychology of strategic management: diversity and cognition revisited. In *International Review of Industrial and Organizational Psychology*, Cooper C, Robertson IT (eds). Wiley: Chichester; 65–120.
- Huff AS, Jenkins M (eds). 2002. *Mapping Strategic Knowledge*. Sage: London.
- Ketchen DJ, Thomas JB, Snow CC. 1993. Organizational configurations and performance: a comparison of theoretical approaches. *Academy of Management Journal* 36: 1278–1313.
- Lawrence BS. 1997. The black box of organizational demography. *Organization Science* 8: 1–22.
- Levenhagen M, Porac JF, Thomas H. 1993. Emergent industry leadership and the selling of technological visions: a social constructionist view. In *Strategic Thinking: Leadership and the Management of Change*, Hendry J, Johnson G (eds). Wiley: New York; 69–87.
- Levy O. 2005. The influence of top management team attention patterns on global strategic posture of firms. *Journal of Organizational Behavior* 26: 797–819.
- Lewis P, Thomas H. 1990. The linkage between strategy, strategic groups, and performance in the U.K. retail grocery industry. *Strategic Management Journal* 11(5): 385–397.
- Miles RE, Snow CC. 1978. *Organizational Strategy, Structure and Process*. McGraw-Hill: New York.
- Miller CC, Burke LM, Glick WH. 1998. Cognitive diversity among upper-echelon executives: implications for strategic decision processes. *Strategic Management Journal* 19(1): 39–58.
- Miller D. 1993. The architecture of simplicity. *Academy of Management Review* 18: 116–138.
- Misangyi VF, Elms H, Greckhamer T, Lepine JA. 2006. A new perspective on a fundamental debate: a multilevel approach to industry, corporate and business unit effects. *Strategic Management Journal* 27(6): 571–590.
- Moors A, De Houwer J. 2006. Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin* 132: 297–326.
- Morris R. 1994. Computerized content analysis in management research: a demonstration of advantages and limitations. *Journal of Management* 20: 903–931.
- Nath D, Gruca TS. 1997. Convergence across alternative methods for forming strategic groups. *Strategic Management Journal* 18(9): 745–760.
- Ocasio W. 1997. Towards an attention-based view of the firm. *Strategic Management Journal*, Summer Special Issue 18: 187–206.
- Olson EM, Slater SF, Hult GT. 2005. The performance implications of fit among business strategy, marketing organization structure, and strategic behavior. *Journal of Marketing* 69: 49–65.
- Osborne JD, Stubbart CI, Ramaprasad A. 2001. Strategic groups and competitive enactment: a study of dynamic relationships between mental models and performance. *Strategic Management Journal* 22(5): 435–454.
- Palmer I, Kabanoff B, Dunford R. 1997. Managerial accounts of downsizing. *Journal of Organizational Behavior* 18: 623–639.
- Porac J, Mishina Y, Pollock TG. 2002. Entrepreneurial narratives and the dominant-logics of high-growth firms. In *Mapping Strategic Knowledge*, Huff AS, Jenkins M (eds). Sage: London; 112–136.
- Porac J, Thomas H. 2002. Managing cognition and strategy: issues, trends and future directions. In *Handbook of Strategy and Management*, Pettigrew A, Thomas H, Whittington R (eds). Sage: London; 165–181.
- Porac J, Thomas H, Badden-Fuller C. 1989. Competitive groups as cognitive communities: the case of the Scottish knitwear manufacturers. *Journal of Management Studies* 26: 397–415.
- Porac JF, Thomas H, Wilson F, Paton D, Kanfer A. 1995. Rivalry and the industry model of Scottish knitwear producers. *Administrative Science Quarterly* 40: 203–227.
- Porter ME. 1980. *Competitive strategy*. Free Press: New York.
- Reger RK, Huff AS. 1993. Strategic groups: a cognitive perspective. *Strategic Management Journal* 14(2): 103–124.
- Sebastiani F. 2002. Machine learning in automated text categorization. *ACM Computing Surveys* 34: 1–47.
- Smith AE, Humphreys MS. 2006. Evaluation of unsupervised semantic mapping of natural language with Leximancer concept mapping. *Behavior Research Methods* 38(2): 262–279.
- Sparrow PR. 1994. The psychology of strategic management: emerging themes of diversity and cognition. In *International Review of Industrial and Organizational Psychology*, Cooper C, Robertson IT (eds). Wiley: Chichester; 147–181.
- Stone PJ, Dunphy DC, Smith MS, Ogilvie DM. 1966. *The General Inquirer: A Computer Approach to Content Analysis*. MIT Press: Cambridge, MA.

- Stubbart CI, Knight MB. 2006. The case of the disappearing firms: empirical evidence and implications. *Journal of Organizational Behavior* **27**: 79–100.
- Walsh JP. 1995. Managerial and organizational cognition: notes from a trip down memory lane. *Organization Science* **6**(3): 280–321.
- Weber RP. 1990. *Basic Content Analysis*. Sage: London.
- Williams K. 2003. AI: Categorizer-0.07 automatic text categorization (computer software). <http://search.cpan.org/~kwilliams/AI-Categorizer-0.07/> [July 2003].
- Zajac EJ, Shortell SM. 1989. Changing generic strategies: likelihood, direction, and performance implications. *Strategic Management Journal* **10**(5): 413–430.

APPENDIX 1: Descriptive Statistics and Correlations for Themes and Performance

Variable	Mean ^a	S.D.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
1. Cost Reduction ^b	0.05	0.04	1.00										
2. Productivity ^b	0.03	0.03	0.28**	1.00									
3. Marketing ^b	0.01	0.02	-0.11**	0.04*	1.00								
4. Self Comparison (Financial) ^b	0.26	0.13	-0.10**	-0.25**	-0.16**	1.00							
5. Peer Comparison ^b	0.02	0.03	-0.18**	0.02	0.31**	-0.24**	1.00						
6. Collaborations/Alliances ^b	0.03	0.04	0.09**	-0.03	0.06**	-0.25**	0.05**	1.00					
7. Acquisitions ^b	0.04	0.05	-0.19**	-0.22**	-0.05**	0.08**	-0.07**	-0.10**	1.00				
8. Restructuring ^b	0.01	0.02	0.00	0.12**	0.03*	-0.04*	-0.04**	-0.08**	-0.04*	1.00			
9. Divestment/Closure ^b	0.01	0.01	0.15**	0.06**	-0.08**	0.03	-0.10**	0.05**	0.04*	0.21**	1.00		
10. Infrastructure/Capabilities ^b	0.11	0.11	0.22**	-0.02	-0.16**	-0.21**	-0.15**	0.38**	-0.06**	-0.17**	0.06**	1.00	
11. New Markets ^b	0.02	0.02	-0.19**	0.02	0.18**	-0.23**	0.27**	-0.01	-0.05**	0.00	-0.07**	-0.14**	1.00
12. Competition ^b	0.06	0.05	-0.23**	0.05**	0.21**	-0.24**	0.27**	-0.19**	-0.10**	0.09**	-0.11**	-0.42**	0.34**
13. Research and Development ^b	0.01	0.03	-0.03	0.07**	0.22**	-0.25**	0.21**	0.27**	-0.09**	-0.01	-0.04*	-0.04**	0.14**
14. Product Focus ^b	0.02	0.03	-0.14**	0.10**	0.30**	-0.23**	0.19**	-0.09**	-0.03	0.04*	-0.07**	-0.26**	0.26**
15. Customer Service ^b	0.05	0.05	-0.05**	0.14**	0.14**	-0.28**	0.11**	-0.16**	-0.09**	-0.01	-0.13**	-0.29**	0.12**
16. Core Business/Prudence ^b	0.06	0.04	-0.14**	-0.04*	-0.03	-0.13**	0.02	-0.12**	-0.02	0.13**	0.01	-0.28**	0.04*
17. Market/Economic Conditions ^b	0.17	0.08	-0.12**	-0.04*	-0.03	-0.13**	-0.03*	-0.19**	0.00	-0.01	-0.03*	-0.27**	0.11**
18. Governance/Management ^b	0.09	0.07	-0.12**	-0.04**	0.01	-0.29**	0.03	-0.06**	-0.04*	0.05**	-0.09**	-0.18**	-0.07**
19. Shareholders ^b	0.05	0.04	-0.09**	-0.10**	-0.10**	-0.14**	-0.03	-0.10**	-0.06**	0.00	-0.08**	-0.17**	-0.09**
20. Communities ^b	0.00	0.01	0.12**	0.15**	0.00	-0.16**	0.06**	0.01	-0.12**	0.02	0.03	0.03*	-0.03
21. Employees ^b	0.04	0.03	-0.04*	0.16**	0.02	-0.27**	0.08**	-0.04*	-0.23**	0.01	-0.09**	-0.02	0.03
22. ROA ^c	3.95	14.10	-0.12**	-0.02	-0.03	0.18**	-0.05*	-0.23**	0.07**	-0.02	-0.01	-0.11**	0.01
23. ROE ^d	8.77	28.88	-0.05**	-0.01	-0.03	0.12**	-0.04*	-0.21**	0.01	-0.02	-0.05**	-0.13**	0.05*
24. PER ^c	20.47	25.08	0.02	-0.01	0.03	-0.09**	0.04*	0.15**	0.00	-0.03	0.02	0.09**	0.02

APPENDIX 1: (Continued)

Variable	Mean ^a	S.D.	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.
12. Competition ^b	1.00												
13. Research and Development ^b	0.09**	1.00											
14. Product Focus ^b	0.32**	0.24**	1.00										
15. Customer Service ^b	0.25**	0.03	0.36**	1.00									
16. Core Business/Prudence ^b	0.16**	-0.04*	0.01	0.05**	1.00								
17. Market/Economic Conditions ^b	0.22**	-0.09**	-0.02	-0.08**	-0.04*	1.00							
18. Governance/Management ^b	-0.05**	0.01	-0.04*	0.01	0.11**	-0.17**	1.00						
19. Shareholders ^b	0.02	-0.05**	-0.08**	0.03*	0.38**	-0.08**	0.20**	1.00					
20. Communities ^b	-0.06**	-0.03	-0.02	0.15**	-0.09**	-0.04**	0.02	0.00	1.00				
21. Employees ^b	0.03	0.03	0.07**	0.20**	-0.01	-0.17**	0.20**	0.08**	0.21**	1.00			
22. ROA ^c	0.05**	-0.22**	0.00	0.02	0.03	0.02	-0.03	0.01	-0.02	0.01	1.00		
23. ROE ^d	0.09**	-0.20**	0.02	0.16**	0.01	-0.02	-0.04*	0.01	0.06**	0.07**	0.70**	1.00	
24. PER ^e	-0.01	0.14**	-0.04	-0.02	-0.01	-0.02	0.00	0.01	0.00	-0.01	-0.20**	-0.14**	1.00

^a These means reflect theme frequencies adjusted for document length but prior to standardization across all annual reports. It should be noted that this is on an annual report and not firm basis; that is, there are multiple observations for most organizations.

* $p < 0.05$ (2-tailed); ** $p < 0.01$ (2-tailed); ^b $n = 3, 663$; ^c $n = 3, 065$; ^d $n = 3, 048$; ^e $n = 2, 201$.

APPENDIX 2: VALIDATION OF STRATEGIC FACTORS

Four strategic management academics took part in a two-stage validation of the uncovered factor structure involving reading whole paragraphs from 210 annual reports that had the largest loadings on one of the seven strategic factors (i.e., 30 paragraphs for each factor); with the proviso that at least five industries were represented in each set of 30. Each set of 30 paragraphs was further split into two subsets of 15, so that we were able to assess consistency or 'test-retest reliability' of judgments. The average number of words per paragraph was 73, with a range of 36–104. The four judges received a general overview of the nature of our research and written instructions that we sought their assistance in identifying the main strategic emphasis within the extracts. The first task involved reading each set and providing a brief written description of the main strategic orientations they perceived in each set.

Judges were then given a 'factor key' that named and described each one of the seven factors by describing the main themes loading on the factor; for example, Operational Efficiency was described as loading on two themes: 'Costs: reducing/minimizing/cutting costs, losses, overheads and expenses' and 'Productivity: increasing productivity, mechanical/operational efficiency, streamlining processes, improving production technology.' We actually included eight dimensions in the factor key, the eighth being a 'dummy' factor we named 'Stable Administration Structures' that we believed was plausible in this context (it comes from Miles and Snow, 1978). We used this mild deception in order to examine the extent to which raters' judgments were influenced by our suggesting the presence of a theme in the extracts. Raters were asked to read the factor descriptions carefully, look at each set of 30 paragraphs again, and

rate the extent to which each factor was present in each of the 30 sets using a five-point, Likert-type scale.

Stage one provided clear, qualitative indications that raters perceived that each set had content relating to the factor we had identified. For example, for the first set of 15 paragraphs loading on the Operational Efficiency dimension all four raters described the strategic orientation in terms such as: 'Themes are operational. Cost reduction. Improved efficiency. Continuous improvement. Productivity gains.' Stage two provided quantitative evidence in support of this view. This evidence takes three forms: evidence of *inter-rater* accuracy and agreement in their identification of the 'correct' factor content in each subset of paragraphs; evidence that they correctly rated the 'dummy' factor as having low prevalence in all the paragraphs; and finally, evidence of *intra-rater* reliability in how each one rated both halves of the 30 paragraphs. On the five-point Likert-type scale (1 = not at all; 5 = a very large extent) the average rating for the 'correct factor' in each set was 4.8, while for the 'incorrect factors' it was 2.7. The mean rating for the 'dummy factor' was 1.8. The mean ratings indicate a high degree of accuracy among raters as a group, with the highest average values always being obtained by the factor for which paragraphs had been selected. The dummy factor rated lowest for all but one set. Correlations between set one and set two scores for individuals ranged between $r = 0.60$ and $r = 0.81$. Reliability for the scale as a whole, averaging corresponding items from the two sets, was high (intra-class correlation coefficient = 0.88). These results show that, both with and without prompting, our four raters agreed that the strategic orientation that best typified each set was the factor intended to be represented by that set.