

활용과 탐색 측정을 위한 방법론으로써 콘텐츠 분석 :세계 광디스크 라이브러리장치 산업

Content Analysis as a Method for Measuring Exploitation and Exploration:
Discussion with Example Application to the Worldwide Optical Library Industry

유건재
홍익대학교 경영학부

Gun Jea Yu(morz9492@gmail.com)

요약

활용과 탐색을 측정하는 방법은 특허 기반의 방법, 설문 기반의 방법, 언론 기반의 방법 세 가지로 나뉜다. 다양한 방법론이 나오게 된 이유는 활용과 탐색의 정의에 대한 합의가 아직 이루어지지 않았기 때문이다. 활용과 탐색 활동은 기업이 장기적으로 행하기 때문에 언론 기반의 방법이 가장 적절하며, 이 논문에서는 세계 광디스크 라이브러리 장치 산업의 데이터를 활용하여 유효하고, 신뢰성 있는 언론 기반의 방법에 의해 활용과 탐색을 어떻게 측정하는 것인지를 보여주고자 한다.

■ 중심어 : | 활용과 탐색 | 방법론 | 콘텐츠 분석 |

Abstract

Measures of exploration and exploitation fall into three categories, patent-based measures, survey-based measures, and press-based measures. Such variety stems from the lack of consensus on definitions of exploration and exploitation. Given a dynamic nature of exploration and exploitation, I suggested how to improve existing press-based measures by suggesting strategies and procedures for constructing valid and reliable press-based measures in a single industry context. I illustrate my arguments through a study of the worldwide optical disk industry

■ keyword : | Exploration and Exploitation | Methodology | Content Analysis |

I. Introduction

Since March[1] proposed the concepts of exploitation and exploration, a diverse range of operationalizations have emerged for these two concepts, and there is little agreement on the ideal measurement approach. Research using survey-based

measures captures exploration and exploitation through cross-sectional snapshots[2-4]. As a result, this work does not capture the evolution of exploitation and exploitation over time. This is potentially problematic because organizations adjust their exploratory and exploitative activities over time

* This work was supported by 2014 Hongik University Research Fund.

접수일자 : 2014년 06월 19일

수정일자 : 2014년 07월 15일

심사완료일 : 2014년 07월 17일

교신저자 : 유건재, e-mail : morz9492@gmail.com

to cope with environmental change. Some studies have used patent-based measures to capture exploitation and exploration[5-7]. Patent data have considerable variation in terms of representation of a firm's technology because firms in different industries have different strategies to protect their technologies. Such variation results in less generalizability across industries[8].

Recently, Uotila et al.[9] suggested press-based measures using news documents as an alternative to survey-based and patent-based measures of exploitation and exploration. This method is designed to broaden the coverage of firm actions aimed at exploitation and exploration, increase generalizability, and measure these activities longitudinally. However, the press-based measures Uotila et al.[9] used also have limitations such as limited coverage by news sources, the selection of search terms for exploration and exploitation, and sample bias.

This study contributes to the current exploration and exploitation literature by three folds. First, this study suggests the way to improve press-based measures, the most recently developed approach, in greater depth and apply the technique to a study of innovation in a single industry. This implies that this study helps researchers to use more validated and reliable measures for exploration and exploitation. Second, this study provides an overview of the operationalization of exploration and exploitation. I review the empirical strategies, the data sources, and the construction of various measures that have been employed by past scholars. I identify three distinct approaches -- 1) survey-based measures; 2) patent-based measures; and 3) press-based measures -- and highlight the strengths and weaknesses of each. Third, this study also adds to the empirical evidence for the use of press-based measures and shows the usefulness of contents analysis in the

management literature.

II. Theoretical Background

2.1 The Definitions of Exploration and Exploitation

March[1] developed the original concepts of exploitation and exploration. Exploration is associated with "search, variation, risk-taking, experimentation, play, flexibility, discovery, and innovation," whereas exploitation refers to activities such as "refinement, choice, production, efficiency, selection, implementation, and execution". These definitions of exploration and exploitation can be interpreted in various ways due to their quite broad scope. In a subsequent study, Levinthal and March [9] narrowed down the scope of the two activities to the knowledge domain, noting that exploitation refers to "the use and development of things already known," whereas exploration is related to "a pursuit of new knowledge". Although subsequent studies have expanded the notions of exploration and exploitation, this study used Levinthal and March's[9] definitions of exploration and exploitation and March's[1] for the operationalization of exploration and exploitation.

The effects of exploitation and exploration on organizational performance have been the primary concerns in the exploration and exploitation literature. Researchers have argued that the impacts of exploration and exploitation on organizational performance differ in terms of long-term and short-term outcomes. Exploration activities help the firm search for new capabilities and enhance long-term performance, such as survival[11]. However, the chance of a payoff is uncertain and distant compared with that of exploitation. Thus, the uncertain long-term performance from exploration

can lead to vicious cycles (i.e., failure of exploration results in more exploration)[12]. On the other hand, exploitation enhances short-term performance by increasing efficiency, decreasing performance variance, and improving adaptation to the current environment. However, high efficiency and reduced variation might become an organizational liability that prevents firms from adapting to changing environments.

2.2 Three Approaches to Measuring Exploration and Exploitation

The increasing attention to exploration and exploitation has contributed to conceptual refinement[16]. In addition, many empirical studies have provided evidence of conceptual work. However, less attention has been given to methodology, and the understanding of methodology in the exploration and exploitation literature has remained under-explored. In this study, I intend to clarify what measures are applicable to exploration and exploitation studies and to review the strengths and weaknesses of different measures. Before reviewing the measures, criteria to evaluate them will be addressed.

2.2.1 Three criteria to evaluate measures of exploration and exploitation

Three different approaches to measuring exploration and exploitation have been used in the exploration and exploitation literature: survey-based measures, patent-based measures, and press-based measures. These three approaches are compared using three criteria: easiness of collecting longitudinal data, generalizability, and validity. In this section, I will define the three criteria and explain why these three criteria were adopted.

One of the important criteria is whether or not a

measure of exploration and exploitation can provide longitudinal data. A static perspective on exploitation and exploration has been prevalent[2][3][13-18]. This perspective assumes that organizations pursue exploration and exploitation activities to cope with demands of internal and external environments and that these two activities are constant over time [16]. However, organizations need to adjust their exploration and exploitation activities to satisfy changing demands in their internal and external environments[20][21]. The static perspective has limitations in that it cannot capture the evolution of exploration and exploitation. Methodologies using longitudinal data can provide the historical pattern of exploitation and exploration and help to clarify their evolution.

Generalizability is defined as the extent to which inferences can be made from a sample and extended to the population as a whole and is often referred to as one criterion for the quality of measures[22]. Lack of generalizability derives from various sources such as the research sample, the time of study, and the operational definitions[22]. In this study, I evaluated generalizability across industries in terms of operational definitions because the three measures using archival data are limited in terms of operationalization and generalizability across firms within industry by examining whether a measure can be applicable to diverse types of firms within industry.

Content validity is one of two validity tests based on the subjective evaluation of an operational definition. Content validity is defined as “the extent to which a measure adequately represents all facets of a concept”[22]. In this paper, content validity is defined as the extent to which a measure of exploration and exploitation represents the concepts of exploration and exploitation that March[1] proposed, because

most researchers have conceptualized exploration and exploitation based on his definitions. With regard to archival data (i.e., patent-based measures, and press-based measures), judgments of content validity can be made by comparing March's[1] conceptualizations of exploration and exploitation and the operationalizations of exploration and exploitation made using the three approaches that rely on archival data because most studies in the exploration and exploitation literature conceptualized and operationalized the two concepts based on the ones March[1] defined. The content validity of survey-based measures can be made by analyzing content validity of previous studies that used surveys. Validity testing is frequently used as a criterion to judge the quality of various measures.

2.3 Reviews of the three measures of exploitation and exploration

I reviewed the three approaches to measuring exploration and exploitation using three criteria. [Table 1] summarizes the strengths and weaknesses of the three approaches based on these three criteria.

Table 1. Comparison of three Approaches for Measuring Exploration and Exploitation

Criteria	Survey-based measure	Patent-based measure	Press-based measure
Longitudinal data	M*	H	H
Generalizability across industries	H	L*	M
Generalizability across firms within industry	H	M	H
Content Validity	H	M	H

H (High): Fully support, M (Medium): Support, but not fully, L (Low): Hardly support

*These are not inherent weaknesses of measures, but a weakness with the design choices in the exploration and exploitation literature

2.3.1 Survey-based measures of exploration and exploitation

A large number of studies have used surveys of key personnel to capture firms' exploitation and exploration activities[3][4][14-19]. Most studies operationalized exploration and exploitation using market and product concepts. For example, He and Wong[3] measured exploitation by asking whether organizations focused on cost reduction and existing products and measured exploration by inquiring whether their products or services were targeted at new markets or customers or they entered into new technology fields. Subsequent studies which adopted a survey methodology have utilized similar operationalizations. In sum, operationalization using technology and market concepts has been dominant in survey-based measures of exploration and exploitation and studies using surveys focused on the organizational level to reveal the consequences or antecedents of exploration and exploitation.

Studies using survey-based measures enhanced content validity by asking managers for items of exploration and exploitation, which are developed based on March's[1] definitions and by conducting factor analysis of exploration and exploitation items. For example, He and Wong[3] developed items for exploration and exploitation at the organization level and asked key personnel to inspect the items for exploration and exploitation in terms of appropriateness, redundancy, and ambiguity. Next, they ran factor analysis to determine whether items captured exploration and exploitation distinctly. Many subsequent studies provided evidence of the content validity of survey-based measures of exploration and exploitation. This implies that survey-based measures have a high content validity. Some studies using surveys were conducted in multiple industries. This implies that studies using survey-based

measures can develop survey items to cover multiple industries and these items can be generalizable across industries. In addition, the samples in several studies included both large and small firms. Thus, survey-based measures are also generalizable across firms within industry. However, surveys are less applicable for obtaining longitudinal data than for patent data because it is difficult for informants to retrospectively recall exploitation and exploration activities.

2.3.2 Patent-based measures of exploration and exploitation

Some studies have relied on patent data as objective proxies for exploration and exploitation [5-7][13][22]. These studies operationalized exploration and exploitation mainly using citations in patents. For example, Katila and Ahuja[12] measured exploitation as “the average number of times a firm repeatedly used the citations in the patents it applied for” and measured exploration as “the proportion of previously unused citations (new citations) in a firm’s focal year’s list”. Wang and Li[7] constructed exploration as the number of citations in a firm’s patent that came from outside of the firm’s main class of patented technology and constructed exploitation as the number of citations in the patent that originated within the main class of the firm’s patented technology.

Patent data have several strengths as a measure of exploration and exploitation. Patents provide rich and detailed information about technologies such as the people, places, times, and technological characteristics involved in new product development. In addition, patent data are not static, providing longitudinal data. When a patent is submitted for approval, the patent should show the citations that the applicant referred to. Citations in patents provide antecedents and

descendants of patented technologies[23].

Patent-based measures also have limitations in that patent data do not fully represent firms’ innovative activities. First, although patent data can provide technological information about firms’ innovation, firms’ innovative activities have been conceptualized very broadly, including search behavior[6] and entrance into new markets[3] as well as technological advancement. Thus, operationalization using patent-based measures may be too narrow to measure exploration and exploitation. Such narrow operationalization implies that patents may not represent all facets of exploration and exploitation defined by March[1]. In addition, diverse operationalizations of exploitation and exploration using patent data mean that there has been little consensus on how to measure both of these activities. Therefore, I argue that patent-based measures for exploration and exploitation have a low content validity. Second, not all innovative activities are patented. Each firm has a different strategy and motivation to manage its technological assets[23]. Some firms do not patent their technologies or inventions if they can protect them in other ways such as know-how[24]. Such differences result in considerable variation across industries in terms of the number of patents and the value of patents. For example, in the pharmaceutical and chemical industries, inventions are protected by a small number of patents, whereas in the electronics industry, a large number of patents protect inventions. This also implies that researchers should control for unobserved heterogeneity in patent data if multiple industries are studied. Thus, patent-based measures for exploration and exploitation are not generalizable across industries. In addition, firms need resources and personnel to manage their patents. Large firms may have more resources and personnel than small ones.

Thus, there is some variation across firms within industry in terms of the number of patents. This implies that researchers also need to control unobserved heterogeneity in patent data if they used the patent-based measures within industry.

2.3.3 Press-based measures of exploration and exploitation

The use of content analysis in management studies has been growing[25]. Content analysis enables researchers to explore a wide range of strategy topics such as downsizing[26], corporate reputation[27], and strategy reformulation. Recent research suggests that a simple quantitative count of words in news documents that capture a negative evaluation of firms can be used to predict individual firms' financial performance[28]. In addition, simple word counting using computers minimizes reliability problems in text classification because computerized simple word counting is highly reproducible[26].

Most studies which utilized content analysis to measure exploration and exploitation fit into a specific context (e.g., firm alliance) or a specific industry (e.g., professional service)[9][18][29]. Uotila and colleagues[9] were the first to use press-based measures for exploitation and exploration in order to make their operationalization applicable to multiple industries. They developed measures that "(1) cover a broad scope of corporate actions, (2) are available for a large number of companies over an extended period of time, and (3) are applicable across a range of industries". The operational definitions of exploration and exploitation that Uotila and colleagues[9] adopted stems from March's notion of exploitation and exploration. This suggests that press-based measures of exploration and exploitation can capture exploration and exploitation relatively accurately because the conceptualizations of

exploration and exploitation in most research were done based on the notions of exploration and exploitation by March [1]. In addition, the collection of longitudinal data is relatively easy because news articles are highly accessible.

However, the press-based measures of exploration and exploitation that Uotila and colleagues[9] employed are not generalizable across industries, although they argue otherwise. On the other hand, the press-based measures that Uotila and colleagues[9] used are generalizable across firms within industry if visibility of a firm is controlled because large firms are more likely to appear in news documents. I will delineate why their measures of exploration and exploitation are not generalizable across industries in the following section.

In sum, the three approaches for measuring of exploration and exploitation have strengths and limitations. A comparison of the three approaches indicates that press-based measures of exploration and exploitation are the most reliable measures of exploration and exploitation for studies conducted in a single industry (See [Table 1]). Press-based measures can provide longitudinal data based on accurate operationalization of exploration and exploitation. In addition, data for press-based measures are highly accessible. However, press-based measures need to be carefully used when researchers consider measures applicable across industries.

Patent-based measures can also offer longitudinal data. However, the narrow operational definitions of exploration and exploitation are a weakness of patent-based measures. Thus, patent-based measures are recommended when technological aspect in innovation is the focus of study and longitudinal approach is needed.

Although surveys can provide longitudinal data, it

is difficult to collect longitudinal data using surveys. However, survey-based measures have high levels of generalizability across and within industry and validity. Therefore, survey-based measures are recommended when researchers need high levels of generalizability and validity, but don't need longitudinal approach. In the following section,

I will explain how press-based measures of exploration and exploitation can be used in a single industry by improving on the press-based measures used by Uotila et al [9] and how to adjust press-based measures in other industry.

III. The Use of Press-based Measures for Exploration and Exploitation

A broad range of methods and analytical techniques have been categorized as content analysis[30]. Shapiro and Markoff proposed a concise and complete definition of content analysis: "any methodological measurement applied to text (or other symbolic materials) for social science purposes". According to Deffner[31], content analysis is classified into three types: (1) human scored schema; (2) individual-word-counting systems; and (3) computerized systems using artificial intelligence. Uotila and colleagues [9] adopted the computerized individual-word-counting method to measure exploration and exploitation. Words in a text are assigned to pre-specified, semantically similar categories in individual-word-counting systems[32]. Although human coders can conduct this type of analysis, computerized coding systems are superior because they ensure higher reliability and better cost effectiveness[33].

Uotila and colleagues[9] collected news documents from the Reuters News archive in the Factiva

database. Exploitative activities were captured by counting the number of times March's eight words appeared in news documents for each company-year and exploratory activities were measured by counting the number of times March's nine words appeared in news documents for each company-year. Uotila and colleagues[9] performed several validity and reliability tests to determine whether the seventeen words identified actual activities of exploration and exploitation because previous studies did not use content analysis for the operationalization of these activities.

The strengths of the approach that Uotila and colleagues[9] suggested are twofold. First, operationalizations using press-based measures of exploration and exploitation are highly generalizable compared with those of survey or patent data methodologies. The search terms for content analysis cover a broader range of firms' exploration and exploitation activities, including strategic action as well as technological aspects, while patent-based measures capture only technological activities. Second, researchers can implement a longitudinal research design because news documents for content analysis are available over a long period of time.

While the press-based measures of exploration and exploitation that Uotila and colleagues[9] used have several strengths, they also have weaknesses. First, Uotila and colleagues[9] argued that their press-based measures of exploration and exploitation are highly generalizable across industries. However, the search terms they used may not be generalizable because not all search terms may represent exploration and exploitation in every industry. For example, 'play', which is one of the search terms for exploration, can have various meanings such as to make something operate (as in to play a CD), a performance of a role or character, a form of amusement, and to engage in

competitive sports (as in to play against another team or player). In the video game industry (e.g., PlayStation and Xbox), 'play' usually means a form of amusement or making something operate. Another example is 'search'. 'Search', which is one of the search terms for exploration, is widely used in the Internet industry, including the term 'search engine'. This usage suggests that 'search' is not relevant to explain exploration in the Internet industry. Therefore, the appropriate search terms for measuring exploration and exploitation are likely to vary by industry. This implies that press-based measures of exploration and exploitation are not as generalizable as Uotila et al. [9] suggested.

Second, the selection of news sources may need to be altered in order to capture exploration and exploitation in small firms. Uotila et al.[9] collected data from 279 manufacturing firms in the Standard & Poor 500 index. The firms that they chose are relatively large. Large firms may have higher visibility than small ones in major news sources such as Reuters News, and some small firms may not be covered by one particular news source. They suggested that it would be worthwhile to expand their results to small and medium size firms.

3.1 Enhancement of a content analytic measure for exploration and exploitation

I used the same search terms for exploration and exploitation used by Uotila and colleagues[9] to add empirical evidence from a different sample and to find a way to tailor the operationalization of exploitation and exploration using content analysis in a particular industry or sector. I collected data for exploration and exploitation in the worldwide optical library industry.

The optical library industry is an appropriate setting for investigating the effects of exploration and exploitation for several reasons. First, this industry

has experienced high rates of innovation change (e.g., changes in optical disk technology and network technology). Second, the data on this industry cover every firm during the period 1990 to 1998, making it possible to avoid sample selection bias. Third, firms in the optical library industry have been diverse in terms of size, providing an appropriate setting for extending the results of Uotila and colleagues[9] to small firms.

The population I studied (optical library industry) included 100 firms from 1984 to 1998, which generated a total of 461 firm-year observations. Among these 100 firms, 15 firms were de novo (start-up firms), 85 firms were de alio (diversified firms); 24 firms were public, and 72 firms were private (three firms went public during this period and one firm was separated from the parent firm, resulting in its going private). Companies were located in many different countries (Canada, France, Germany, Israel, Japan, the Netherlands, Singapore, South Korea, Spain, Sweden, Taiwan, the United Kingdom, and the United States), among which the U.S. and Japan hosted 74% of all producers over time. The average sales of public firms were 2.3 times those of private firms.

Content analysis was done using "Concordance,"[34] which is a software package widely used in many disciplines including accountancy, history, marketing, musicology, politics, geography, and media studies.

The individual firm data were collected by year. I used the same word roots that Uotila and colleagues [9] used. Exploration was captured by the word roots explor*, search*, variation*, risk*, experiment*, play*, flexib*, discover*, and innovat* (The wildcard '*' can represent any character). Exploitation was captured by the word roots exploit*, refine*, choice*, production*, efficien*, select*, implement*, and execut*. All words that have

the word roots included in the search were identified from collected news documents by firm-year using content analysis software. However, inappropriate forms of words were identified after checking all of the words that appeared in the software: executive(s) for 'execut*', player(s), playstation and playback for 'play*', and explorer(s) for 'explor*'. Because these words do not represent exploration or exploitation activities, they were excluded from the calculation of the number of exploitation and exploration words. The total number of exploitation words (60,618) surpassed the total number of exploration words (33,419) from 1984 to 1998. Over this time period, the total number of firms in the optical library industry increased from 2 firms to 67 firms.

I pointed out that the press-based measures of exploration and exploitation that Uotila and colleagues [9] used have two weaknesses: the selection of journals and the selection of word roots used to search news documents. I will explain the process I used to overcome these two weaknesses.

The selection of word roots. One weakness of content analysis for the operationalization of exploitation and exploration is that not all word roots may represent exploration and exploitation in every industry. To determine the appropriate word roots that capture exploration and exploitation of firms for the optical library industry, an understanding of the industry is needed.

The key product of the optical library industry is an optical library, also called an optical jukebox or an optical tower, designated for storage and retrieval of massive amounts of data. An optical library is designed to automatically pick, load, unload, and re-file media units for an optical disk drive to write or read. Optical libraries are useful for record retention, backup systems, desktop publishing and many more applications.

The term 'play' is frequently used in the optical disk industry as a synonym for reading an optical disk. The term 'search' also often appeared with the meaning of finding an optical disk in an optical tower that contains multiple optical disks. Therefore, these two terms needed to be examined to determine whether they represent exploration activities of firms in the optical disk industry.

In the optical disk industry, 'play' is the dominant word root among the word roots that capture exploration activities, accounting for 27%. However, the meaning of play in most of the news documents that were searched is making something operate. I found that relatively few of the mentions of 'play' were related to exploration activities. Therefore, it is appropriate to exclude 'play' from the exploration word roots as a search term.

Examples:

"The bus supports hot-swapping, and its recognition support allows plug-and-play operation. This type of peripheral is worth examining this year, but don't expect a flood of products immediately. SCSI will remain the best choice for at least a few more years. Maxopix 1998"

"NEC has a 4x speed CD-ROM with 7 disc changer that allows you to play music CDs while running software applications. NEC 1995"

An optical library is used to write or read information. To read information, optical libraries must have a searching function. In addition, 'search' refers to a search engine in many news articles. Only a small portion of news items identified using the term 'search' were in fact related to exploration activities. The following examples show when the term 'search' does not mean exploration:

Table 2. The Number of Exploration and Exploitation Words by Year (All Firms)

	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
1	105	472	639	925	1,318	2,691	2,850	4,534	4,405	5,704	7,627	5,114	6,385	7,838	10,011
2	14	133	201	347	653	1,169	1,433	2,336	2,815	3,736	3,598	2,708	3,200	4,812	6,264
3	119	605	840	1272	1,971	3,860	4,283	6,870	7,220	9,440	11,225	7,822	9,585	12,650	16,275
4	2	6	6	9	12	17	24	28	28	37	44	49	59	73	67
5	60	101	140	141	164	227	178	245	258	255	255	160	162	173	243

1. The number of exploitation words
2. The number of exploration words
3. Sum of exploration and exploitation words
4. The number of firms
5. Words per firm

“It comes bundled with Alchemy Personal, text-based file retrieval software that lets you build databases of images and documents that can be searched and displayed in under three seconds, and Easy-CD Pro, which lets you copy data or audio files to CD from Windows File Manager or Explorer. Kintronics 1997”

“End users can search across approved supplier catalogs through an interface on the company’s intranet. Fustitju, 1998”

Therefore, the terms ‘play’ and ‘search’ should be excluded from the data because these terms have different meanings that cannot be interpreted as exploration.

The selection of journals. As mentioned above, the visibility of small and large firms is different across news sources. Large firms are more likely to appear in sources such as The Wall Street Journal and Reuters News. To capture small firms’ exploitation and exploration, the range of publications needed to be expanded. I collected textual data in the form of all news articles and newswire publications in the Factiva database from 1984 to 1998. A total of 109,490 news documents were collected. In the first data collection, all publications in the Factiva database were included to identify publications which cover all firms in the optical disk industry. 180 journals were

identified in the first analysis. Then, out of these 180 journals, I selected 103 journals that fully covered the years from 1990 to 1998. Finally, news articles were collected from these 103 journals with the 15 terms noted above, excluding ‘play’ and ‘search’. The 103 journals included major journals (e.g., Dow Jones News Service and PR Newswire), industry specialized journals (e.g., CD Computing News and CD-ROM Professional), and regional journals (e.g., Kyoto News and Business Times Singapore).

[Table 2] describes the number of firms, total number of words that measure exploration of all firms, total number of words that measure exploitation of all firms, and the sum of exploration and exploitation words of all firms by year. The total number of words increased by a factor of 137, from 119 (1984) words to 16,275 (1998). The total number of words per firm also rose from 60 words to 243 words. While the number of exploitation words increased by a factor of 95, from 105 (1984) to 10,011 (1998), the number of exploration words rose from 14 (1984) to 6,264 (1998), an increasing of 447 percent. While exploitation is more prevalent in the optical library industry, the relative portion of exploration to the sum of exploration and exploitation activities increased.

3.2 Validity tests

To establish the convergent validity of the

exploration and exploitation score, I measured exploration and exploitation using patent data. The exploration and exploitation scores using patent data were calculated using the approach of Wang and Li [7]. The exploration score in a certain year was calculated by the number of citations in a firm's patent that were outside of the firm's main class of patented technology, whereas the exploitation score in a certain year was calculated by the number of citations in the patent that were within the main class of the firm's patented technology. The firm's main class of patented technology was determined by the class of patents that the firm applied for before a certain year.

I computed the correlation between the exploration and exploitation scores using content analysis and the exploration and exploitation scores using patent data by firm-year (28 firms out of the 100 have patents in the optical disk industry). The correlation of the exploration scores between the two measures is 0.77 (p=0.00) and that of the exploitation scores is 0.77 (p=0.00), suggesting a high level of convergent validity.

To assess the content validity of the exploration and exploitation measures, fourteen firms were selected out of the 100 firms in 1998. The sample consisted of 5 public firms and 9 private firms to reflect the ratio of public to private firms in the population. Two coders read all news in the Factiva database of the 14 firms selected for the validity test to determine whether these news items captured exploitation and exploration activities based on the concepts of exploration and exploitation that March [1] proposed. If a coder made a decision that a certain news article did not represent exploration and exploitation activities, this entire news article was removed from the news articles for content analysis.

The correlation between this manual classification and the automated classification was calculated. The correlation of exploitation is 0.67 (p=0.01) and that of exploration is 0.93 (p=0.00). This high correlation provides evidence of validity for the automated content analysis. In addition, the correlation of exploration is higher than that of exploitation. This implies that words that capture exploration are a more accurate measure than words that capture

Table 3. Descriptive Statistics and Correlation Coefficients

No	Variable	M	SD	1	2	3	4	5	6	7	8	9	10
1	Relative portion of exploration	0.45	0.19	1									
2	Relative portion of exploitation ²	0.24	0.21	-0.18*	1								
3	firm number of patent (t)	1.09	2.07	-0.17*	0.93*	1							
4	Density all firms (t)	50.52	17.12	0.36*	-0.15*	-0.2*	1						
5	WW industry sales	380.5	134.7	-0.09*	0.08	0.08	-0.61*	1					
6	Joint venture dummy	0.03	0.18	-0.21*	0.05	0.05	-0.26*	0.3*	1				
7	Public firm dummy	0.37	0.48	0.19*	0.21*	0.2*	-0.14*	0.03	-0.12*	1			
8	Firm entry mode dummy	0.14	0.35	0.05	0.17*	0.15*	-0.1*	0.06	-0.09*	0.02	1		
9	Firm number of products (t)	6.18	6.44	0.05	0.07	0.06	-0.24*	0.2*	0.15*	-0.2*	0.2*	1	
10	Firm number of new products (t)	2.48	3.88	-0.06	0.04	0.03	-0.21*	0.2*	0.14*	-0.2*	0.3*	-0.1	1

* p<.05

exploitation.

Additionally, I tested the theoretical expected relationship between exploration and exploitation scores and organization level variables such as age and public status for further validity test. First, I calculated the correlation between exploration (exploitation) and both age and public status. Research suggests that old firms prefer exploitation to exploration[18]. Thus, the correlation between exploitation and age is expected to be higher than the correlation between exploration and age. The results show that the correlation between exploration and age is .62 ($p=0.00$) and the correlation between exploitation and age is .53 ($p=0.00$). To test the significance of the difference between these two correlations, I used the Fisher r -to- z transformation. The results show that there is a significant difference between the two correlations ($z=1.83$, $p=.03$).

Public firms have responsibilities and obligations to their shareholders and thus are more likely to be subject to structural inertia. As a result, they are more likely to be involved in exploitation than in exploration. The correlation between public status and exploitation is .61 ($p=0.00$) and the correlation between public status and exploration is .52 ($p=0.00$). To test the significance of the difference between

these two correlations, I used the Fisher r -to- z transformation. The results show that there is a marginally significant difference between the two correlations ($z=1.4$, $p=.08$).

I also replicated the empirical test which Uotila et al[9] conducted using regression analysis [35]. They investigated relationship between the relative portions of exploration over the sum of exploration and exploitation and long-term organizational performance. In this test, the long-term organizational performance is not available. Thus, innovation performance measured by patent is used. [Table 4] presents statistics and correlations for variables. I checked the multicollinearity, using variance inflation factors (VIF). The range of VIF factors is from 1.14 to 8.02, which is below the rule-of-thumb cut-off of 10. Model 1 shows the results for control variables. Among control variables, public firms are positively related to firm's innovation performance and the number of products is positively related to firm's innovation performance. Model 2 shows the results of main effect. This model shows that there is a curvilinear relationship between the portions of exploration and innovation performance. The effect of the portions of exploration on innovation performance is same with the results of Uotila et al.[9].

Table 4. Results of Regression Analyses for Innovation Performance (DV: Ln firm number of patent)

	Model 1	Model 2
Number of firm (t)	-0.011 (-0.006)	-0.018 (-0.008) *
WW industry sales [in mln. U.S.\$] (t)	0.000 (-0.001)	0.000 (-0.001)
Joint venture dummy = 1	0.531 (-0.351)	0.182 (-0.339)
Public firm dummy = 1	2.426 (-0.559) ***	2.144 (-0.599) ***
Firm entry mode dummy: de novo = 1	-0.630 (-0.430)	-0.922 (-0.644)
Firm number of products (t)	0.074 (-0.036) *	0.100 (-0.054)
Firm number of new products (t)	-0.051 (-0.030)	-0.071 (-0.039)
Relative portion of exploration		4.675 (-1.533) **
Relative portion of exploration²		-2.762 (-1.206) *
Constant	0.13 (-0.35)	-1.002 (-0.583)
R-squared	0.48	0.51
F improvement of fit	4.32	4.54
N	394	394

* $p < .05$; ** $p < .01$; *** $p < .001$

3.3 Reliability test

Seventeen word roots were used to identify exploration (eight word roots) and exploitation (nine word roots) in the news documents collected. The portion of individual exploitation (exploration) word roots out of the total number of exploration (exploitation) word roots during 1984 to 1998 was calculated. For example, 'production' was the most dominant word root out of the eight word roots that capture exploitation, consisting of about 59% of all exploitation words. 'Choice (10%)', 'implement (11%)' and 'select (13%)' followed 'production'. The rest of the word roots, 'efficiency (8%)', 'execute (3%)', 'exploit (3%)' and 'refine (3%)', made up less than 10%. In the exploration word roots, 'play (26%)' was the most dominant. 'Flexibility (14%)', 'innovation (15%)', 'risk (13%)', and 'search (13%)' followed 'play'. The rest of the word roots such as 'discover (6%)', 'explore (5%)', 'variation (2%)', and 'experiment (6%)' comprised less than 10% of the total.

A single influential word such as 'production' or 'innovation' can have spurious effects on results. For example, the word root 'production' accounted for 59% of the seven word roots for exploitation and 'innovation' accounted for 15% of the six word roots for exploration. If such influential words have spurious effects on results, the exploration and exploitation scores using the multiple word roots do not measure exploration and exploitation activities consistently.

The exploration and exploitation scores were calculated by excluding each word root from the previous content analysis. The correlations among the ten exploration scores and the nine exploitation scores were calculated. The correlations among the ten exploration scores are between 0.97 and 0.99. The correlations among the nine exploitation scores are between 0.85 and 0.99. These results support the

reliability for measures of exploration and exploitation using content analysis.

IV. Conclusions

Research in the exploration and exploitation literature has depended on survey-based measures and patent-based measures. The survey-based measures have been used extensively by researchers because they measure exploration and exploitation in terms of market and product conceptualizations. However, survey-based measures cannot easily provide longitudinal data, which make it possible to understand the evolution of exploration and exploitation. Patent-based measures can provide longitudinal data on exploration and exploitation. However, patent-based measures do not fully represent a firm's innovative activities[23].

Press-based measures were adopted to overcome the weaknesses of the two approaches mentioned above[9]. This approach provides broad coverage of firms' activities for exploration and exploitation and enables the collection of longitudinal data.

However, the content analytic approach Uotila and colleagues[9] used also has its limitations in that the selection of news journals did not adequately capture exploration and exploitation of both large and small firms, and some of their search terms that represent exploitation and exploration are not appropriate in every industry. In addition, although Uotila and colleagues[9] argued that the operationalization of exploration and exploitation using content analysis is highly generalizable, I showed that the operationalization Uotila et al.[9] used is not generalizable across industries because the search terms need to be altered according to industry.

In this paper, I compared three approaches for

measuring exploration and exploitation. In addition, I showed how to improve on the press-based measures used by Uotila and colleagues[9] and to use them in a single industry study. First, the press-based measures used by Uotila and colleagues[9] have the fewest weaknesses when studies focus on a single industry. However, in cross industry studies, researchers need to compare strengths and weaknesses of the three different measures of exploration and exploitation. Second, researchers need to change the selection of journals and search terms in order to accurately capture exploration and exploitation using the press-based measures that Uotila et al.[9] used. The selection of journals should be determined according to the industry, time period, and characteristics of the research sample (size and nationality).

Seventeen terms were suggested to capture exploration and exploitation. However, some terms do not accurately measure exploitation and exploration in particular industries. Thus, the selection of search terms also should be based on an analysis of the particular industry being studied.

In addition, search terms in service industries such as banking and restaurants need to be chosen carefully. For example 'production' is the most dominant search word in this study. It is reasonable to choose 'production' as an exploitation word because previous studies examined the effect of exploration and exploitation in the manufacturing industry[9]. However, 'production' may not represent exploitation activities in service industries because service is provided or offered, not produced, to satisfy the needs of customers. Thus, researchers need to fully understand the industry they are examining to choose appropriate search terms.

This research provides an improved way for utilizing press-based measures for exploitation and

exploration in a single industry study. Uotila et al.[9] opened new avenues for the measure of these activities. The content analysis they used captured exploration and exploitation in a relatively easy way because the data were collected from publicly available news articles[9]. This study adds to the empirical evidence for the appropriateness of press-based measures and suggests ways to accurately capture exploration and exploitation using this approach. I hope that this paper motivates researchers to enhance their measurement of exploitation and exploration using a content analysis approach.

참고 문헌

- [1] J. G. March, "Exploration and exploitation in organizational learning," *Organization Science*, Vol.2, No.1, pp.71-87, 1991.
- [2] A. S. Alexiev, J. J. P. Jansen, F. A. J. Van den Bosch, and H. W. Volberda, "Top Management Team Advice Seeking and Exploratory Innovation: The Moderating Role of TMT Heterogeneity," *Journal of Management Studies*, Vol.47, No.7, pp.1343-1364, 2010.
- [3] Z. L. He and P. K. Wong, "Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis," *Organization Science*, Vol.15, No.4, pp.481-494, 2004.
- [4] J. J. P. Jansen, M. P. Tempelaar, F. A. J. van den Bosch, and H. W. Volberda, "Structural Differentiation and Ambidexterity: The Mediating Role of Integration Mechanisms," *Organization Science*, Vol.20, No.4, pp.797-811, 2009.
- [5] P. G. Audia, and J. A. Goncalo, "Past success and creativity over time: A study of inventors

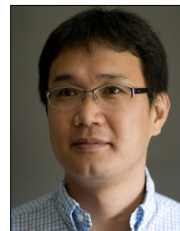
- in the hard disk drive industry,” *Management Science*, Vol.53, No.1, pp.1-15, 2007.
- [6] L. Rosenkopf and A. Nerkar, “Beyond local search: Boundary-spanning, exploration, and impact in the optical disk industry,” *Strategic Management Journal*, Vol.22, No.4, pp.287-306, 2001.
- [7] H. L. Wang and J. T. Li, “Untangling the effects of overexploration and overexploitation on organizational performance: The moderating role of environmental dynamism,” *Journal of Management*, Vol.34, No.5, pp.925-951, 2008.
- [8] A. Gambardella, D. Harhoff, and B. Verspagen, “The value of European patents,” *European Management Review*, Vol.5, No.3, pp.85-89, 2008.
- [9] J. Uotila, M. Maula, T. Keil, and S. A. Zahra, “Exploration, exploitation, and financial performance: Analysis of S&P 500 corporations,” *Strategic Management Journal*, Vol.30, No.2, pp.221-231, 2009.
- [10] 박재민, 이종만, “기업의 혁신 활동이 기업성과에 미치는 영향”, *한국콘텐츠학회논문지*, 제11권, 제3호, pp.339-350, 2011.
- [11] D. A. Levinthal and J. G. March, “The Myopia of Learning,” *Strategic Management Journal*, Vol.14, pp.95-112, 1993.
- [12] R. Katila and G. Ahuja, “Something old, something new: A longitudinal study of search behavior and new product introduction,” *Academy of Management Journal*, Vol.45, No.6, pp.1183-1194, 2002.
- [13] Z. L. He and P. K. Wong, “Exploration vs. exploitation: An empirical test of the ambidexterity hypothesis,” *Organization Science*, Vol.15, No.4, pp.481-494, 2004.
- [14] S. Auh and B. Menguc, “Balancing exploration and exploitation: The moderating role of competitive intensity,” *Journal of Business Research*, Vol.58, No.12, pp.1652-1661, 2005.
- [15] J. J. P. Jansen, D. Vera, and M. Crossan, “Strategic leadership for exploration and exploitation: The moderating role of environmental dynamism,” *Leadership Quarterly*, Vol.20, No.1, pp.5-18, 2009.
- [16] S. Raisch and J. Birkinshaw, “Organizational ambidexterity: Antecedents, outcomes, and moderators,” *Journal of Management*, Vol.34, No.3, pp.375-409, 2008.
- [17] J. J. P. Jansen, F. A. J. Van den Bosch, and H. W. Volberda, “Managing potential and realized absorptive capacity: How do organizational antecedent’s matter?,” *Academy of Management Journal*, Vol.48, No.6, pp.999-1015, 2005.
- [18] F. T. Rothaermel and D. L. Deeds, “Exploration and exploitation alliances in biotechnology: A system of new product development,” *Strategic Management Journal*, Vol.25, No.3, pp.201-221, 2004.
- [19] J. S. Sidhu, H. W. Volberda, and H. R. Commandeur, “Exploring exploration orientation and its determinants: Some empirical evidence,” *Journal of Management Studies*, Vol.41, No.6, pp.913-932, 2004.
- [20] N. Siggelkow and D. A. Levinthal, “Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation,” *Organization Science*, Vol.14, No.6, pp.650-669, 2003.
- [21] M. L. Tushman and P. Anderson, “Technological discontinuities and organizational environments,” *Administrative Science Quarterly*, Vol.31, No.3, pp.439-465, 1986.
- [22] R. Singleton and B. C. Straits. *Approaches to*

- social research*. New York: Oxford University Press, 2010.
- [23] M. Gittelman, "A Note on the Value of Patents as Indicators of Innovation: Implications for Management Research," *Academy of Management Perspectives*, Vol.22, No.3, pp.21-27, 2008.
- [24] W. Cohen, R. Nelson, and J. Walsh, *Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not)*. NBER. Cambridge, MA. 2000.
- [25] V. J. Duriau, R. K. Reger, and M.D. Pfaffer, "A content analysis of the content analysis literature in organization studies - Research themes, data sources, and methodological refinements," *Organizational Research Methods*, Vol.10, No.1, pp.5-34, 2007.
- [26] I. Palmer, B. Kabanoff, and R. Dunford, "Managerial accounts of downsizing," *Journal of Organizational Behavior*, Vol.18, pp.623-639, 1997.
- [27] C. Fombrun and M. Shanley, "What's in a name - reputation building and corporate strategy," *Academy of Management Journal*, Vol.33, No.2, pp.233-258, 1990.
- [28] P. C. Tetlock, M. Saar-Tsechansky, and S. Macskassy, "More than words: Quantifying language to measure firms' fundamentals," *Journal of Finance*, Vol.63, No.3, pp.1437-1467, 2008.
- [29] P. Ming, "Thriving in the New: Implication of Exploration on Organizational Longevity," *Journal of Management*, Vol.36, No.6, pp.1529-1554, 2010.
- [30] N. K. Denzin and Y. S. Lincoln, *Handbook of qualitative research*, Thousand Oaks, Calif.: Sage Publications. 1994.
- [31] G. Deffner, "Microcomputers as aids in GOTTSCHALK-GLESER ratings," *Psychiatry Research*, Vol.18, No.2, pp.151-159, 1986.
- [32] R. P. Weber, *Basic content analysis*, Newbury Park, Calif.: Sage Publications. 1990.
- [33] S. D. Rosenberg, P. P. Schnurr, and T. E. Oxman, "Content analysis: A comparison of manual and computerized systems. *Journal of Personality Assessment*," Vol.54, No.1, pp.298-310, 1990.
- [34] P. Warwick, N. Mercer, R. Kershner, and J. K. Staarman, "In the mind and in the technology: The vicarious presence of the teacher in pupil's learning of science in collaborative group activity at the interactive whiteboard," *Computers in Education*, Vol.55, No.1, pp.350-362, 2010.
- [35] 대맹, 강호정, "중국기업의 사회적 책임이 기업 가치에 미치는 영향", *한국콘텐츠학회논문지*, 제 13권, 제12호, pp.422-427, 2013.

저 자 소 개

유 건 재(Gun Jea Yu)

정회원



- 2000년 2월 : 서강대학교 경영학과(학사)
 - 2008년 5월 : Purdue University (경영학석사)
 - 2013년 8월 : Cornell University (School of Industrial and Labor Relations(ILR) 박사, 거시조직 행동 전공)
 - 현재 : 홍익대학교 경영학과 조교수
- <관심분야> : 혁신 (탐색과 활용), R&D 매니지먼트