

# 정보검색에서 실험연구 방법을 이용한 연구의 분석\*

## Experiments in Information Retrieval

곽 철 완(Chul-Wan Kwak)\*\*

### 목 차

- |                  |              |
|------------------|--------------|
| 1. 서 론           | 3.1 사용된 시스템  |
| 2. 분석방법          | 3.2 연구분석의 대상 |
| 2.1 자료의 수집       | 3.3 변인       |
| 2.2 자료의 분석       | 4. 결 론       |
| 3. 정보검색 실험연구의 분석 |              |

### 초 록

실험연구방법을 사용한 정보검색분야의 연구를 분석하여, 이를 각 항목에 따라 체계화하였다. 사용된 연구 방법은 국내외 정보검색 분야의 논문 중에서 실험연구 방법을 이용한 연구 논문을 연구자 임의로 수집하여, 각 논문을 사용된 시스템, 연구분석의 대상, 변인으로 크게 구분한 후, 구체적으로 무엇이 사용되었는가 체계적으로 나열하고, 범주화하였다. 특히, 연구분석의 대상은 크게 검색결과물, 탐색질의, 이용자 등으로 구분하였으며, 변인은 인간 변인과 비인간 변인으로 구분되고, 각각은 독립변인과 종속변인으로 나누어 기술하였다.

### ABSTRACT

The purpose of this study is to analyze the papers of experiments in information retrieval for systematic categorization. The result shows that broad categories are systems, units of analysis, variables. It found three subcategories in the units of analysis: search results, queries, user. Variables are divided into human variable and non-human variables and each variables are divided into independent and dependent variables as subcategories.

키워드 : 정보검색, 실험연구, 연구방법 정보검색에서 실험연구 방법을 이용한 연구의 분석

\* 이 연구는 1998년도 강남대학교 일반연구비 지원에 의하여 연구되었음.

\*\* 강남대학교 문헌정보학과 전임강사

■ 논문접수일 : 1998년 11월 19일

## 1. 서론

정보검색은 도서관 활동의 중심이다. 모든 도서관 활동은 직접적으로 혹은 간접적으로 정보검색활동과 연관되어 있다. 즉, 장서개발은 검색되어질 정보를 선택하고 수집하는 과정이며, 정보조직은 정보를 체계적으로 배열하여 저장하고 이용자에게 접근될 수 있도록 한다. 정보봉사는 이용자가 어떤 정보를 요구하며, 요구 될 것인지 혹은 어떤 형식의 정보를 요구하는지에 관련된다. 그리고, 다양한 종류의 정보검색시스템은 정보를 검색하기 위해 사용된다.

이와 같이 도서관 활동의 핵심인 정보검색에 관하여 몇 가지 질문이 가능하다. 시스템 측면에서 보면, 어떻게 정보가 검색되는가? 이용자 측면에서 보면, 어떻게 정보를 검색해야 하는가? 그리고, 공통적으로 어떻게 하면 정보를 가장 잘 검색할 수 있는가? 여기에서, 시스템 측면은 정보검색시스템의 발전에 관련된 답을 제공한다. 이용자 측면은 도서관의 본질적인 내용에 의해 여러 가지 답을 제공한다. 그리고, 가장 잘 검색한다는 것은 과학적인 방법을 통하여 검증하는 활동이다.

과학적인 연구방법중의 하나인 실험연구는 변인들의 함수관계를 비교하기 위해 설계된 기법이다. 특히, 데이터는 두 가지 혹은 세 가지의 다른 상황에서 개체를 관찰하여 얻은 것이다. 이 개체가 연구분석의 대상이 된다. 일반적으로 연구 질문은 개체의 상태가 어떤 원인에 의한 것인지 혹은 그 원인과 어떠한 관련이 있는가에 초점을 맞춘다. 개체의 상태는 원인에 의해서 변하거나, 원인과 연결되어 변하거나, 혹은 원인과 함께 변하기도 한다. 그리고, 이 모든 가능한 연결들은 각각 중요할 수 있거나 혹은 그렇지 않을 수 있다.

이러한 내용으로 볼 때, 정보검색에 대한 실험 연구 내용을 분석하는 작업은 중요한 의미를 가지고 있다. 정보를 잘 검색하기 위해서, 과거에 무엇이 연구되었으며, 어떤 변인이 사용되고 그 변인들의 관계는 어떠한가를 파악하는 작업은 정보검색을 연구하는 연구자에게 기초 지식을 제공한다. 그러므로, 본 연구의 목적은 정보검색분야 연구에서 실험연구방법을 사용한 논문들이 무엇을, 어떻게 연구하였는가 분석하려 한다.

## 2. 분석방법

### 2.1 자료의 수집

분석 대상은 국내 5가지와 국외 52가지의 논문들을 선택하였는데, 80년대 이후의 것을 연구자 임의로 선정하였다. 하지만 몇몇 논문들은 그 이전 것도 선택되었는데, 이는 본 논문의 목적과 부합되는 것으로 한정하였다. 최근의 논문으로 한정된 이유는 오늘날 주류를 이루고 있는 정보검색 연구 대상이 전자매체를 비추어 현실적인 활용을 염두에 두었기 때문이다. 그리고 각각의 논문에 대한 서술적인 면이 중시되었다. 즉, 문헌정보학 연구자들이 무엇을 연구하였는가 이해하는데 초점을 맞추었다. 또한, 사회과학분야에서 실험연구를 기술한 Campbell and Stanley(1969)의 구분에 따른, 순수한 실험연구로 한정하지 않고 광범위한 실험연구를 포함하였다. 이는 도서관 현장의 연구자가 실제 활용할 수 있거나 문헌정보학의 현실적인 제한점과 관련되기 때문이다.

### 2.2 자료의 분석

#### 2.2.1 정보검색 실험연구에 사용된 시스템

본 연구에서 분석된 논문들은 크게 색인, OPAC, 데이터베이스(온라인 탐색), 기타의 네 가지 집단으로 구분되었다. 본질적으로 이 네 가지 집단은 서로가 연결되어 있다. 예를 들면, OPAC의 경우, 색인과 관련되며, OPAC은 데이터베이스의 한 종류이다. 하지만, 도서관 활동은 일반적으로 현장 지향적인 분야이기 때문에, 이론적인 관련성은 강조되지 않았다.

### 2.2.2 정보검색 실험연구에서 연구분석의 대상 및 실험참가자 선정

연구분석의 대상에 관해서는 두 가지의 질문을 기준으로 사용하였다. 첫째, 무엇이 연구분석의 대상이며, 정보검색 연구에서 사용된 연구분석의 대상은 무엇인가? 둘째, 어떻게 연구분석의 대상이 선정되었는가? 즉, 실험참가자가 어떻게 선정되었는가에 대한 설명이다.

연구분석의 대상이란 어떤 변인이 측정되고 분석되는 “개체”이다. 연구분석의 대상은 크게 검색 결과물, 탐색질의, 사람, 기타 등으로 범주화되었다.

실험참가자 선정은 연구자가 무엇을 연구하기 원하느냐와 밀접한 관련이 있다. 즉, 연구질문이 무엇에 대한 것인가 하는 내용이다. 실제로 연구에서 실험참가자의 선정은 수집된 데이터가 과연 그 연구의 신뢰도(reliability)와 타당도(validity)를 충족시키는가에 커다란 영향을 미친다.

전형적인 연구설계의 목적은 두 가지 질문에 대해 답을 준다. 첫째, 우리가 보는 현실이 우리가 조사하고자 생각했던 것과 일치하는가? 둘째, 우리가 조사하는 내용이 그 연구의 외적 상황에 대한 정보를 우리에게 줄 수 있는가? 첫 번째 질문은 내적 타당도(internal validity)로서 원인과 결과의 형태로 정의되어 있으면 내적으로 타당하

다고 말한다. 두 번째 질문은 외적 타당도(external validity)로 연구결과를 일반화시킬 수 있으면 외적으로 타당하다고 말할 수 있다(이두영 외, 1997).

실험참가자는 크게 두 가지로 구분되었다. 첫째, 인간을 실험에 참가시킨 경우이다. 실험에 참가한 인간이 그 실험에 직접적인 영향을 미치고 경우이다. 둘째, 탐색질의를 실험에 사용하고 인간은 단순히 탐색만 행하여, 실험에 어떠한 영향을 미치지 않는 경우로 구분하였다.

### 2.2.3 정보검색 실험연구에 사용된 변인

정보검색 연구에 사용된 변인들은 수없이 많다. 하지만 본 연구에서는 변인들에 대한 정의와 측정에 관한 것으로 한정하였다. 변인들은 인간과 비인간 변인들로 크게 구분하였다. 그리고, 인간과 비인간 변인 각각을 독립변인과 종속변인으로 구분하였다. 독립변인은 종속변인에 영향을 미치는 변인으로 설명된다. 종속변인의 목적은 독립변인이 영향을 미치는가를 조사하는 것이다. 그래서, 종속변인의 측정은 신뢰성이 있어야 하며, 실행되는 경우 타당성을 가지고 있어야 한다. 즉, 측정하고자 하는 것이 측정되어야 하며, 그 측정이 정확하게 이루어졌는가 하는 것이다.

## 3. 정보검색 실험연구의 분석

### 3.1 사용된 시스템

정보검색을 연구하기 위해서는 시스템 선택이 매우 중요하다. 왜냐하면 어떤 시스템을 선택하였는가에 따라 원하는 연구결과가 달라질 수 있기 때문이다. 연구할 시스템을 선택하기 위해서

는 다음의 세 가지 질문이 필요하다. 첫째, 무엇이 더 일반적인 연구의 대상인가? 둘째, 왜 그 대상을 선택하였는가? 셋째, 선택된 시스템을 통하여 어떤 연구질문이 나올 수 있으며, 그리고 그 답이 발견될 수 있는가?

무엇이 가장 일반적인 시스템인가에 대한 답을 살펴보면, 색인과 온라인 탐색을 연구한 연구자들 대부분은, 일반적으로 사용되는 시스템을 사용하였다. 예를 들면, Medline, ERIC 등이 주로 사용되었다. OPAC에 관한 연구는 실제 사용하는 도서관에서 실행되었다.

왜 그 대상을 선택하였는가에 대한 답은 그 시스템이 연구자의 직장이나 혹은 학교에 설치되어

있어서, 연구자가 자유롭게 사용할 수 있는 여건이 마련되었기 때문이라 생각할 수 있다. 이는 정보검색 연구의 특징으로 설명할 수 있다. 그중 한 가지는 정보검색 연구는 실용성이 있어야 한다. 도서관 사서들은 자관의 이용자가 정보를 획득하고 만족을 얻는데 보람을 느끼지, 어떤 정교한 실험시설을 만들어서 실험을 하여 얻는 결과를 가지고 보람을 느끼지 않는다. 그리고, 공통성은 신뢰도를 향상시킨다는 것이다. 한 도서관의 OPAC 사용법을 배운 사람은 다른 도서관의 OPAC을 사용하는데 적용할 수 있다. 또한, Medline이나 혹은 ERIC을 연구한 연구자들은 그 데이터베이

〈표 1〉 실험연구에 사용된 시스템의 종류

사용된 시스템			
색 인	OPAC	온라인 탐색	기 타
Bartell,	곽철완(1996)	김선호(1992), 이명희(1994)	Allen(1989, 1991)
Cotterell,	Blackshaw &	이관주(1994), 이상복(1994)	-Online
&Belew(1995)	Fischhoff(1982)	Boyce & McLain(1989), Brooks(1995)	service
Chan(1989)	Borgman(1986)	Coll, Coll & Nandavar(1993)	Dimitroff &
Chen, et	Byrne & Micco(1988)	Fenichel(1981), Fidel(1985)	Wolfram(1995)
al.(1995)	Connell(1991)	Howard(1982), Hsieh-Yee(1993)	-HyperLynx
Kessler(1965)	Dalrymple(1990)	Katzeff(1988),	Khan &
Knutson(1991)	Henty(1986)	Katzer et al.(1982), Kim(1997)	Locatis(1998)
Markey(1984)	Hunter(1991)	Kirby & Miller(1986)	-WWW
Rajashekar &	Kaske(1988),	Marcus(1983)	Pajmans(1993)
Croft(1995)	Kwak(1995)	Marcus & Reintjes(1981)	-IR system
Schabas(1982)	Larson(1991)	Markey, Atherton & Newton(1980)	Qui(1993)
Sievert &	Lipetz & Paulson(1987)	McKinin et al.(1991), Oh(1998)	-Hypertext
Andrew(1991)	Matthew(1984)	Sewell & Teitelbaum(1987)	
	Pease & Gouke(1982)	Stielow & Tibbo(1988)	
	Van Pulis &	Tenopir(1985), Wilbur(1998)	
	Ludy(1988)	Yoon(1994)	
	Voorbij(1998)	Young & Shneiderman(1993)	

스에 익숙하기 때문에 연구를 하겠지만, 동시에 많은 도서관 사서들이 Medline이 중요하기 때문에 많이 이용하여 익숙하다는 이유도 있다(표 1) 참조).

### 3.2 연구분석의 대상(units of analysis)

#### 3.2.1 종류

##### 검색 결과물

연구분석의 대상으로 검색 결과물이 이용되는 경우는, 분석을 위해 각각의 결과물에 대한 적합성(relevance)이 측정되었다. 적합성은 탐색질의와 비교하여 재현율과 정도율이 계산되었다. 특히, 색인연구에서는 검색결과물이 연구대상으로 자주 사용되었다. Bartell, Cottrell, and Belew(1995)는 각각의 검색결과물을 가지고 재현율과 정도율을 측정하여 3가지 다른 색인방법을 연구하였다. 또한, 색인의 일관성 연구에서는 검색결과물의 짝이 연구대상으로 사용되었는데, 일치되는 정도가 측정되었다(예, Markey, 1984; Chan, 1989).

##### 탐색질의

각각의 탐색질의를 서로 다른 탐색방법으로 사용하였을 때 결과물에 대한 재현율, 정도율, 비용, 사용된 명령어 등이 연구되었다. 이 경우, 실험에 참가한 사람들이, 결과물에 대한 평가는 하지만, 실험참가자가 연구대상이 되는 경우는 아니다. 예를 들면, Dalrymple(1990)은 두 가지 목록에 대한 재탐색에 관하여 연구를 하였는데, 데이터는 실험참가자들에 대한 관찰과 생각을 큰 소리로 말하여 기록한 방법으로 수집되었다. 하지만, 연구목적은 각 탐색질의에 대한 재탐색이 얼마나 발생하는 가를 조사하는 것이기 때문에,

연구대상은 탐색질의가 된다.

##### 인간

인간을 연구분석의 대상으로 삼은 연구는 Allen(1991)에 의한 것이 있다. 이 연구는 탐색질의를 이용하여 재현율을 조사하였다. 하지만, 독립변수는 "질문주제에 대한 지식의 차이"였으며, 분석된 결과는 주제에 관한 높은 지식을 가진 사람은 낮은 재현율을 나타낸다고 하였다. 또한, 이용자의 검색 결과에 대한 만족도나, 탐색에 대한 개개인의 반응 등이 관찰되는 경우가 있다. 이 경우, 표현되는 형식은 '검색 결과에 대해 자연어를 이용하여 탐색을 한 이용자는 통제어휘를 이용하여 탐색을 한 이용자보다 더 많은 만족을 얻는다.'라고 나타난다. 하지만, 결과가 '통제어휘 탐색이 더 높은 만족도를 제공한다'라는 표현된다면 연구분석의 대상은 탐색질의가 된다(예, Sullivan, Borgman & Wippem, 1990).

탐색질의와 인간에 대한 연구의 구별은 쉽지가 않다. 특히, 한 사람에 대하여, 한 탐색질의가 사용된 연구에서는 구분하는 의미가 없다. 이 경우, 각각을 연구분석의 대상으로 해석할 수 있기 때문이며, 연구자는 바꾸어 설명할 수도 있다.

##### 기타

몇몇 연구의 결과들은 불명확한 연구분석의 대상이나 사람이 포함되었기 때문에 명백하지 못하는 경우가 있다. 예를 들면, 색인의 일관성 연구에서 색인 작성자에 대한 내용을 포함시키지 않고, 두 번 색인된 레코드만 조사한 연구가 있다(예, Sievert & Andrews, 1991). 다른 연구분석의 대상으로 검색화면의 인터페이스가 사용된 경우도 있었다(예, Coll, Coll & Nandavar, 1993; Young & Shneiderman, 1993). 또한,

〈표 2〉 연구분석의 대상 종류

연구분석의 대상			
검색 결과물	탐색질의	인 간	기 타
김선호(1992), 이란주(1994), 이명희(1994) Bartell, Cottrell & Belew(1995) Chan(1989) Chen, et al.(1995) Markey(1984) Oh(1998) Pajmans(1993) Qui(1993) Rajashekar & Croft(1995) Voorbij(1998)	Boyce & McLain(1989) Dalrymple(1990) Fenichel(1981) Kirby & Miller(1986) Markey, Atherton & Newton(1980) Sewell & Teitelbaum(1987) Tenopir(1985) Yoon(1994)	Allen(1989, 1991) Blackshaw & Fischhoff(1982) Borgman(1986) Dimitroff & Wolfram(1995) Fidel(1986) Howard(1982) Hsieh-Yee(1993) Katzeff(1988) Lipetz & Paulson(1987) Marcus(1983) Marcus & Reintjes(1981) Matthew(1984) McKinin et al.(1991) Pease & Gouke(1982) Stielow & Tibbo(1988) Van Pulis & Ludy(1998) Wilbur(1998)	곽철완(1996)-OPAC 탐색절차 이상복(1990)-설문지 Byrne & Micco(1988) Brooks(1995)-서지레코드 Coll, Coll & Nandavar(1993) - 화면 Connell(1991) Henty(1986)-OPAC 탐색오류 Hunter(1991)-OPAC 탐색절차 Kaske(1988)-OPAC 탐색절차 Katzner, et al.(1982)-문헌 Kessler(1965)-문헌 Khan & Locatis(1998)-WWW Kim(1997)-조직방법 Knutson(1991) Kwak(1995)-OPAC 탐색절차 Larson(1991)-OPAC 탐색절차 Nye(1988)-OPAC 시스템 소프트웨어 Schabas(1982)-문헌 Sievert & Andrews(1991) -색인된 레코드 Young & Shneiderman(1993) -인터페이스

Kwak(1995)은 OPAC 탐색절차를 연구분석의 대상으로 이용하였다. OPAC을 이용할 때, 탐색질의와 사람에 따라 탐색행태가 어떠한 차이가 나는가 조사하였다(〈표 2〉 참조).

### 3.2.2 실험참가자 선정(recruitment)

#### 인간

대부분의 연구는 연구자가 실험을 위하여 인간을 인위적으로 선정하였다. 선정된 실험참가자들

은 학생, 정보탐색 업무 담당자 등으로 구분된다. 특히, 많은 연구자들이 대학생을 실험참가자로 사용하였는데(예, Coll, Coll & Nandavar, 1993; Kim, 1997; Kwak, 1995; Young & Shneiderman, 1993), 이는 쉽게 구할 수 있었기 때문이라 생각된다.

#### 탐색질의

탐색질의를 이용한 연구는 탐색질의를 만든 방

〈표 3〉 실험참가자의 종류

실험참가자			
인 간	탐색질의		기 타
	연구자에 의해	실험참가자에 의해	
광철완(1996)-대학생 김선호(1992)-탐색 업무 담당자 유재욱(1990)-대학생 이상복(1994), Allen(1989)-대학원생 Allen(1991)-대학생 Blackshaw & Fishhoff(1982)-공공도서관이용자 Borgman(1986)-대학생 Brooks(1995)-문헌정보학과 학생 Chen, et al. (1995)-전문가, 초심자 등 Coll, Coll & Nandavar(1993)-대학생 Dalrymple(1990)-대학, 대학원생 Dimitroff & Wolfram(1995) -문헌정보학과 학생 Fidel(1985)-탐색업무 담당자 Howard(1982)-탐색업무 담당자 Hsieh-Yee(1993)-대학원생 Katzeff(1988)-대학생 Kirby & Miller(1986)-동료학자 Khan & Locatis(1998)-고등학생 Kim(1997)-대학생, Kwak(1995)-대학생 Lipetz & Paulson(1987)-이용자 Marcus(1983)-박사과정학생 Marcus & Reintjes(1981)-연구자 Markey(1984)-목록작성자 Matthew(1984) Oh(1998)-대학원생 Pease & Goulke(1982) Qui(1993)-대학생 Sewell & Teitelbaum(1987)-유경험이용자 Stielow & Tibbo(1988)-교수 Van Pulis & Ludy(1989) Young & Shneiderman(1993)	이란주(1994), 이명희(1994) Bartel, Cottrell & Belew(1995) Boyce & McLain(1989) Croft(1995) Fenichel(1981) Sewell & Pajmans(1993) Rajashekar & Teitelbaum(1987) Yoon(1994)	Katzer, et al.(1982) Markey, Atherton & Newton(1980) McCain, White & Griffith(1987) McKinin et al.(1991) Schabas(1982) Tenopir(1985) Voorbij(1998)	Byrne & Micco(1988)-교수가 제공한 어휘 Chan(1989)-OCLC 레코드 Henty(1986)-OPAC transaction Hunter(1991)-OPAC transaction Kaske(1988)-OPAC transaction Knutson(1991)-문헌 Larson(1991)-OPAC transaction Nye(1988)-vendors Sievert & Andrew(1991)-CA 서지사항

법에 따라 구분되었다. 구체적으로 살펴보면, 연구자가 인위적으로 만든 경우와 실험참가자 스스로 탐색질의를 만든 경우로 구분되었다. 특히, 연구자가 인위적으로 만든 탐색질의는 선행의 정보 검색 실험에서 사용되었던 것을 사용하는 경우(예, Bartell, Cottrell & Belew, 1995)와 연구자가 ERIC이나 JASIS 같은 데이터베이스에서 선정한 경우(예, Brooks, 1995; Yoon, 1994), 그리고 연구자 스스로 만든 경우(예, 이명희, 1994; 이란주, 1994; Rajashekar & Croft, 1995; Pajmans, 1993)로 구분할 수 있다.

기타

그 밖의 범주에 속하는 것으로는, OPAC의 트랜잭션(예, Kaske, 1988; Larson, 1991)이나, 서지정보(예, Chan, 1989; Andrew, 1991) 등이 있다(〈표 3〉 참조).

3.3 변인

3.3.1 인간변인

독립변인

인간을 독립변인으로 사용한 연구들은 크게 3가지 범주로 나눌 수 있다. 첫째, 인간의 개인적 특성에 대한 범주이다. 이 범주에는 성별과 나이가 포함되며, 학력이나 직업 혹은 몇몇 안되지만 인간성도 포함되기도 한다. 둘째, 경험과 관련된 사항이다. 그리고 셋째는 운영변인이 포함된다(〈표 4〉 참조).

가. 인간의 개인적 특성

연령이나 성별 등은 정의 내리기가 매우 힘들다. 성별의 예를 들면, 어떤 연구에서는 두 가지로 구분하는 대신, 3가지로 구분하여, “표시되지 않음”으로 나타내기도 한다(예, Van Pulis & Ludy, 1988). 연령에 대해서는 ratio-level 변인이 사용되나, 만약 범주를 사용한다면, 연구에 다른 영향을 미친다. 예를 들면 대학생의 경우, 오늘날 다양한 연령층으로 구성되어 있기 때문에 특정한 연령층으로 한정하기 매우 힘들며, 일반적인 대학생 연령과 틀린 다른 연령층의 학생들이 커다란 영향을 미치고 있다고 볼 수 있다.

〈표 4〉 인간 독립변인의 종류

인간 독립변인		
인간의 개인적 특성	경험	운영변인
이상복(1994) Lipetz & Paulson(1987) Matthews(1984), Qui(1993) Sewell & Teitelbaum(1987) Stielow & Tibbo(1988) Van Pulis & Ludy(1988)	곽철완(1996), 김선호(1992) Allen(1991) Dimitroff & Wolfram(1995) Howard(1982), Kwak(1995) Matthew(1984), Qui(1993) Sewell & Teitelbaum(1987) Van Pulis & Ludy(1988) Wilbur(1998)	유재욱(1990), 이명희(1994) Allen(1989) Blackshaw & Fischhoff(1982) Borgman(1986) Hsieh-Yee(1993) Katzeff(1988)



몇몇 연구들은 인간을 교수, 대학원생 등으로 구분하거나(예, Sewell & Teitelbaum, 1987), 사서와 이용자(예, Lipetz & Paulson, 1987), 혹은 전공 영역별로 구분하였다(예, Stielow & Tibbo, 1988).

나. 경험

경험은 도서관이나 정보검색 경험에 관련된 변인으로, 직접적으로 그 연구에 관련이 있거나, 선행연구와 관련이 있다. 경험에 관련된 변인을 사용한 연구는 목록이나 온라인 탐색의 경험이었다(예, 광철완, 1996; Dimitroff & Wolfram, 1995; Kwak, 1995; Matthew, 1984; Qui, 1993; Van Pulis & Ludy, 1988). 또 다른 경험은 정보검색에 관련된 교육을 받은 경험이 있는가 하는 내용이다(예, Sewell & Teitelbaum, 1987; Van Pulis & Ludy, 1988). 이러한 변인 대부분은 실험참가자 자신이 이야기하는 내용으로 정확성에 대해서는 문제점을 가지고 있다. 여기에는 개개인의 기억력, 사회적인 과시(내가 OPAC을 잘 사용한다고 하면 남들이 얼마나 감탄 할 것인가?), 순종(연구자의 의도에 따름), 묵

인(무조건 “예”라 대답하는 경향) 등이 있다. 이러한 문제점들은 부차적인 측정으로 해결될 수 있다. 예를 들면, 시험을 통하거나(예, Allen, 1991), 관찰을 통한 방법이 있다(예, Sewell & Teitelbaum, 1987).

다. 운영변인

운영(manipulated) 독립변인은 실험연구에서 매우 중요한 역할을 한다. 이는 교육, 작업, 기타의 3가지로 구분되었다. 교육에 관련된 연구는 측정 이전에 주어진 교육의 형태나 시간에 따라 다양하다. 그리고 측정은 주어진 작업 내에서 이루어진다. 실험참가자는 동일한 작업을 수행하지만, 주어진 교육의 정도에 따라 구별된다(예, Borgman, 1986). 작업 처리에 있어서는 교육은 동일하지만, 주어지는 작업이 틀리는 경우이다(예, Allen, 1989). 그리고, 기타의 범주에는 다양한 정보봉사 상황에서 이용자의 반응을 조사하였는데 데이터의 해석은 여러 가지 결합된 분석으로 이루어졌다. 인간을 이용한 종속 변인은 크게 내적 반응, 활동 선택, 탐색 과정, 탐색 결과로 구분되었다(〈표 5〉 참조).

〈표 5〉 인간 종속변인의 종류

인간 종속변인			
내적반응	활동의 선택	탐색과정	탐색결과
Allen(1989) Blackshaw & Fischhoff(1982) Borgman(1986), Chen, et al.(1995) Dalrymple(1990) Dimitroff & Wolfram(1995) Lipetz & Paulson(1987) Stielow & Tibbo(1988)	이상복(1994) Matthew(1984)	광철완(1996), 유재욱(1990) Allen(1989, 1991) Dalrymple(1990) Hsieh-Yee(1993), Katzeff(1988) Kwak(1995) Sewell & Teitelbaum, (1987) Van Pulis & Ludy, (1988)	Blackshaw & Fischhoff(1982) Borgman(1986) Howard(1982) Katzeff(1988) Lipetz & Paulson(1987)

가. 내적 반응

내적 반응 변인은 만족도, 신뢰도(confidence), 그리고 감정적이며 인지적인 중요성 등이 있다(예, Lipetz & Paulson, 1987; Blackshaw & Fischhoff, 1982; Allen, 1989; Stielow & Tibbo, 1988).

나. 활동의 선택

활동의 선택 변인은 시작에 대한 내용이다(예, Matthew, 1984). 예를 들면, 이상복(1994)은 정보탐색을 위한 시스템 선택을 독립변인으로 사용하였다.

다. 탐색 과정

이는 사용된 어휘(예, Allen, 1989; 1991), 접근점(예, 광철완, 1996; Kwak, 1995; Sewell & Teitelbaum, 1987; Van Pulis & Ludy, 1988), 과정에서 정보의 변화(예, Katzeff, 1988; Dalrymple, 1990), 탐색전략 등이 포함된다.

라. 탐색 결과

탐색 결과의 변인은 정확성(예, Borgman, 1986; Blackshaw & Fischhoff, 1982; Katzeff, 1988)과 효율성(예, Howard, 1982) 등이 있는데 이는 비인간 변인의 결과변인과 유사하다.

3.3.2 비인간 변인

독립변인

인간 변인에서와 마찬가지로 비인간 변인에 있어서도 독립변인은 모두가 동일하게 사용되지는 않는다. 어떤 변인들은 기존에 존재해 있는 것이며, 또 다른 것은 선정된 것이다. 이러한 구별은 다소 명백하지는 않는다. 기존에 존재하는 독립변인을 사용하는 방법은 연구자가 현존하는 상황을 연구 목적에 적합하도록 이용하는 경우이다. 예를 들면, 모집단이나 표본이 파악되면, 이를 연구 목적에 맞도록 하부범주로 나누어 독립변인으로 사용하는 방법이다. 이와는 반대로, 선정된 방

〈표 6〉 비인간 독립변인의 종류

비인간 독립변인			
시스템의 종류	탐색방법	레코드의 특징	기 타
Bartell, Cottrell & Belew(1995) Chen et al.(1995), Connell(1991) Coll, Coll & Nandavar (1993) Darlymple(1990), Kim(1997) Kwak(1995), Larson(1991) Lipetz & Paulson(1987) Oh(1998), Pajmans(1993) Pease & Goulke(1982) Rajashekar & Croft(1995)	이란주(1994) Lipetz & Paulson(1987)	Brooks(1995) Chan(1989) Knutson(1986) Larson(1991) Sievert & Andrews (1988) Voorbij, 1998	Byrne & Micco(1988) Khan & Locatis(1998) Kaske(1988), Kessler(1965) Knutson(1986) McKinin et al.(1991) Nye(1988), Yoon(1994) Young & Shneiderman(1993)

법에 의한 변인은 연구자가 오직 몇 가지 상황에 만 적합하도록 선정하여 실험참가자를 구별하는 방법이다. 일반적으로 사용한 독립변인은 네 종류로 구별할 수 있다: 시스템 종류, 탐색 방법, 레코드 특징, 기타(〈표 6〉 참조).

#### 가. 시스템의 종류

이는 온라인 목록 혹은 카드 목록의 구별(예, Pease & Goulke, 1982; Connell, 1991)과, 실험용 시스템과 현재 사용하고 있는 시스템(예, Larson, 1991), 메뉴방식 혹은 키워드 방식 등(예, Coll, Coll & Nandavar, 1993; Kim, 1997)으로 구분된다.

#### 나. 탐색 방법

일반적으로 탐색 방법은 운영적(manipulated)인 독립변인인데, 연구자에 의해서 만들어 진 것이 아니라, 실험참가자가 행하는 내용이기 때문이다. 또한 연구에서 사용된 방법 역시, 모든 다양한 탐색 방법의 가능성을 배제하고 있지 않다. 예를 들면, 이란주(1994)는 인용문헌의 사용 수를 독립변인으로 이용하였다.

#### 다. 레코드의 특징

이는 색인의 깊이(예, Sievert & Andrews, 1988), 목록에 나열된 주제명의 수(예, Knutson, 1986), 본문의 특징(예, Brooks, 1995), 주제의 종류(예, Voorbij, 1998) 등이 여기에 포함된다.

#### 라. 기타

기타 변인으로 Nye(1988)는 가능한한 존재하는 모든 소프트웨어의 정보원을 수집하여 사이트 별 혹은 상업적으로 개발된 업체별로 크게 두 가

지 범주로 구별하였다. 여기에서, 모든 정보원을 포함하려 시도하였기에 "선정"된 변인이 아니라 "존재"하는 변인에 속한다.

선정된 변인을 사용한 연구들에 있어서는, 연구자가 관심 있는 특정한 변인을 사용하였다. 그러므로, 연구에 사용된 변인은 실제 존재하는 다양한 변인보다 크게 한정되었다. 이러한 선정된 변인을 사용하는 연구에서 발생하는 질문은 왜 그 대상이 선정되었는가, 그리고 왜 다른 대상은 선정되지 않았는가 고려해야 한다. 몇몇 연구들은 소장장서 중에서 변인으로 선정하였고(예, Kaske, 1988), 몇몇은 실험참가자 중에서(예, Byrne & Micco, 1988), 그리고 몇몇은 처리과정에서 선정하였다(예, Kessler, 1965).

#### 중간 변인

중간 변인은 종속변인을 설명하는 역할을 담당하는 변인으로 독립변인 혹은 종속변인에 속하기도 한다. 비인간 변인을 사용한 연구에서 가장 많이 사용된 중간변인은 적합성(relevance) 변인이다. 검색된 서지정보를 가지고 탐색질의와 관련 있는지 혹은 관련 없는지 결정하는 것은 탐색질의를 통한 종속변인을 연구하기 위해서 필수적인 것이다. 이 주제에 대한 연구는 Cuadra and Katter(1967)의 적합성 판단의 경험적 연구가 있다. 그리고, 몇몇 이론적인 비평연구도 있다(예, Saracevic, 1975; Harter, 1992).

적합성에 대한 가장 명백하고 일반적인 접근은 이분법과 이용자 판단을 통한 방법이다. 대부분의 연구에서는 결과에 대해서 단순히 예/아니오를 사용하여 판단하였다. 하지만, 몇몇 연구들은 다양한 스케일을 이용하여 결과를 비교하였다(예, Katzer et al., 1982; Schabas, 1982). 몇몇 연구들은 전문가에 의해서 판단되었는데, 이

경우는 대부분 데이터베이스 연구의 경우였다 (예, Markey, Atherton & Newton, 1980).

**종속변인**

연구자가 독립변인의 중요성을 결정하는데 사용되는 종속변인은 크게 두 가지 범주로 구분된다. 첫째로, 탐색과정 변인은 어떻게 탐색 혹은 검색이 실행되는가 하는 문제를 조사하는데 사용된다. 하지만 이 변인은 때때로 독립변인으로 사용되는데, 다음에 설명되는 탐색결과 변인에 영향을 미치는 여부를 파악하기 위해 이용되기도 한다. 둘째, 탐색결과 변인은 검색과정의 결과에 대한 조사를 위해 사용된다(〈표 7〉 참조).

**가. 탐색과정**

탐색과정 변인은 '어떤 일이 수행되었는가?(예, 사용된 탐색 종류)', '얼마나 많이 사용되었는가?(다양한 요소들의 양)', 그리고 중복되는 특수한 범주에 속하는 변인들이 있는데, 이는 각 상황에서 '어떤 일이 수행되었는가'와 '얼마

나 많이 사용되었는가'를 비교하는 변인으로 생각할 수 있다. '어떤 일이 수행되었는가'는 독립변인으로 생각될 수 있으나, 여기에서는 독립변인의 결과물로 구분하였다. 그러므로, 주제 탐색의 양은 연도에 따라 변화 하든가(예, Kaske, 1988), 혹은 사용하는 목록의 종류가 변화하는 연구(예, Lipetz & Paulson, 1987)가 여기에 속한다. '얼마나 많이 사용되었는가'는 개념에 대해(예, Kirby & Miller, 1986; Dalrymple, 1990), 용어에 대해(예, Fidel, 1985), 시간에 대해(예, Marcus & Reintjes, 1981; Oh, 1998), 그리고 오류에 대해(예, Henty, 1986) 조사되었다. 이 변인은 트랜잭션 로그를 분석하여 데이터를 수집하기도 하였다(예, Qui, 1993).

**나. 탐색결과**

탐색결과는 크게 검색 결과의 양, 시간, 정도를, 중복율, 재현율 등으로 구분된다. 검색 결과의 양은 직접적으로 보여주지 때문에 구별하는데 있어서 이점을 가지고 있으며, 적합성 질문에 대

〈표 7〉 비인간 종속변인의 종류

비인간 종속변인	
탐색과정	탐색결과
Chan(1989)	김선호(1992), 이란주(1994), 이명희(1994), Allen(1989, 1991)
Dalrymple(1990)	Bartell, Cottrell & Belew(1995), Boyce & McLain(1989), Brooks(1995)
Fidel(1985), Henty(1986)	Chen, et al.(1995), Coll, Coll & Nandavar(1993), Dalrymple(1990)
Kaske(1988), Kim(1997)	Dimitroff & Wolfram(1995), Fenichel(1981), Howard(1982), Hunter(1991)
Kirby & Miller(1986)	Katzer et al.(1982), Khan & Locatis(1998), Kim(1997), Kirby & Miller(1986), Larson(1991), McKinin et al.(1991), Marcus(1983)
Lipetz & Paulson (1987)	Marcus & Reintjes(1981), Oh(1998), Pajmans(1993), Rajashekar & Croft(1995), Schabas(1982), Tenopir(1985), Van Pulis & Ludy(1988)
Marcus(1983)	Marcus & Reintjes(1981), Oh(1998), Pajmans(1993), Rajashekar & Croft(1995), Schabas(1982), Tenopir(1985), Van Pulis & Ludy(1988)
Marcus & Reintjes(1981)	Marcus & Reintjes(1981), Oh(1998), Pajmans(1993), Rajashekar & Croft(1995), Schabas(1982), Tenopir(1985), Van Pulis & Ludy(1988)
Oh(1998), Qui(1993)	Voorbij(1998), Wilbur(1998), Yoon(1982), Young & Shneiderman(1993)

한 문제점을 해결하기 위해 효과적인 대안을 제공해준다. 이는 두 가지 방법으로 얻어지는데, 첫째는 적합성이 포함되지 않고, 검색된 양만을 구분하는 경우이다. 예를 들면, Marcus & Reintjes(1981)는 인쇄된 페이지 수만 분석하였다. 둘째, 적합성 판정의 문제를 피하기 위해 탐색을 성공했는가 하지 못했는가의 이분법을 적용하였다. 이는 실험참가자가 스스로 탐색에 성공했다고 이야기하는 경우(예, Kirby & Miller, 1986; Van Pulis & Ludy, 1988)와 전문가가 판단(예, Van Pulis & Ludy, 1988)하거나 혹은 기계적으로 시스템이 결과물을 보여주는지 여부(예, Hunter, 1991)로 나눌 수 있다.

시간은 탐색결과물이 얻는데 얼마나 소요되었는가 하는 문제로서, Dimitroff & Wolfram(1995), Coll, Coll & Nandavar(1993) 등이 연구에 사용하였다. 하지만, 시간 측정은 환경에 많은 영향을 받으므로, 다른 환경을 정확하게 통제하여야 하므로 사용하기가 쉽지 않다.

정도율은 검색된 총 문헌의 수와 검색된 문헌 중 관련된 문헌의 수와의 비율이다. 이는 적합성과 직접적으로 연관이 되는 것으로 많은 연구에서 사용되었다. 정도를 측정함에 있어서 문제는 어떤 방법으로 계산할 것인가에 있다. 정도를 측정을 위해 여러 번의 탐색이 이루어졌을 때(예, 디스크립터 탐색과 원문탐색 각각 11회; McCain, White & Griffith, 1987), 결과에 대한 통계를 각각의 평균으로 할 것인가 아니면, 전체 평균으로 할 것인가 하는 문제가 있다. 각각의 평균치를 사용하였을 때, 정도율은 각각 탐색의 정도율이 서로 합해진 후 총 탐색횟수로 나눈 값이다. 하지만, 전체 평균을 사용하는 방법은 검색된 모든 결과물을 함께 모아, 적합성에 따라 정도율을 측정하는 것이다. 일반적으로 전체 평균이 정도율 계

산에 사용(예, Katzer et al., 1982; Schabas, 1982)되었지만, 몇몇은 각각 평균치가 사용(예, Tenopir, 1985)되기도 하였다. 하지만, 대부분은 계산방법을 기술하지 않았다.

정도율은 탐색방법의 효과성을 측정하지만, 동시에 탐색결과물의 중복 정도에 대해서 필수적으로 비교되어야 한다. 몇몇 연구들은 탐색방법을 비교하는데 중복율을 사용하였다(예, 이란주, 1994; 이명희, 1994; Yoon, 1994).

재현율은 측정하기 어려운 변인중의 하나이다. 이론적으로는 잘 정의되어 있으나, 실제로는 절대 재현율과 상대 재현율로 구분된다. 절대 재현율을 계산하기 위해서는 탐색질의와 연관이 있는 모든 자료의 적합성을 조사하여야 한다. 이 방법은 Cranfield 연구에 사용되었고, ONTAP ERIC 데이터베이스 연구에서도 사용되었다(예, Fenichel, 1981; Boyce & McLain, 1989).

다른 연구자들은 모든 자료를 다 조사하지 않고, 대략적인 추정을 통하여 사용하였다. 예를 들면, Marcus의 두 번의 연구(Marcus & Reintjes, 1981; Marcus 1983)에서 “대략적인 재현율 기본자료(estimated recall base)”를 설정하고, 연구자가 반복적인 탐색을 통하여 이를 결정하였다. 비슷한 예로, Allen(1991)은 두 명의 훈련받은 사서가 탐색한 결과를 기준으로 이용자의 결과물을 비교하였다.

#### 4. 결 론

본 연구는 정보검색분야에서 사용되는 실험연구 방법의 다양성을 체계적으로 분석하였다. 실험연구 방법을 통하여 정보검색을 연구할 때, 선행 연구들은 색인, OPAC, 온라인탐색(데이터베

이스) 등의 시스템을 이용하여, 검색결과물, 탐색질의, 인간 등을 연구 대상으로 삼아 분석하였다. 그리고, 실험을 위하여 인간 혹은 탐색질의를 선정하였고, 변인으로는 인간과 비인간들을 이용하였다.

추후에 정보검색을 연구하려는 연구자들에게 효과적인 도움을 주기 위해서는, 연구 방법에 관련된 다양한 정보들이 포함된 데이터베이스 구축이 필요하다. 오삼균(1998)은 정보검색 효과에

대한 연구에서 변인들의 관계를 보여주는 정보검색시스템이 일반적인 정보검색시스템에 비해 더 효과적이라는 주장을 하였다. 이러한 주장은 본 연구가 실험연구 방법을 사용한 정보검색분야 데이터베이스 구축을 위한 구조 모델이 될 수 있으며, 효과적인 정보검색시스템 구축을 위한 기초연구가 될 수 있다. 그러므로, 다른 연구방법에 대한 체계적인 분석과 분석된 내용의 상호적합성에 대한 후속연구가 필요하다.

## 참 고 문 헌

- 곽철완. 1996. "온라인 목록의 탐색 유형: 관련된 변수와 사용된 서지사항을 중심으로", 한국문헌정보학회지, 30(2): 87-106.
- 김선호. 1992. "온라인 정보탐색의 효과변인 분석", 한국문헌정보학회지, 22: 361-396.
- 오삼균. 1998. "실험변수를 이용한 정보검색 시스템의 구축 및 평가", 한국문헌정보학회 학술발표논문집, 6: 41-53.
- 이두영. 1997. 문헌정보학 연구방법론. 서울: 한국도서관협회.
- 이관주. 1994. "인용문헌에 의한 정보검색 효과에 관한 고찰", 한국문헌정보학회지, 26: 265-289.
- 이명희. 1994. "수자원문헌의 주제탐색과 인용탐색의 검색효율 비교 연구", 한국문헌정보학회지, 26: 213-233.
- 이상복. 1994. "최종이용탐색시스템의 이용과 이용자의 인지적 특성간의 관계 연구", 한국문헌정보학회지, 26: 291-339.
- Allen, B. 1989. "Questions on search request forms". *Online Review*, 13: 283-289.
- Allen, B. 1991. "Topic knowledge and online catalog search formulation", *Library quarterly*, 61: 188-213.
- Bartell, B., Cottrell, G., and Belew, R. 1995. "Representing documents using an explicit model of their similarities", *Journal of the American Society for Information Science*, 46(4): 254-271.
- Blackshaw, L. and Fishhoff, B. 1982. "Decision making in online searching", *Journal of the American Society for Information Science*, 33: 369-389.
- Borgman, C. 1986. "The user's mental model of an information retrieval system: An experiment on a prototype online catalog", *International Journal of Man-Machine Studies*, 24: 47-64.
- Boyce, B. and McLain, J. 1989. "Entry point depth and online search using a controlled vocabulary", *Journal of the American Society for Information*

- Science, 40: 273-276.
- Brooks, T. 1995. "People, words, and perceptions: A phenomenological investigation of textuality", *Journal of the American Society for Information Science*, 46(2): 103-115.
- Byrne, A. and Micco, M. 1988. "Improving OPAC subject access: the ADFA experiment", *College & Research Libraries*, 49: 432-441.
- Campbell, D., and Stanley, J. 1969. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Chan, L. 1989. "Inter-indexer consistency in subject cataloging", *Information Technology and Libraries*, 8: 349-358.
- Chen, H., Yim, T., Fye, D., and Scharz, B. 1995. "Automatic thesaurus generation for an electronic community system", *Journal of the American Society for Information Science*, 46(3): 175-193.
- Coll, J., Coll, R., and Nandavar, R. 1993. "Attending to cognitive organization in the design of computer menus: A two-experiment study", *Journal of the American Society for Information Science*, 44(7): 393-397.
- Connel, T. 1991. "User acceptance of library catalog results: An exploratory study", *Library Resources and Technical Services*, 35: 191-201.
- Cuadra, C., and Katter, R. 1967. "Opening the black box of relevance", *Journal of Documentation*, 23: 291-303.
- Dalrymple, P. 1990. "Retrieval by reformulation in two library catalogs: Toward a cognitive model of search behavior", *Journal of the American Society for Information Science*, 41: 272-281.
- Dimitroff, A. and Wolfram, D. 1995. "Searcher response in a hypertext-based bibliographic information retrieval system", *Journal of the American Society for Information Science*, 46(1): 22-29.
- Fenichel, C. 1981. "Online searching: Measures that discriminate among users with different types of experience", *Journal of the American Society for Information Science*, 32: 23-32.
- Fidel, R. 1985. "Individual variability in online searching behavior", in *ASIS 85: Proceedings of the 48th ASIS Annual Meeting*, ed. Carol A. Parkhurst. White Plains, NY: Knowledge Industry Publications.
- Harter, S. 1992. "Psychological relevance and information science", *Journal of the American Society for Information Science*, 43: 602-615.
- Henty, M. 1986. "The user at the online catalogue: A record of unsuccessful keyword searches", *LASIE*, 17: 47-52.
- Howard, H. 1982. "Measures that discriminate among online searchers with different training and experience", *Online Review*, 6: 315-327.

- Hsieh-Yee, I. 1993. "Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers", *Journal of the American Society for Information Science*, 44(3): 161-174.
- Hunter, R. 1991. "Successes and failures of patrons searching the online catalog at a large academic library: a transaction log analysis", *RQ*, 30: 395-402.
- Kaske, N. 1988. "A comparative study of subject searching in an OPAC among branch libraries of a university library system", *Information Technology and Libraries*, 7: 359-372.
- Katzeff, C. 1988. "The effect of different conceptual models upon reasoning in a database query writing task", *International Journal of Man-Machine Studies*, 29: 37-62.
- Katzer, J., McGill, M., Frakes, W., and DasGupta, P. 1982. "A study of the overlap among document representations", *Information Technology, Research and Development*, 2: 261-274.
- Kessler, M. 1965. "Comparison of the results of bibliographic coupling and analytic subject indexing", *American Documentation*, 16: 223-233.
- Khan, K. and Locatis, C. 1998. "Searching through cyberspace: The effects of link display and link density on information retrieval from hypertext on the World Wide Web", *Journal of the American Society for Information Science*, 49(2): 176-182.
- Kim, Y. 1997. "A method for the classification and retrieval problem of reusable software resources", *Information Processing & Management*, 33(4): 513-522.
- Kirby, M and Miller, N. 1986. "Medline searching on colleague: Reasons for failure or success of untrained end users", *Medical Reference Services Quarterly*, 5: 17-34.
- Knutson, G. 1991. "Subject enhancement: Report on an experiment", *College and Research Libraries*, 52: 65-79.
- Krikelas, J. 1980-81. "Searching the library catalog—a study of user's access", *Library Research*, 2: 215-230.
- Kwak, C. 1995. Search patterns using online catalogs, Ph. D. thesis. University of Wisconsin-Madison.
- Larson, R. 1991. "The decline of subject searching: Long-term trends and patterns of index use in an online catalog", *Journal of the American Society for Information Science*, 42: 197-215.
- Lipetz, B. and Paulson, P. 1987. "A study of the impact of introducing an online subject catalog at the New York state library", *Library Trends*, 35: 597-617.
- Marcus, R. 1983. "An experimental comparison of the effectiveness of computers and



- humans as search intermediaries", *Journal of the American Society for Information Science*, 34: 381-404.
- Marcus, R. and Reintjes, J. 1981. "A translating computer interface for end-user operation of heterogeneous retrieval systems", *Journal of the American Society for Information Science*, 32: 304-317.
- Markey, K. 1984. "Inter-indexer consistency tests: A literature review and report of a test of consistency in indexing visual materials", *Library and Information Science Research*, 6: 155-177.
- Markey, K., Atherton, P., and Newton, C. 1980. "An analysis of controlled vocabulary and free text search statements in online searches", *Online Review*, 4: 225-236.
- Matthews, J. 1984. *Online public catalogs: A nationwide survey*.
- McCain, K., White, H., and Griffith, B. 1987. "Comparing retrieval performance in online data bases", *Information Processing and Management*, 23: 539-553.
- McKinin, E., Sievert, M., Johnson, E., and Mitchell, J. 1991. "The Medline/full-text research project", *Journal of the American Society for Information Science*, 42: 297-307.
- Nye, J. 1988. "User interaction with the authority structure of the online catalog: results of a survey", *Information Technologies and Libraries*, 7: 313-316.
- Oh, S. 1998. "Document representation and retrieval using empirical facts: Evaluation of a pilot system", *Journal of the American Society for Information Science*, 49(10): 920-931.
- Pajmans, H. 1993. "Comparing the document representations of two IR-systems: CLARIT and TOPIC", *Journal of the American Society for Information Science*, 44(7): 383-392.
- Pease, S. Gouke, M. 1982. "Patterns of use in an online catalog and a card catalog", *College and Research Libraries*, 43: 279-291.
- Rajashekar, T. and Croft, W. 1995. "Combining automatic and manual index representation in probabilistic retrieval", *Journal of the American Society for Information Science*, 46(4): 272-283.
- Qui, L. 1993. "Markov models of search state patterns in a hypertext information retrieval system", *Journal of the American Society for Information Science*, 44(7): 413-427.
- Saracevic, T. 1975. "Relevance: A review of and a framework for the thinking on the notion in information science", *Journal of the American Society for Information Science*, 26: 321-343.
- Schabas, A. 1982. "Post coordinate retrieval: A Comparison of two indexing languages", *Journal of the American*

- Society for Information Science, 33: 32-37.
- Sewell, W. and Teitelbaum, S. 1986. "Observations of end-user online searching behavior over eleven years", *Journal of the American Society for Information Science*, 37: 234-245.
- Sievert, M. and Andrews, M. 1991. "Indexing consistency in information science abstracts", *Journal of the American Society for Information Science*, 42: 1-6.
- Stielow, F. and Tibbo, H. 1988. "The negative search, online reference, and the humanities: A critical essay in library literature", *RQ*, 27: 358-365.
- Sullivan, M., Borgman, C., and Wippern, D. 1990. "End-users, mediated searches, and front-end assistance programs on DIALOG: A comparison of learning, performance and satisfaction", *Journal of the American Society for Information Science*, 41: 27-42.
- Tenopir, C. 1985. "Full text database retrieval performance", *Online Review*, 9: 149-164.
- Van Pulis, N. and Ludy, L. 1988. "Subject searching in an online catalog with authority control", *College & Research Libraries*, 49: 523-33.
- Voorbij, H. 1998. "Title keywords and subject descriptors: A comparison of subject search entries of books in the humanities and social science", *Journal of Documentation*, 54(4): 466-476.
- Wilbur, W. 1998. "A comparison of group and individual performance among subject experts and untrained workers at the document retrieval task", *Journal of the American Society for Information Science*, 49(6): 517-529.
- Yoon, L. 1994. "The performance of cited references as an approach to information retrieval", *Journal of the American Society for Information Science*, 45(5): 287-299.
- Young, D. and Shneiderman, B. 1993. "A graphical filter/flow representation of Boolean queries: A prototype implementation and evaluation", *Journal of the American Society for Information Science*, 44(6): 327-339.