

국어교육학 실증적 연구의 종합을 위한 메타분석 방법론*

조재윤**

〈 차 례 〉

- I. 서론
- II. 연구 종합 방법
- III. 메타분석 및 효과크기의 개념
- IV. 메타분석의 절차
- V. 결론

I. 서론

1. 메타분석 방법론의 필요성

이 글의 목적은 국어교육학 실증적 연구의 종합을 위한 메타분석 방법에 대하여 구체적으로 소개하는데 있다.

시대의 변화에 따른 학문의 발달과 그 복잡성으로 인하여 각 학문의 개개 연구물이 축적되어 그 연구물들에 대한 종합적인 연구가 필요한 상황 하에서 메타분석이라는 새로운 연구 방법이 생기게 되었다.

* 이 논문은 제29회 국어교육학회(2004. 11. 19. 서울대학교) 학술발표대회에서 발표한 것을 수정·보완한 것이다. 토론을 해 주신 전주교대 서혁 교수님께 감사드린다.

** 서울수암초등학교 교사

메타분석 방법은 다음과 같은 상황에서 유용하고 중요하게 쓰인다. 첫째, 단일 주제에 관해 두 개의 상반된 결론이나 논쟁이 야기되었을 때 이의 해결을 위한 보다 신뢰할 수 있고 타당성 있는 대결론(big decision)을 내려야 할 필요가 있을 때이다. 둘째, 현장으로부터의 원 자료(raw data)를 수집할 만한 시간적 여유가 없거나, 경비와 노동력의 절약이 요구되는 상황에서의 2차 자료를 이용한 문제 해결이 필요할 때이다. 셋째, 각종 학술정보의 범람 속에서 체계적으로 압축된 지식 또는 정보를 필요로 할 경우이다.

2. 메타분석 방법론의 국어교육적 의의

우리나라에서는 메타분석 방법을 황정규(1988) 등의 소개로 의학, 간 호학, 교육학 등에서 유용하게 사용하고 있으나, 국어교육학 분야에서는 그 연구의 예가 매우 드물다. 윤준채(2002)와 윤준채 외(2004)가 외국의 독서 연구를 바탕으로 한 연구가 있고, 조재운(2005)이 국내의 국어교육 학 실증적 연구인 '질문생성전략'을 변수로 하는 연구물을 대상으로 수행한 연구가 있다.

국어교육학의 실증적인 연구물의 수가 계속하여 늘어가는 추세인 만큼 연구 종합을 위한 메타분석 방법을 활용한 연구가 필요하다. 메타분석 방법론이 일반 교육학의 연구 방법론을 따르는 것이지만 국어교육학 분야에서 이 방법론을 사용한 연구가 충분히 이루어진다면 의학이나 간호학 등의 분야에서 매우 유익한 연구 방법이 된 것처럼 국어교육학의 새로운 연구 방법으로 정착되어 유용하게 쓰일 것이다.

실증적 연구 중에 메타분석 방법론을 적용할 수 있는 국어교육 연구 물에는 독해력 변수¹⁾와 관련된 연구, 정교화 학습 전략 변수와 관련된

1) 국립중앙도서관을 통한 논문 검색에서 '독해력'과 관련된 연구물은 300여 편이다. 이 연구물들을 엄밀히 분석한다면 메타분석 방법을 사용하기에 부적절한 것이 많을 것이지만 한 가지의 변수에 대한 연구물이 꽤 많이 축적되어 있는 것은 메타분석의 필요성을 보여주는 예라고 할 수 있다.

연구, 읽기 지도 방법 변수와 관련된 연구 등이 있다. 이 연구물들은 현재까지의 연구물만으로도 메타분석이 가능할 것이다.

II. 연구 종합 방법

1. 연구 문헌의 종합 방법

기존에 이루어진 개개의 연구 문헌을 종합하고 리뷰하는 방법은 이론적 리뷰, 방법론적 리뷰, 통합적 리뷰로 나눌 수 있다.(황정규, 1988:2, 박도순, 2002:253-254)

가. 이론적 리뷰(theoretical review)

리뷰하는 사람이 어떤 특정한 현상을 설명하기 위해 이미 발표된 이론들을 이론의 범위, 내적 일관성, 예언의 성질 등에 관해 서로 비교하기 위하여 사용하는 방법으로, 새로운 이론적 체계를 시사하거나 상이한 이론이 지닌 추상적 관점을 재구성하거나 통합하려고 한다.

나. 방법론적 리뷰(methodological review)

특정 연구 문제에서 사용되었던 연구 방법과 조작적 정의를 검증하려는 목적을 가진 리뷰 방법이다. 이를 통하여 연구 방법론의 적절성, 새로운 연구 방법에 대한 시사를 하려고 한다.

다. 통합적 리뷰(integrative review)

관련되거나 동일한 가설들을 다루고 있다고 판단되는 과거에 행해진 많은 독립된 연구들에서 전체적인 결론을 얻으려는 리뷰 방법이다.

2. 메타분석 이전의 통합적 리뷰 방법

메타분석 방법이 출현하기 훨씬 이전부터 이미 많은 학자들에 의해 기존의 연구 결과들을 통합하고자 하는 노력은 끊임없이 진행되어 왔다. 전통적인 통합적 리뷰 방법으로 불리는 이들 방법은 오늘날 행해지고 있는 메타분석이 계량적인 데 비해 다분히 질적인 분석의 성격을 띠고 있었다. 따라서 분석 기준이 주관적이고 보수적인 성격을 지니고 있다. 이러한 방법으로는 나열식 방법, 평균법, 투표식 방법, 군집 접근법이 있다.(오성삼, 2002:7-12)

가. 나열식 방법(The Narrative Method)

나열식 방법은 연구 결과들을 연도순에 따라 이야기체로 기술해 놓은 방법으로 흔히 수많은 연구들로부터 도출된 지식들의 통합된 형태를 이룩하기 위해 문헌 리뷰를 통하여 일련의 연구들 가운데서 정보를 추출해 내는 기법이다. 나열식 방법의 첫째는 특정 주제와 관련된 실증적 연구들을 수집하는 단계, 둘째는 실험연구 설계가 부적절한 연구물들을 제거해 버리는 단계, 마지막은 나머지 연구들로부터 결론을 도출해 내는 단계이다.

나열식 방법이 전제하고 있는 기본적 가정은 수집된 연구 가운데 위에서 열거한 이유들로 인해 많은 연구들이 제거되었다고 하더라도 리뷰하는 연구의 결론에는 아무런 영향을 미치지 않았다는 점을 전제로 하고 있다. 왜냐하면 연구자는 이미 주관적인 개인적 결론에 맞추어 나열을 진행시켜 나가고 있기 때문이다.

나. 평균법(The Averaging Method)

평균법은 나열식 방법에서 시도하는 단순한 나열보다는 체계적인 방법이라고 할 수 있다. 이 방법은 많은 수의 연구들을 통계적으로 사용하

기 위해 전체적인 평균치를 계산하는 방법이라 할 수 있는데, 이 방법의 단점은 중앙 집중치를 측정하는 과정에서 실험처치 효과의 실제 측정값들에 관한 정보를 상실하게 된다.

다. 투표식 방법(The Voting Method)

투표식 방법은 단순히 하나의 독립변수에 대해 어느 측정의 종속변수에 유의한 차이를 보이는가를 검정하는 방법으로, 사용 방법이 단순하여 여러 연구들에서 이용되어 왔다. 이 방법은 변수들 사이의 관계에서 어느 한쪽을 기준으로 통계적 유의도를 검정함으로써 정적으로(+) 유의한 차, 부적으로(-) 유의한 차를 보이거나 정적으로도 부적으로도 유의한 차를 보이지 않는 경우(0) 등 세 가지 결과 가운데 하나에 속하게 된다. 이처럼 각각의 연구 결과들이 위의 세 가지 유목 중 어느 유목에 속하는가를 투표하는 방식에 의해 분류한 후 각 유목에 속하는 연구 결과들의 숫자를 세어 세 가지 유목(+, -, 0)중 가장 많은 수를 확보한 유목을 우승자로 결정하는 방법이다.

투표식 방법은 사용이 간단하고 결론이 명쾌하다는 장점에 비해 몇 가지 단점이 있다. 첫째, 값진 기술적 정보(descriptive information)들을 상실하게 된다. 둘째, 이 방법은 각 단편적인 연구들의 표집의 크기를 고려하지 않기 때문에 적은 수의 표집에서 나온 결과와 많은 수의 표집을 사용한 연구 결과를 동등하게 비중을 줌으로써 경우에 따라서는 실제의 결론과는 반대의 결론에 도달할 가능성을 배제할 수 없다. 마지막으로 이 방법은 실험연구 결과와 상관관계의 연구를 구별하지 않는다는 점이다.

라. 군집 접근법(The Cluster Approach)

서로 상반되는 연구 결과들을 통합하기 위한 하나의 방법이다. 군집이란 군집 표집(cluster sampling)과 유사한 개념으로 쓰이는데, 이 접근

방법은 많은 모집단들을 군집이라 하고 하위 집단들로 세분될 수 있다는 실용적 통찰(pragmatic insight)에서 나왔다.

연구 결과 통합의 군집 접근법은 개별 연구 결과들 간의 어떠한 차이점들이 존재하는지를 고찰해야 한다는 점을 강조하고 있다. 여기에서 군집이란 하나의 모집단에서 추출된 표집이 아니라 모집단에서 자연적으로 통합된 상태를 의미하는데, 흔히 군집들은 광범위하고 체계적인 방법들과는 구분된다. 따라서 한 연구는 그 모집단 내에서 통합되는 여러 통합들에 따라 몇 개의 군집을 포함할 수 있다.

군집 접근법은 실제로 적용할 때 다음과 같은 단점이 있다. 첫째, 실제의 원 자료가 요구되는데, 과거에 이루어진 연구에서 원 자료를 구한다는 일은 사실상 어렵다. 둘째, 통계적·방법론적인 제약점으로 인해 연구 결과들을 제외시킨다. 셋째, 연구의 표집 대상들의 모집단을 잘 알고 있어야 한다는 점이다.

III. 메타분석 및 효과크기의 개념

1. 메타분석의 개념

Glass(1976)는 자료의 분석을 제1차 분석(primary analysis)과 제2차 분석(secondary analysis) 그리고 메타분석(meta analysis)의 세 가지로 구분하고 있다.

제1차 분석이란 연구에서 얻은 원 자료(raw data)들을 분석하는 것이며, 제2차 분석은 제1차 분석에서 제기되었던 연구 질문을 보다 나은 통계적 방법으로 사용하여 해답을 얻으려 하거나 기존의 자료들을 가지고 새로운 연구 질문에 해답을 구하려는 것이다. 그리고 메타분석이란 분석들의 분석(analysis of analysis)을 한다는 의미로 사용되고 있는데, 이것은 낱낱의 연구 결과들을 통합할 목적으로 많은 수의 개별적 연구나

결과들을 통계적 방법을 사용하여 분석하는 것이다.

Glass는 메타분석의 기법을 설명하면서 메타분석의 특이한 점들을 다음의 세 가지로 요약하였다. 첫째, 메타분석은 수량적²⁾이다. 메타분석에서는 단순한 자료들을 나열해 놓은 것이 아니라 다른 방법에 의존해서는 거의 파악하기 어려울 정도의 수많은 연구 결과들을 함축성 있게 분류하고 모종의 의미를 추출하기 위하여 계량적·통계적 방법을 사용한다. 따라서 메타분석은 그 통합하는 과정에 있어 원 자료를 그대로 사용하는 대신에 요약 통계(summary statistic)를 사용한다.

둘째, 효과크기(effect size)를 계산하기 위해 서로 상이한 연구들이 한데 모아진다는데 있다. 따라서 많은 수의 연구 결과들을 통합함에 있어 어떤 종류의 연구 결과가 기대하는 결과와 대다수의 연구 결과 간에 차이가 있다는 이유로 인해 분석 대상에서 제외되지 않는다. 다시 말해 서로 상이한 연구 결과라 할지라도 전체적인 결론을 도출해 내기 위해 분석 대상이 될 수 있다. 따라서 개개의 연구들을 연구자의 주관적 판단에 의해서 연구에 포함시켜야 할 연구와 포함시키지 말아야 할 연구들을 사전에 결정하지 않아야 한다.

셋째, 메타분석을 통해 일반적 결론을 도출해 낸다. 어떤 한 분야의 연구 결과에 대해 일정수의 연구 결과들이 긍정적인 효과를 산출해 낸다고 하더라도 그 효과크기가 각기 다른 경우에 서로 상이한 방향이나 서로 다른 효과크기들에 대해 일반적 결론을 찾기 위해서는 각각의 단편적인 연구들 사이에 존재하는 작은 차이는 무시되어도 무방하다는 전제하에서 일반화가 가능하다. 따라서 메타분석은 일반화와 실제적 간결성의 두 조건을 동시에 충족시키는 함수를 찾으려 노력한다.

메타분석의 개념을 이해함에 있어 메타분석의 학문적 용어는 그 사용되는 학문 분야에 따라 각기 다른 용어로 표현되는데, 일반적으로 Social Sciences에서는 meta-analysis, quantitative research synthesis를

2) 앞으로 더욱 효과적이고 의미 있는 연구 종합이 되기 위해서는 수량적 연구뿐만 아니라 질적인 연구물도 분석 대상이 되어야 한다는 학계의 요구가 있다. 이러한 연구의 하나가 Slavin(1986)의 'best-evidence synthesis'이다.

사용하고, Biomedical Sciences에서는 overview, pooling of results로 사용하며, Physical Sciences에서는 research synthesis, research review 등의 용어로 표현한다. 그러나 이와 같은 구분은 학문 분야뿐만 아니라 사용하는 학자들의 선호도에 따라 용어가 여러 가지로 혼용된다.(오성삼, 2002:14-15)

메타분석이란 같은 연구 주제로 실시된 많은 통계적 연구를 다시 통계적으로 통합하는 방법으로, 기존의 연구 방법이 갖는 제한적이고, 국부적이고, 편파적일 수 있는 방법과는 달리 통계적 기법을 사용하여 포괄적이며 거시적이고 객관성을 지닌 결론을 이끌어 내고자 하는 방법이다.(한국교육학회 교육평가연구회, 1995)

2. 메타분석의 장단점

가. 장점

메타분석은 각각의 연구들을 가지고 통계적인 방법으로 가설을 검정할 수 있으므로 보다 강력한 결론에 도달할 수 있다는 장점을 지닌다. 통계적 방법은 연구 결과들이 일관성이 없을 때 좀더 예리하게 연구의 초점을 맞출 수 있으므로 보다 더 체계적이며 정확한 결론을 내릴 수 있다. 연구물의 효과를 실증적으로 검정할 수 있기 때문에 메타분석은 전통적인 문헌연구보다 더욱 더 객관적이라 할 수 있다.

Jackson(1980)은 메타분석의 장점으로 일련의 연구 결과들을 통합하는데 있어 체계적이며 정확한 결론을 내릴 수 있고, 특정 주제에 관련된 여러 연구들의 결과를 정보자료로 사용하는 것이 가능하며, 사용 가능한 자료들의 여러 결점이 지니는 편견을 통제할 수 있으며, 복잡한 통계를 사용할 때 메타분석은 연구 방법, 연구 대상, 범위, 조건 그리고 실험 기간 사이의 관계를 동시에 규명할 수 있다고 보고 있다.

Pillemer와 Light(1980)는 여러 연구 결과를 종합함으로써 얻게 되는 이점으로, 개별 연구의 표집을 종합하여 좀더 큰 표집을 가지고 가설

을 검정하기 때문에 통계적 검정력을 높일 수 있고, 좀더 정확한 효과크기를 추정할 수 있으며, 관심 있는 두 변수 간에 존재하는 관계의 양상을 잘 파악할 수 있는 동시에, 상충하는 연구 결과가 있을 때 그 원인이 어디에 있는지 살펴볼 수 있는 기회를 줄 수 있다고 하였다.(오성삼, 2002:17-18 재인용)

나. 단점

메타분석이 지닌 단점은 지나치게 결과에 초점을 맞추고 있고 중재변수의 개입이나 상호작용 효과를 무시함으로써 한 연구의 결과를 지나치게 단순화시킨다. 또한 메타분석을 통하여 상호작용의 존재 여부를 실증적으로 규명하기 위하여 연구의 특성을 코딩하고 있으나, 실제로 대부분의 메타분석은 상호작용 효과를 거의 무시할 수 있다.

메타분석에 관한 Eysenck(1978)와 Glass(1982)의 비판은 첫째, 마치 사과와 오렌지를 한데 섞는 것과 같이 메타분석에서는 서로 비교할 수 없는 다른 성질의 연구 결과들을 종합하려는 점이다. 연구의 성격이나 방법이 현격하게 서로 다른 연구의 결과들을 한데 섞어서 종합하는 것은 무리이므로, 극단적인 자료는 제외시키는 것이 바람직하다는 지적한다. 둘째, 소위 ‘우수한’ 연구와 ‘줄렬한’ 연구의 결과를 구별하지 않고 그대로 종합하는 데에 있다는 점이다. 각각 개별 연구의 질적 수준에 따라 적절한 비중을 두고 종합하는 것이 좀더 합리적이다. 연구의 질과 관련하여 사례수가 큰 연구 결과는 사례수가 적은 연구 결과에 비해 보다 큰 가중치를 부여해야 한다는 논리에 의해 가중치를 적용해야 하는 방법이 있다. Hedges는 표준오차 값을 사용하여 메타분석에서의 적절한 가중치를 산출해 내는 공식을 제안하였다.(오성삼, 2002:179-184) 효과크기에 역변량가중치(W)를 적용하면 음수(-)값은 모두 양수값으로 변하게 되고, 그 해석도 일반적인 효과크기 해석(4장 참조)의 방법도 달라진다.

$$w = \frac{1}{SE^2} \quad [\text{역변량가중치 공식}]$$

3. 효과크기의 개념

효과크기란 실험집단의 평균값에서 비교집단의 평균값을 뺀 값을 비교집단(또는 두 집단)의 표준편차 값으로 나눈 값을 의미하는데, 공식으로 표시하면 $ES = \frac{\overline{X}_e - \overline{X}_c}{Sd_c}$ 이다. 여기서 $\overline{X}_e$ 는 실험집단의 평균값을, $\overline{X}_c$ 는 비교집단의 평균값, 그리고 Sd_c 는 비교집단의 표준편차를 의미한다. Glass와 그의 동료들에 의해 제안된 이 효과크기의 개념은 각기 다른 통계적 방법에 의해 이루어진 연구 결과들을 수량적으로 통합하고자 할 때 서로 다른 척도와 방법을 사용하여 얻은 연구 결과를 의미 있게 비교할 수 있도록 하나의 공통 척도로 나타내는 집단간 표준화된 평균치의 차이를 표시하는 방법이다. 효과크기(ES)는 메타분석을 가능하게 해주는 단위이다. 즉, 효과크기는 분석 대상이 되는 선행 연구의 다양한 형태의 결과들을 통합 또는 비교가 가능하도록 공통의 단위로 변환시켜 놓은 것이다. 그러므로 효과크기는 전통적인 리뷰 방법이 표집 크기에 따른 영향을 너무 많이 받을 뿐만 아니라 효과의 방향성과 크기를 제시해 주지 못하는 제한점을 극복하기 위한 대안으로 개발된 방법이다. 효과크기를 표현하는 방식에는 각각의 연구 상황에 따라 여러 가지 방식이 존재하며, 효과크기를 계산하는 방식 또한 다양하다.(오성삼, 2002:15-16)

4. 효과크기의 해석

메타분석의 결과로 제시되는 효과크기의 해석으로는 먼저, Cohen이

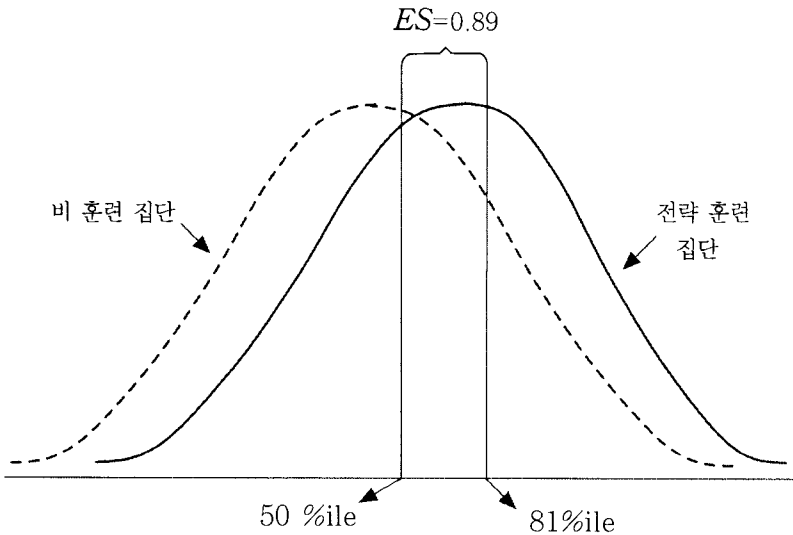
제시한 개괄적 기준을 사용하는 방법이 있다. Cohen은 표준화된 효과크기의 해석을, $ES=0.2$ 이하는 작은 영향, $ES=0.5$ 는 중간 정도 영향, $ES=0.8$ 이상은 큰 영향이라고 해석한다.³⁾

다음으로, 앞의 Cohen(1969)의 방법으로 효과크기를 의미 있게 해석하기 어렵기 때문에 Cohen(1977, Lipsey & Wilson, 2001:147 재인용)이 제시한 보다 정밀한 기준으로 비중복 계수(U_3 : percentile of nonoverlap)를 사용하여 효과크기를 해석하는 방법을 사용한다. 비중복 계수는 효과크기를 확률적으로 해석하는 방법으로 효과크기의 개념이 Z 점수와 동일함을 이용한 방법이다. 비중복 계수는 ES 값에 상응하는 Z 점수를 찾아 이를 다시 백분율로 표시한다.⁴⁾

비중복 계수를 사용한 예를 들어 보면, 선행 논문을 수집하여 메타분석을 한 결과 질문생성전략 훈련 집단이 훈련을 하지 않은 집단에 비해 독해력에서 0.89의 효과크기($ES=0.89$)가 있음을 밝혔다.(조재운, 2005) 이 효과크기를 이해하기 쉽도록 U_3 로 나타내고, 이를 정상분포곡선 그림으로 표시할 수 있다. $ES=0.89$ 에 해당하는 $F(Z)$ 값은 .8132671이므로 이를 백분율로 고치면 81.33%가 된다. 이 81.33%가 비중복 계수가 되며, 이는 질문생성전략 훈련을 받은 학생이 이 훈련을 받지 않은 학생보다 독해력에서 31%ile 정도 향상되었음을 뜻한다. 이를 정상분포곡선으로 그려보면 다음과 같다.

3) 표준화된 효과크기에 따른 것 외에 상관계수 효과크기에 따라 $ESr=0.1$ 이하는 작은 영향, $ESr=0.25$ 는 중간 정도의 영향, $ESr=.04$ 이상은 큰 영향을 준 것이라고 해석한다.(Lipsey & Wilson, 2001:147)

4) 효과크기, 상관계수, PV, U_3 , BESD의 관계는 Lipsey & Wilson(2001)의 153쪽 참조.



[그림 1] 질문생성전략 훈련이 독해력 향상에 미친 효과

IV. 메타분석의 절차

메타분석을 위해서는 통계적인 분석 이전에 거쳐야 할 여러 단계가 있다.⁵⁾ 이 글에서는 Cooper(1998)의 연구 단계를 제시하고자 한다. Cooper는 연구의 단계를 연구 문제 설정(Problem Formulation), 자료 수집(Data Collection), 자료 평가(Data Evaluation), 통계 분석 및 해석(Analysis and Interpretation), 자료 제시(Public Presentation)로 나누고

5) Jackson(1980)은 통합 연구를 위해 거쳐야 할 단계로 첫째, 연구를 위한 질문이나 가설을 선정하는 단계, 둘째가 통합 연구를 위한 연구 논문을 표집·선정하는 단계, 셋째가 연구의 자료들을 코딩하거나 연구의 특성들을 밝혀내는 단계, 넷째가 연구 결과들을 통계적으로 분석하여 통합하는 단계, 다섯째가 통계 분석된 자료들을 해석하는 단계, 여섯째가 메타분석을 통해 얻어진 결과를 작성하는 단계로 제시하고 있다.(오성삼, 2002:39 재인용)

있다. 이 같은 단계는 일반적으로 행해지는 실증적 연구 절차와 크게 다르지 않으나 근본적인 차이는 분석 대상이 한 편의 연구 결과라는 점이다.

1. 연구 문제 설정

일반적으로 어떠한 형태의 연구를 막론하고 연구가 전개되는 첫 번째 단계는 연구를 통해 밝혀내고자 하는 연구 문제의 진술 또는 연구의 가설 설정으로부터 출발한다. 이와 관련하여 메타분석에서 영 가설이나 대안 가설이 반드시 사전에 설정되어야 하느냐의 물음에 대해 똑같은 대답은 있을 수 없다. 왜냐하면 일반적으로 영 가설이나 대안 가설의 설정은 집단간의 차가 우연에 의해 발견되어질 가능성이 어느 정도인지를 규명할 수 있기 때문이다. 뿐만 아니라 영 가설이나 대안 가설의 설정, 이를테면 일방점증에 있어 연구 결과에 관한 알파수준과 메타수준을 미리 정해 놓을 수 있다. 그러나 만약 연구자가 메타분석에 있어서 통계 결과의 추리에 별다른 관심이 없다면 연구의 형식적인 영 가설이나 대안 가설을 반드시 설정할 필요는 없다.

2. 자료 수집

통합 연구의 성패는 연구하고자 하는 주제와 관련하여 얼마나 많은 양의 자료, 즉 연구 결과들을 수집했는가에 의해 좌우된다. 다시 말해서 선정된 주제와 관련된 발표 또는 미발표 연구물들을 가능하면 하나도 빠짐없이 수집함으로써 통합 연구가 의도하는 보다 신뢰성 있고 객관성 있는 포괄적 결론을 내릴 수 있는 논거를 제공하기 때문이다.

연구 논문의 선정 문제와 관련해서 생각해야 할 다른 문제는 한 연구의 결과를 몇 개의 서로 다른 저널이나 학회에 발표한 경우이다. 이 경우 흔히 동일한 연구의 내용을 각각 학술 저널의 성격이나 학회의 주제에

따라 조금씩 시각을 달리하는 제목으로 발표되는 경우가 있어 연구 결과의 검토 없이 각기 다른 연구로 간주하여 이중 삼중으로 한 편의 연구가 분석될 수 있다. 석·박사 학위논문이 학술지에 내용을 요약하여 게재된 경우가 그 예이다. 이렇게 한 편의 연구, 즉 동일한 데이터를 사용한 연구가 이중 삼중으로 중복되어 각기 다른 연구의 결과처럼 분석됨으로써 연구 결과를 어느 한쪽으로 유리하거나 불리하게 만드는 데 일익을 담당하게 되는 오류를 범하게 된다. 따라서 메타분석을 위해 선정된 자료의 내용들은 그 중복 가능성을 배제하기 위해 엄밀하게 검토할 필요가 있다.

가. 연구 자료의 출처와 소재 파악

메타분석을 시도하는 연구자는 메타분석에 사용될 연구 자료들의 출처를 우선 규정지어야 하는데 일반적으로 선정된 주제와 관련한 선행 연구물들을 찾아내는 출처로는 학술지에 실린 논문 목록이나 단행본으로 발간된 서적들, 석·박사 학위논문집, 학회에서 발표된 논문들 그리고 심지어는 현재 출판 예정 중에 있는 논문이나 미발표 논문까지도 발굴하여 연구의 대상에 포함시켜야 한다.

메타분석의 주제와 관련된 연구 결과들이 발표된 학술 논문집이나 자료들의 소재를 확인하는 작업은 인내와 끈기가 필요한 과정이다. Cooper (1982)는 이와 같은 작업에 효과적인 접근 방법으로 상향추적(ancestry) 접근 방법과 하향추적(descendancy) 접근 방법을 제시하고 있다. 상향추적 접근 방법은 한 연구발표 논문에서 인용한 연구를 추적하는 방법으로 연구 논문마다 그 논문과 관련이 있는 선행 연구의 결과를 인용한 출처나 참고 문헌을 제시하고 있는데, 이들을 근거로 하여 관련 주제의 연구를 추적하는 방법이다. 하향추적 접근 방법은 각종 연구주제별 목록집 등을 이용하여 우리가 찾고자 하는 연구들을 발견해 내는 방법이다.

메타분석을 위해 선정된 주제와 관련된 자료가 너무 방대할 경우 그 방대한 자료를 모두 수집하기란 경제적인 면에서나 시간과 노력의 측면

에서도 그리 쉬운 일이 아니다. 따라서 이런 경우에는 합리적인 방법에 의해 분석의 적정 규모를 선정해야 하는데, 이 때 쓰는 방법이 무선 표집 방법에 의해 제한된 수의 연구 논문을 표집하는 방법, 연도순에 따라 일정 기간으로 연구의 범위를 제한하는 방법⁶⁾, 연구 주제를 포괄적인 것으로부터 보다 구체적인 주제에 초점을 맞추므로써 선행 연구 결과의 수집을 제한하는 방법이 있다.

나. 자료 수집시의 유의점

메타분석의 자료를 수집할 때 생기는 문제에 따른 의 유의점은 다음과 같다.

1) 연구 논문의 질 판단 문제

연구의 질이 떨어지는 논문을 메타분석에 포함시킬 것인가 아니면 배제시킬 것인가의 문제와 관련하여 어려움을 겪게 된다. 그러나 모든 연구들이 오차를 내포하고 있다는 사실을 염두에 두고 결정해야 한다. 연구 방법론상 질의 문제는 양질의 논문과 불량 논문을 분류하는 식의 양분된 개념이 아니라 연속선상에서 결정되는 사안이다. 따라서 연구 논문의 질을 결정할 때 지나치게 엄격한 제약을 전제로 하면 메타분석 연구 결과의 일반화 문제가 제기될 것이고 너무 느슨하여 허용적일 경우에는 불량한 연구 결과의 혼용으로 인해 결론적으로 제시되는 효과크기의 신뢰구간 설정에 문제가 생기게 된다. 따라서 분석 대상 연구논문의 질을 결정하는 문제는 연구자 자신이 적절한 선에서 메타분석의 균형을 유지하는 방안을 마련해야 한다.

6) 질문생성전략이 독해력이 미치는 영향에 대한 메타분석인 조재운(2005)의 연구는 연구 대상을 2003년까지로 하고 있다. 이 연구 후에도 질문생성전략과 관련된 실증적 연구로 고도연(2004), 한유경(2004) 등이 있다.

2) 분석 대상 논문 선정의 범위 문제

대부분의 메타분석 대상으로 선정되는 논문은 기존에 발표된 논문만을 대상으로 하게 되지만 일반적으로 발표된 논문의 경우에 유의성 검정에 성공한 논문이 주류를 이루고 있어 연구 결과 유의하지 않은 사실에 대한 통합을 간과하는 경향이 있다. 따라서 메타분석의 의미를 최대한 살리기 위해서는 가능하다면 기존에 출판된 논문뿐만 아니라 미출판 논문을 함께 포함시키기 위한 노력이 따라야 한다.

3) 극단의 효과크기 처리 문제

극단의 효과크기들이 메타분석에 부적절한 비율을 초래하기 때문에 분석 대상이 되는 효과크기가 극단의 값을 가질 경우에는 극단의 값을 분석 대상에서 제외시키든지 아니면 극단의 점수를 교정점수로 전환하여 사용하는 방법을 마련해야 한다. 이 경우에는 극단의 점수를 어떻게 처리하였는가에 대하여 진술해야 한다.

4) 한 개의 연구에서 산출할 효과크기 개수 문제

한 개의 연구에서 한 개의 효과크기만을 산출해야 하기 때문에 단일 연구 내에서 하위 표집에 대한 효과크기를 한 개로 해야 한다.

연구의 특성이란 개별적 연구가 지니고 있는 특징적인 정보를 지칭한다. 예를 들어, 피험자의 수, 연구의 형태, 피험자의 연령, 결과의 보고 등이다. 이 정보 속에는 이미 계량적으로 기술된 것도 있지만, 유목적인 것도 있다. 계량화되지 않은 특성은 메타분석을 위해 계량화의 방법, 즉 코딩을 해야 한다.

3. 자료 평가

메타분석에 사용될 자료는 이론적인 것이 아니라 실증적 자료이어야 하며, 결과를 질적 형태가 아니라 수량적 형태로 제시된 것으로서 동일

한 구인(construct)이나 관계성(relationship)이 검증된 자료이어야 한다. 뿐만 아니라 상호 비교가 가능하도록 연구 결과를 제시하는 통계적 방법인 효과크기(effect size)나 상관계수(correlation Coefficients) 등으로 나타낼 수 있는 것이어야 한다.

메타분석을 위해서는 다음과 같은 전제 조건이 충족되어야 한다.

첫째, 통합하고자 하는 주제와 관련된 선행 연구물의 수가 충분해야 한다. 선정된 연구 주제와 관련하여 충분한 선행 연구를 수집할 수가 없다면 표본수의 부족으로 인하여 검정력이 약화되며, 대결론(big decision)을 위한 메타분석 본래의 뜻이 상실된다.

둘째, 연구의 설계 방식은 통제집단과 실험집단이 존재하는 실험연구이어야 한다. 따라서 메타분석을 위해 선정된 주제와 아무리 관련이 있는 연구라 할지라도 그 연구 자체가 실험연구(통제집단과 실험집단을 가진)가 아닌 사례 연구나 관찰·면접에 의한 기술적 연구 등의 질적 연구는 메타분석의 대상에서 제외한다.

셋째, 통제집단과 실험집단에 모두 평균점수와 표준편차, 사례수 및 유의도 수준이 밝혀진 연구 결과이어야 한다. 실험집단과 통제집단의 평균치, 표준편차 및 사례수나 유의도 등에 대한 정보가 제시되지 않더라도 T검정, F검정 상관계수(r) 등에 의한 통계값이 제시된 경우라면 통계적인 변환공식에 의해 각종 통계값의 결과를 효과크기(ES)로 고쳐서 사용해야 한다.

가. 연구 특성의 선택

연구 특성을 수량화하는 데 주로 포함되는 속성에는 연구 논문의 발표일, 연구 대상의 연령 계층, 남녀간의 표집 비율, 표집의 크기나 실험 대상의 탈락 등이 있다. 또한 연구 대상의 배정의 무선 또는 비무선 표집, 연구 논문의 발표 형태(학술지, 단행본, 학위논문 등) 등의 속성들이 있다. 메타분석을 시도하는 연구자에게는 분명한 준거를 가지고 연구 특성을 분류해야 한다. 그리고 연구의 특성을 선정하여 코딩하는 작업은 될

수 있으면 두 사람 이상의 평가자에 의해 분류 및 코딩을 하여, 앞에서 열거한 속성들을 분류하는 데 따른 평가자 상호간의 합치도를 계산하여 밝혀주는 것도 바람직하다.

나. 코딩 매뉴얼과 코딩표의 개발

최종 분석 대상 선행 연구 논문이 결정되고 나면 각각의 연구 논문에 대해 설정된 연구의 특성에 따라 코딩하는 작업에 들어가면 각각의 코딩 내용을 세분화하여 <표 1>과 같은 형태의 코딩표를 개발해야 한다. 그런데 여러 사람들이 단일 주제 하에 수집된 각각의 논문들을 코딩하다 보면 개개인의 생각과 판단이 달라 동일한 내용을 달리 코딩하는 결과를 범하게 될 경우가 발생할 수 있다. 뿐만 아니라 코딩 분류가 애매한 경우에는 동일인이 하는 코딩일지라도 동일한 내용을 앞서 코딩한 논문과 나중에 코딩하는 논문의 경우에 다르게 코딩할 경우도 발생하게 되어 연구의 신뢰도를 떨어뜨릴 수 있다. 따라서 이와 같은 실수를 방지하고 코딩하는 사람들의 판단을 명료화하기 위하여 코딩 매뉴얼 작성이 중요하다. 코딩 매뉴얼은 코딩표에 나타난 각각의 코딩 내용들을 보다 상세하게 설명해 놓은 것으로써, 경우에 따라서는 코딩에 제시된 변수가 구체적으로 무엇을 의미하는지에 대한 내용 설명과 함께 여러 가지 경우의 발생 가능한 예외적 상황에 대해 어떻게 코딩 처리해야 할 것인가의 지침서이다. 따라서 코딩 매뉴얼이 구체적이고 명료하게 만들면 만들수록 코딩의 오류를 방지할 수 있고 메타분석 연구의 신뢰도를 높일 수 있다.

<표 1> 코딩표 예시(조재운, 2004:25-26)

1. 일련 번호:
2. 저 자 명:
3. 논문 제목:
4. 대학명(학술지):
5. 출판 년도:

6. 독립 변수:
7. 종속 변수:
8. 학년 구분: ① 초1 ② 초2 ③ 초3 ④ 초4 ⑤ 초5 ⑥ 초6
⑦ 중1 ⑧ 중2 ⑨ 중3 ⑩ 고1 ⑪ 고2 ⑫ 고3 ⑬ 기타:
9. 실험 기간: ① 1~5주 ② 6~10주 ③ 11주 이상
10. 연구 대상수: ① 40명 이하 ② 41~80명 ③ 81~120명
④ 121~160명 ⑤ 161명 이상(명)
12. 자료 출처: ① 석사논문 ② 박사논문 ③ 학회지
13. 연구 대상(성별): ① 남 ② 여 ③ 남녀 공학
14. 연구 실시 지역: ① 특별시·광역시 ② 도시 지역 ③ 군 단위 이하
15. 통계 방법: ① t-검사 값:
② CR-테스트 자유도(df):
③ Correlation 표준편차(SD):
④ ANOVA 평균(M):
⑤ 카이 제곱
⑥ 실험집단 평균: , 표준편차:
 비교집단 평균: , 표준편차:
16. 검사 도구: ① 독립변수: a 표준화 검사지 b 비표준화 검사지
 ② 종속변수: a 표준화 검사지 b 비표준화 검사지
17. 효과크기(Effect size):

4. 통계 분석 및 해석

가. 메타분석의 통계 절차

메타분석을 위한 통계적 방법의 주된 작업은 분석 대상이 되는 기존의 선행 연구 결과에 제시된 다양한 형태의 통계값을 하나의 통일된 효과크기로 변환시키는 일이다. 일반적으로 공통된 효과크기로 나타내는 방법에는 표준화된 평균차(standardized mean difference)로 나타내는 효과크기(ES_{sm})와 통계적 유의성 검정(statistical significant test)으로서의 효과크기, 상관계수(correlation coefficient)의 효과크기(ES_r), 그리고

분할표의 메타분석 방법으로서의 승산비(odds-ratio)의 효과크기(ESor) 등이 있다.

선행 연구 결과의 통합 방법으로서의 메타분석 방법 역시 논리적인 면에서 일반적인 통계분석 방법의 절차와 동일한 과정을 거친다. 다시 말해서, 분석 대상으로 최종 선정된 전체 연구 논문에 대한 동질성(homogeneity) 검정을 실시한다. 만약 전체 연구의 동질성 검정이 실패할 경우, 즉 이질 분포를 이루고 있는 자료에 대해서는 각 그룹별, 특성별 분석 범주(category)로 나누어 각 범주에 포함될 연구 결과만을 대상으로 소집단 동질성 검정을 실시한다. 이와 같이 분석 대상 연구 논문의 동질성 유무에 대한 검정을 마친 이후에 전체 집단 또는 그룹별 평균 효과크기의 산출과 평균차의 검정 및 효과크기의 신뢰구간 등을 산출하는 과정을 거쳐 메타분석 통계를 적용한다.

나. 표준화된 평균차의 효과크기 계산 방법

연구물의 통계적 제시 방법에 따라 효과크기의 산출 방식이 다르다.

먼저, 표준화된 평균차의 효과크기는 통합 연구에 있어서 연속형 자료를 사용하여 측정된 두 집단 사이를 비교할 때 사용되는 방법이다. 두 집단은 경우에 따라 실험 상황, 즉 실험집단과 비교집단 또는 남녀의 성별처럼 자연발생적인 경우를 생각할 수도 있다. 표준화된 평균차를 사용한 효과크기에 의한 통합 방법은 메타분석에서 가장 널리 사용되는 방법인데, 일반적으로 분석 대상 논문이 집단간(실험집단 vs 비교집단) 비교를 실시함에 있어 연속형 자료를 통한 측정이 이루어진 경우에 사용되는 방법이다. 이와 같은 효과크기를 구하기 위해서는 [공식 1]처럼 연구에서 사용된 실험집단과 비교집단의 평균점수 간의 차이를 비교집단의 표준편차 값으로 나누어준 지수(index)를 산출한다.

$$\frac{\overline{X_e} - \overline{X_c}}{Sd_c} \quad [\text{공식 1}]$$

여기서, $\overline{X_e}$ 는 실험집단의 평균점수, $\overline{X_c}$ 는 비교집단의 평균점수, 그리고 Sd_c 는 비교집단의 표준편차를 의미한다. 이를 좀더 구체적으로 표현하면 메타분석에서 사용되는 효과크기(ES)란 실험집단의 평균치가 비교집단의 평균치에 비해 얼마나 더 효과적이었는가의 크기를 표준편차라는 공통의 척도로 변화시켜 놓은 것이다.

[공식 1]의 효과크기를 구하는 공식을 전제로 하여 여러 가지 형태의 통계량을 변환하여 효과크기로 환산하는 방법은 다음과 같다.⁷⁾

1) 표준화된 평균차의 직접 계산 방법

이 방법에 의한 계산은 [공식 1]과 같이 두 집단(실험집단, 비교집단)간의 평균값을 통합표준편차 값으로 나누어서 얻게 된다.

$$ES_{sm} = \frac{\overline{X_e} - \overline{X_c}}{S_{pooled}} \quad [\text{공식 2}]$$

[공식 2]에서 ES_{sm} 을 "D" 또는 "g"로 표현하기도 하며, 초기의 메타분석에서는 통합표준편차(S_{pooled}) 대신에 비교집단의 표준편차 값을 사용([공식 1] 참조)하였다. 그러나 요즘에 주로 사용되는 ES_{sm} 의 분모인 S_{pooled} 를 계산하기 위해서는 [공식 3]을 적용하면 된다.⁸⁾

$$S_{pooled} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \quad [\text{공식 3}]$$

7) 효과크기 계산 방법에는 상관계수(correlation coefficient)의 효과크기(ESr)나 승산비(odds-ratio)의 효과크기(ESor) 등이 있으나 이 글에서는 표준화된 평균차와 추정 통계량에 의한 효과크기 계산 방법을 알아보도록 한다.

8) 분모로서의 표준편차 선택 문제의 중요성에 대해서는 황정규(1988)를 참조.

예를 들어 조재윤(2002)의 연구에서 비교집단(31명)의 표준편차가 3.22(평균: 13.10), 실험집단(31명)의 표준편차가 4.06(평균: 15.29)일 때, [공식1]을 적용한 ES는 0.68, [공식2]를 적용한 ES_{sm} 는 0.60으로 두 효과크기에 차이가 있다. 요즘에는 비교집단의 표준편차를 분모로 하는 [공식1]보다는 통합표준편차를 분모로 하는 [공식2]가 연구의 효과성을 더 잘 설명하기 때문에 ES_{sm} 를 연구 분석의 효과크기로 한다.

표준화된 평균차에 의한 효과크기를 산출할 때 유념해야 될 사항은 만약에 표본수가 적은 경우에는 효과크기가 실제 값보다 크게 산출되는 경향을 나타내게 된다. 따라서 소표본(small sample size)⁹⁾으로 인한 문제점을 교정하기 위해서 [공식 4]를 사용한다.

$$ES'_{sm} = ES_{sm} \left[1 - \frac{3}{4N-9} \right] \quad \text{[공식 4]}$$

예를 들어 현애순(2002)의 연구에서 비교집단(12명)의 표준편차가 1.92(평균: 8.33)이고, 실험집단(12명)의 표준편차가 2.89(평균: 13.00)일 때, [공식2]를 적용한 ES_{sm} 는 1.91, 소표본 교정을 사용한 ES'_{sm} 을 적용한 ES_{sm} 는 1.78인데, 이 때 연구의 효과성을 잘 설명하는 것으로 ES'_{sm} 을 사용한다.

2) 추정 통계량의 효과크기 계산 방법

분석 대상이 되는 선행 연구 결과가 메타분석이 요구하는 평균값과 표준편차 값을 제시하지 않고 대신에 실험집단과 비교집단의 평균값의 차에 대한 유의도 검정 결과만을 제시하고 있는 논문의 경우에는 ES_{sm} 대신에 별도의 공식을 적용해 효과크기를 산출한다.

9) 소표본이란 일반적으로 실험집단의 사례수가 20 이하인 경우를 말한다.

가) t -값에 대한 효과크기 계산

효과크기로 전환하기 위한 선행 연구 결과의 자료가 실험집단과 비교 집단 사례수와 두 집단간 평균값에 대한 차이를 t -검정 방법을 사용한 경우에 효과크기를 산출해 내기 위해서는 [공식 5]를 적용한다.

$$ES_m = t \sqrt{\frac{n_1 + n_2}{n_1 n_2}} \quad [\text{공식 5}]$$

나) F -값에 대한 효과크기 계산

분석 대상이 된 선행 연구 결과가 두 집단의 일원분산분석 F 검정을 사용한 경우에는 [공식 6]을 적용하여 효과크기를 구한다.

$$ES_{sm} = 2\sqrt{\frac{F}{N}} \quad [\text{공식 6}]$$

여기서 N 은 (실험집단 사례수 + 비교집단 사례수)이다. [공식 6] 이 외에 [공식 7]과 같이 실험집단과 비교집단의 사례수를 이용하여 구 하기도 한다.

$$ES_{sm} = \sqrt{\frac{F(n_1 + n_2)}{n_1 n_2}} \quad [\text{공식 7}]$$

V. 결론

전통적인 통합 분석 방법과 메타분석 방법의 특징을 비교하여, 전통 적 통합 분석 방법은 연구 문제의 정립이 매우 한정되어 이론 지향적이

고, 분석을 위한 각 기존 연구의 선택이 다분히 연구자에 의해 자의적으로 이루어진다. 또한, 전통적 방법은 기존 연구 결과의 정리와 해석 역시 주관적 판단에 의존하게 되고, 전제된 문제 및 가설 등에 부합하지 않거나 상반되는 것은 무시하거나 배제해 버리는 경우가 보통이다. 이에 비해 메타분석의 특징은 가능한 문제나 가설 등의 범위를 넓게 설정하고, 분석 대상으로서의 기존 연구를 가능한 많이 선정하며, 그 선정은 임의적으로 시도한다. 또한, 메타분석은 전제된 문제 또는 가설 등에 부합하지 않거나 상반되는 결과들도 세심히 탐색함으로써 새로운 주요 독립변수 또는 중재변수 등을 발견함으로써 새로운 설명 가능성까지를 모색한다.(이지훈, 1993)

메타분석은 계량적 접근을 하고, 종합하고자 하는 연구 분야에서 이루어진 모든 연구를 분석 대상으로 삼으며, 체계적이고 반복이 가능하다. 그러므로 메타분석은 어느 한 연구 분야에서 이루어진 연구의 결과를 종합하여 해당 연구 문제에 대한 일반적인 결론을 내리는 것을 주목적으로 한다.

메타분석 방법으로 종합된 연구 결과는 비교적 덜 치우친(less biased) 표본으로부터 얻은 결과로 그 변수나 구인에 대하여 폭넓고 강한 증거로 활용되고, 개개 연구물의 결과보다 통계적 검정력이 높다.

국어교육학 연구 방법론으로 메타분석은 유용한 도구이다. 전통적인 리뷰 방법에만 머물지 않고 국어교육학의 실증적 연구물을 종합하여, 또 때로는 외국의 언어교육 연구물과 종합하여 신뢰가 높고 타당성 있는 대 결론을 이끌어 내는 메타분석 연구가 우리 국어교육학에서도 적극적으로 이루어지길 기대한다.*

* 본 논문은 2005. 2. 19. 투고되었으며, 2005. 3. 12. 심사가 시작되어 2005. 3. 25. 심사가 종료되었음.

□ 참고문헌

- 구병두(1989), “학업성취 관련 변인에 대한 Meta분석”, 건국대 박사학위논문.
- 박도순(2002), 『교육연구방법론』, 문음사.
- 성태제(1999), 『현대 기초통계학의 이해와 적용』, 양서원.
- 오성삼(2002), 『메타분석의 이론과 실제』, 건국대 출판부.
- 윤준채(2002), “SSR이 읽기 태도 및 읽기 이해에 미치는 영향: 효과성에 대한 메타 분석”, 『21세기 국어교육학의 현황과 과제』 587-608, 한국문화사.
- 윤준채·김라연·노명완(2004), “국어교육 연구 방법론으로서의 연구 종합: 독서교육 연구에 대한 종합적 이해”, 『독서연구』 11, 255-273, 한국독서학회.
- 이지훈(1993), 『사회과학의 메타분석방법론』, 충북대 출판부.
- 임채수(2001), “협동학습 수업전략의 학습효과에 관한 메타 분석”, 경남대 박사학위논문.
- 조재운(2002), “질문만들기 지도가 질문만들기 능력과 독해력 향상에 미치는 효과 연구”, 고려대 석사학위논문.
- 조재운(2004), “질문생성전략 훈련의 국어과 읽기 이해 능력 학습 효과에 대한 메타분석”, 『2004년 봄 전국학술대회 자료집』 1-26, 한국국어교육연구학회.
- 조재운(2005), “질문생성전략 훈련이 독해력에 미치는 효과에 대한 메타분석”, 『국어교육』 116, 125-149, 한국어교육학회.
- 차부근(2001), “학업성취와 관계된 지적 특성에 대한 메타분석”, 고려대 석사학위논문.
- 한국교육학회 교육평가연구회(1995), 『교육 측정·평가·연구·통계 용어사전』, 중앙교육진흥연구소.
- 황정규(1988), “메타분석의 이론과 방법론”, 『성곡논총』 19, 1-55.
- 현애순(2002), “자기질문전략이 초등학교 읽기부진아의 요약 능력 및 독해력에 미치는 영향”, 인천교대 석사학위논문.
- Cooper, H. (1982), Scientific guidelines for conducting integrative research reviews. *Review of Educational Research*, 52, 291-302.
- Cooper, H. (1998), *Synthesizing research*(3rd ed). SAGE Publications, Inc.
- Glass, G. V. (1976), Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.
- Glass, G. V., McGaw, B. & Smith, M. L. (1981), *Meta-analysis in social research*. SAGE Publications, Inc.
- Lipsey, M. W., Wilson, D. B. (2001), *Practical meta-analysis*. SAGE Publications, Inc.

Slavin, R. E.(1986). Best-Evidence Synthesis: An alternative to Meta-Analytic and Traditional Reviews. Educational Researcher, 15(9), 5-11.

〈초록〉

국어교육학 실증적 연구의 종합을 위한 메타분석 방법론

조재운

학문이 발달하고 복잡해짐에 따라 각 학문의 연구물이 축적되어 그 연구물들에 대한 종합적인 연구가 필요하게 되었다. 메타분석 방법 이전에 연구 종합 방법으로 나열식 방법, 평균법, 투표식 방법, 군집 접근법이 있었으나 통계적 기법을 사용하여 포괄적이며 거시적이고 객관성을 지닌 결론을 이끌어 내고자 하는 메타분석이 생기게 되었다.

메타분석은 연구 문제 설정, 자료 수집, 자료 평가, 통계 분석 및 해석, 자료 제시의 절차로 이루어진다. 메타분석을 위한 통계적 방법의 주 작업은 분석 대상인 선행 연구 결과에 제시된 다양한 통계값을 통합 또는 비교가 가능하도록 통일된 효과크기로 변환시키는 일이다.

효과크기를 계산하는 방식이 다양한데, 대표적인 효과크기 계산 방식은 표준화된 평균차의 직접 계산법이다. 초기의 메타분석에서는 통합표준편차(S_{pooled}) 대신에 비교집단의 표준편차값을 사용했으나 요즘은 주로 S_{pooled} 을 적용한 효과크기 공식을 사용한다.

$$ES_{sm} = \frac{\overline{X_e} - \overline{X_c}}{S_{pooled}}$$

국어교육학 연구 방법론으로 메타분석은 유용한 도구이다. 국어교육학의 실증적 연구물을 종합하여 신뢰가 높고 타당성 있는 대결론을 이끌어 내는 메타분석 연구가 활발히 이루어지길 바란다.

【핵심어】 메타분석, 연구 종합, 연구 방법, 효과크기, 효과크기 공식, 통계적 검정.

〈Abstract〉

Meta Analytic Methodology for Synthesizing Research in Korean Language Education

Cho, Jae-yoon

As the academy has been developed and complicated, the research results have been accumulated and required to be further studied collectively themselves. Meta analysis has been born as the statistical method of accumulating experimental research.

The meta analysis has been consisted of the research problem formulation, data collection, data evaluation, statistical analysis and interpretation, and finally public presentation.

In the beginning of meta analysis, standard deviation value of comparable group had been used rather than united standard deviation, but the size effect formula has been used nowadays applying S_{pooled} .

$$ES_{sm} = \frac{\overline{X_e} - \overline{X_c}}{S_{pooled}}$$

The meta analysis is a useful method as the Korean language education research method. Thus, it is hopeful to actively study of meta analysis research which leads the results which are high reliable and feasible with synthetic positive research data of Korean language education.

【key Words】 meta analysis, research synthesis, research method, effect size, effect size formula, statistic test.