

통계청 학술연구용역

『통계품질평가지표 개발』

- 통계작성절차를 중심으로 -

(최종보고서)

2003. 9. 30

연구수행기관 : (사)한국조사연구학회

연구 책임자 : 류제복(청주대학교 생명·유전·통계학부 교수)

연 구 원 : 유정빈(서원대학교 응용통계학과 교수)

연 구 원 : 김선웅(동국대학교 통계학과 전임강사)

머 리 말

본 보고서는 (사)한국조사연구학회에서 수행한 통계청의 「통계품질평가지표개발」에 대한 연구용역의 최종보고서이다.

통계품질에 대한 중요성이 90년도 중반부터 주요 국가들로부터 일기 시작하였고, 최근에 들어서는 우리나라에서도 통계품질의 중요성과 필요성을 인식하고 이의 향상을 위한 노력을 기울이고 있다. 통계청에서는 품질을 향상시키기 위해서 현재 통계청에서 생산하고 있는 통계들에 대한 품질을 평가하고 있다. 48개의 품질평가지표를 절차별로 마련하고, 이들 지표를 이용해서 2002년 처음으로 통계청에서 생산한 12개의 통계에 대한 품질을 평가하기에 이르렀다. 통계품질은 제품의 경우와 마찬가지로 사용자들의 사용적합성을 얼마나 만족시키는가로 정의된다. 현재까지는 주로 생산자들의 측면에서 통계생산과 관리가 이루어지고 있으나 앞으로는 이용자 측면에서 다루어져야 할 것이다.

아직까지 통계품질에 대한 인식도 부족하고 이를 평가하는 지표도 미흡한 상태라, 실제로 우리나라의 통계품질 수준이 어느 정도인지를 정확히 측정하기가 곤란하였다. 따라서 우리의 통계품질수준을 다른 나라들과 비교하기가 어려운 실정이다. 이에 통계청에서는 통계품질관리팀을 새로이 만들어 품질향상을 위한 노력을 기울이고 있다.

본 연구에서는 이러한 국내외적인 변화와 요구에 따라 우리나라의 통계품을 평가하는 새로운 지표를 개발하게 되었다. 통계품질평가지표 개발에 국내에서 참고할 만한 자료가 부족한 관계로 대부분을 국제기구나 주요 국가들에서 사용하고 있는 평가지표들을 검토하여 사용하였다. 현행 48개의 평가지표를 수정 보완해서 7개 차원별로 81개의 지표를 개발하고 이들 지표 중에서 필수지표로 48개를 선정하였다. 개별 평가지표의 필요성과 평가내용들을 정리하고 절차별로 구분하였다. 또한 전수조사, 표본조사, 그리고 가공통계로 나누어 평가에 필요한 지표를 분류하였다. 이들 지표는 모든 통계에 대해 동일하게 사용될 필요는 없고 통계별로 달리 사용될 수 있으며, 시간의 흐름과 변화하는 사회적 요구에 따라 새로운 지표가 개발되고 기존의 지표가 수정 보완되는 등의 노력이 지속되어야 할 것이다. 현재로서는 통계별로 품질지수를 산출하는데 많은 어려움이 있지만 향후 연구를 위해서 두 가지 방법의 통계품질지수를 제안하였다. 통계품질평가지표의 개발이 평가보다는 궁극적으로 통계의 품질을 높이는 데 초점을 맞추어야 할 것이다.

국내의 사전 연구가 거의 전무하고 참고자료도 부족한 상황에서 본 연구를 위해 많은 시간을 할애한 연구진들과 연구 진행에 모든 지원을 아끼지 않은 통계청 관계자 여러분들께 감사사를 드린다. 통계의 품질향상을 위해서는 통계이용자, 통계생산부서, 품질관리팀, 그리고 학계 전문가들과의 공동연구가 지속되어야 소기의 목적을 달성할 수 있을 것이다.

2003년 9월 30일

연구책임자

청주대학교 교수 류제복

목 차

제 1장 서론	1
제 2장 현행 통계품질평가지표	3
2.1 통계품질 결정요인 / 3	
2.2 통계품질 평가지표 / 8	
제 3장 2002년도 통계품질평가결과 분석	11
3.1 자체 및 외부 평가결과분석 / 11	
3.2 작성절차에 대한 분석 / 12	
3.3 차원별 분석 / 17	
3.4 기타 분석결과 / 21	
제 4장 외국 사례연구	22
4.1 국제통화기구(IMF) / 22	
4.2 경제협력개발기구(OECD) / 25	
4.3 Eurostat / 28	
4.4 캐나다 / 30	
4.5 영국 / 36	
4.6 스웨덴 / 42	
4.7 네덜란드 / 46	
4.8 미국 / 47	

제 5장 새로운 통계품질평가지표	50
5.1 차원별 품질평가지표 / 50	
5.2 품질평가지표의 필요성과 평가내용 / 53	
5.3 절차별 품질평가지표 / 66	
5.4 통계별 품질평가지표 / 68	
제 6장 통계품질지수	75
6.1 네덜란드 품질지수 / 75	
6.2 적용 가능성 및 제안 / 82	
제 7장 결론 및 제언	91
참고문헌	95
부 록	97
<부록 1> 절차별 세부 분석결과 / 97	
<부록 2> 각 문항 및 평가방법에 따른 비교 / 108	
<부록 3> 평가방법에 따른 평균 점수 차이에 대한 분석 / 124	
<부록 4> 차원별 분석결과(그래프 모음) / 126	
<부록 5> 기타 : '관계없음'에 대한 도수 차이 / 130	

< 표 목차 >

<표 1-1> 기관별 승인통계 작성현황	2
<표 2-1> 통계품질평가 차원별 비교	7
<표 2-2> 대한민국 통계청(절차별)	9
<표 2-3> 대한민국 통계청(차원별)	10
<표 3-1> 기초분석	13
<표 3-2> 자체평가와 외부평가 평균점수 간의 유의성 검정	13
<표 3-3> 3점 척도로 고려할 때, 미흡판정의 비율	13
<표 3-4> 절차별 평균점수 차이(크기 순)	14
<표 3-5> 외부평가자가 지적한 '매우 미흡'과 '미흡'항목	15
<표 3-6> 외부평가가 자체평가보다 점수가 높은 항목	15
<표 3-7> 평가별 평균점수 차이가 큰 항목(0.5이상)	16
<표 3-8> 차원별 문항분포 내역	18
<표 3-9> 평가점수에 의한 비교	19
<표 5-1> 차원별 품질평가지표(81개 항목)	50
<표 5-2> 품질평가지표의 필요성과 평가내용	54
<표 5-3> 절차별 품질평가지표	66
<표 5-4> 전수조사통계 품질평가지표(62개 항목)	69
<표 5-5> 전수조사통계 품질평가지표의 필요성과 평가내용 ...	72
<표 5-6> 가공통계 품질평가지표(41개 항목)	71
<표 5-7> 가공통계 품질평가지표의 필요성과 평가내용	73
<표 6-1> 정확성에 대한 지수	78
<표 6-2> 일관성에 대한 지수	79
<표 6-3> 시의성에 대한 지수	80
<표 6-4> 정보수준에 대한 지수	81
<표 6-5> 최종 통계품질지수	81
<표 6-6> 손실함수를 이용한 품질지수	84
<표 6-7> '100'을 기준으로 변환된 통계품질지수	86
<표 6-8> 절차별 시그마 수준	88
<표 6-9> 각 차원 및 절차별 시그마 수준	90

< 그림 목차 >

<그림 3-1> 12개 통계에 대한 평가방법에 따른 응답현황	11
<그림 3-2> 자료공표 절차와 조사표설계 절차에서의 평가방법에 따른 응답현황	14
<그림 3-3> 차원 및 평가방법별 평균점수	17

제 1장 서론

현대 사회에서 국가의 중요한 기능 중 하나는 국가의 기간산업을 육성하고 발전시키는 일이다. 많은 산업 중에 정보산업이 차지하고 있는 비중은 나날이 커지고 있다. 정보의 근원은 통계이고, 정확하고 신속한 통계는 국가의 각종 정책을 결정하는데 중요한 역할을 담당한다. 뿐만 아니라, 기업이나 개인에 이르기까지 통계의 사용은 그 활용 범위가 점차로 확대되고 있다.

현재 통계청에서는 통계의 품질을 향상시키기 위해서 통계품질관리팀을 운영하고, 통계청에서 생산하고 있는 통계에 대한 품질평가를 2002년부터 실시하고 있다. 이는 정부, 연구기관, 기업, 그리고 개인에게 보다 나은 의사결정을 하기 위한 고품질의 통계정보를 제공하고 나아가 삶의 질을 향상시키기 위함이다.

지금까지는 주로 필요한 통계의 생산에 주력해 왔다. 하지만 정책결정의 성과는 사용된 통계의 질에 의해 큰 영향을 받게 된다. 품질에 대한 문제는 주로 특정부분에 한정되어 왔으나, 현대 사회로 들어와서는 재화나 서비스의 품질향상은 물론이고 정보의 핵심인 통계의 품질을 향상시키는 일이 중요하다는 것을 인식하기에 이르렀다. 품질이 낮은 통계의 사용은 미시적으로 잘못된 결과를 초래하기도 하지만 거시적으로 오류의 범위를 확대시켜서 치명적인 손실을 초래한다. 따라서 통계의 품질을 향상시키는 일은 장기적으로 국가 발전에 필수적이다.

<표 1-1>에 의하면 2003년 8월 1일 현재 통계법 제8조에 의거 승인 받은 통계는 총 446종이고, 이들을 130개(61개의 정부기관과 69개의 지정기관)기관에서 작성하고 있다. 그밖에도 많은 조사기관이나 각종 단체에서 필요한 통계들을 생산하여 사용하고 있다. 따라서 통계의 품질향상을 위해서는 특정 부서만의 노력으로는 불가능하고 모든 통계생산 프로그램과 이들을 관장하고 있는 부서들 간의 통합적인 협조와 연구가 이루어져야 한다. 이를 위해서는 국가적인 차원에서 통계품질관리 및 통계품질 향상에 노력을 기울여야 할 것이다.

조사의 품질(survey quality)은 "Survey Measurement and Process Quality"라는 주제로 1995년 4월 1일부터 4일까지 영국의 Bristol에서 열린 국제회의를 계기로 큰 변화를 갖게 되었다. 오차보다는 품질에 그리고 조사의 모든 단계가 정보 생산의 한 과정으로 보아야 한다는 인식에 초점이 맞추어졌다(Collins와 Sykes, 1999). 그 후 1998년 5월에는 유럽(EUROSTAT)의 통계청장(국가의 통계를 담당하는 기구의 장)들은 "Quality Works and Quality Assurance within Statistics"라는 주제로 모임을 가졌다. 여기에서 적절성(relevance), 전문성(professionalism), 그리고 공개성(openness)을 강조하는 몇 가지 원칙들이 UN에서 승인된 공식통계의 10대 기본원칙¹⁾에 포함되었다(Brackstone, 1999).

통상적으로 통계품질 사용자들의 사용적합성으로 정의한다. 이동명과 김설희(2002)에 의하면, 우리나라에서는 통계품질을 결정하는 요소들로 정확성(accuracy), 시의성(timeliness)/정시성

1) <http://unstats.un.org/unsd/goodprac/bpabout.asp>

(punctuality), 관련성(이후에는 적절성으로 사용, relevance), 접근성(accessibility)/편리성(convenience), 비교성(comparability), 그리고 효율성(efficiency)을 사용하고 있다. 통계청에서는 2002년에 통계청에서 생산하고 있는 12개의 통계에 대한 품질평가에 이들을 사용하였다. 국제기구와 여러 나라들은 통계품질을 평가하는 요소들을 확대 또는 축소하여 정의하기도 하지만 대체적으로 우리나라가 사용하고 있는 6개 차원과 유사하다.

<표 1-1> 기관별 승인통계 작성현황

(2003. 8. 1. 현재)

(단위 : 기관, 종)

구 분	작성 기관수	계	승인통계				
			통계종류		작성방법		
			지정통계	일반통계	조사통계	보고통계	가공통계
계	130	446	85	361	223	173	50
정부기관	61	334	72	262	144	149	41
-중앙행정기관	29	242	56	186	110	117	15
(통계청)	(1)	(52)	(33)	(19)	(42)	(1)	(9)
-지방자치단체	32	92	16	76	34	32	26
지정기관	69	112	13	99	79	24	9

최근에 통계청은 통계품질향상을 위해 조직을 개편하고 자체에서 생산하고 있는 통계에 대한 품질평가를 실시하는 등의 노력을 기울이고 있으나, 이에 대한 연구는 아직 시작단계에 불과하다. 지속적으로 통계품질을 향상시키기 위해서는 현재 우리나라에서 사용하고 있는 방법들에 대한 검토가 필요하고, 통계 선진국가에서 사용하고 있는 통계품질 향상을 위한 다양한 기법들과 통계품질평가지표들에 대한 연구검토가 선행되어야 할 것이다.

본 연구에서는 다음과 같은 세 부분에 중점을 두었다.

첫째, 현행 통계청에서 통계품질의 향상을 도모하기 위해서 사용하고 있는 통계품질평가 시스템에 관한 연구 검토,

둘째, 통계 선진국에서 사용하고 있는 통계품질향상을 위한 다양한 통계품질평가 방법과 지표들에 대한 비교와 국내 적용가능성 검토,

셋째, 새로운 통계품질평가지표를 개발하고 효율적인 평가제도를 제안한다.

이러한 연구를 통해 우리나라의 통계품질이 향상될 수 있는 계기가 마련되리라 기대한다.

제 2장 현행 통계품질평가지표 분석

통계청은 자체에서 생산하고 있는 52종의 승인통계 중에서 40종(2년 이상의 주기 또는 부정기 통계 제외)에 대해 2002년부터 2004년까지(3년간) 품질평가를 계획, 실시하고 있다. 2002년도에는 12개 통계에 대해 평가를 실시한 바 있고, 2003년도에는 11개 통계에 대한 품질평가를 계획하고 있다. 2002년의 통계품질평가는 6개의 품질차원을 8개의 통계작성절차로 분류한 48개 지표에 대해 실시하였다.

고품질의 통계는 부가가치가 높은 제품이라 할 수 있다. 통계의 생산은 통상적으로 많은 단계를 거쳐서 생산되어 최종 이용자들의 욕구에 충족되어야 한다. 따라서 고품질의 통계는 전체 생산 과정에서는 물론 이용자들의 사용적합성을 만족시켜야 한다. 통계가 생산되는 여러 과정 중에서 하나의 과정(절차)이라도 효율적으로 관리되지 못한다면 결과적으로 저 품질의 통계가 생산된다.

품질관리의 중요성은 제2차 세계대전 이후부터 산업계에서 본격적으로 다루어졌고, 그 중요성은 더욱 높아지고 있어 많은 관심과 연구가 진행되고 있다. 품질의 정의는 “사용적합성(Juran)”으로부터, “사회에 끼치는 손실(Taguchi)”로 범위가 넓어지고 있고, 품질의 개념(philosophy)도 생산자 중심에서 소비자 중심으로 변화하고 있다. 최근에는 품질 향상을 위해서, 품질관리(SQC), 실험계획법, 공정관리, 그리고 샘플링기법, 전사적 품질관리(TQM) 등과 같은 다양한 방법들이 사용되고 있다. 한편 개별 국가 단위의 품질관리가 국제단위의 품질기준으로 확대되어 국제적 기구인 ISO(international organization for standardization)가 만들어지고, 이 기구는 1987년에 ISO9000시리즈라는 품질경영과 품질보증에 관한 국제규격시리즈를 제정하였다. 국내의 많은 공공기관과 기업에서는 국제 품질규격인 ISO9000시리즈를 획득하여 국제품질인증을 받으려는 노력을 계속하고 있다.

최근에는 품질향상이 기업의 생사를 좌우하는 것으로 인식되고 있어 전사적인 품질경영시스템이 도입 운영되고 있다. 이러한 현상이 최신 경영기법인 “6시그마”(경영)전략으로 발전하였고 세계유수의 많은 기업에서 이를 시행하고 있다.

2.1 통계품질 결정요인

캐나다, 영국 등 몇몇 나라에서는 통계(정보) 품질에 대한 중요성을 80년대 중반부터 인식하기 시작하였고, 90년대에는 본격적으로 통계품질 향상을 위한 가이드라인을 만들고 지표를 개발하는 등의 노력을 기울여왔다. 2000년에 들어서면서 우리나라에서도 통계품질의 중요성이 부각되기 시작하였다. 국가의 경쟁력과 경제력이 증가하면서 이러한 추세는 극히 자연스러운 것이라 할 수 있다.

통계품질은 제품의 경우와 마찬가지로 사용적합성으로 정의할 수 있다. 각 국에서는 통계 품질을 결정하는 요인들을 설정하여 통계품질관리에 사용하고 있다. 이 절에서는 우리나라에서 사용하고 있는 통계품질의 결정요인을 살펴보고자 한다.

정확성(accuracy) ;

정확성은 통계품질을 결정하는 여러 요인들 중에서 가장 핵심이 되는 것으로, 측정하고자 하는 모집단의 참값을 추정값이 얼마나 정확하게 나타내는가의 정도를 의미한다. 이는 추정에 있어서 오차라는 것으로 표시되는데, 오차는 편향(bias)과 분산(variance)으로 구성된다. 오차의 주요한 원천은 표본오차(sampling error)와 비표본오차(non-sampling error)로 구분된다. 비표본오차에는 포함오차(coverage error), 무응답오차(nonresponse error), 측정오차(measurement error), 처리오차(process error), 그리고 모형가정오차(model assumption error) 등이 있다. 따라서 정확성을 높이기 위해서는 이들 오차가 어느 정도인가를 측정하고 이들 오차를 줄이는 방법이 마련되어야 한다.

부정확한 통계를 사용하게 되면, 이들로부터 얻게 되는 결과는 진실을 심각하게 왜곡하게 되어 오히려 통계를 사용하지 않은 것보다 못하게 된다. 정확성을 높이기 위해서는 조사설계, 조사 실시, 그리고 평가에 각별한 주의를 기울여야 한다.

시의성(timeliness)/정시성(punctuality) ;

필요한 통계자료를 필요한 시점에 사용할 수 있도록 하는 것은 통계의 품질과 매우 관련이 높다. 통계자료의 시의성은 통계자료의 기준시점과 이용자들이 유용하게 사용할 시점과의 차이와 관계가 있다. 아무리 중요하고 정확한 통계라 하더라도 필요한 시점에 제공되지 못하거나 변화하는 사회를 시의적절하게 반영하지 못한다면 그 통계자료는 가치를 잃게 된다. 시의성은 측정하고자 하는 현상의 변화율, 측정의 빈도, 그리고 응답의 즉시성에 따라 변하게 된다.

시의성은 정확성이나 비용과 상반되는 측면이 있다. 정확성에 치중하다보면 시의성이 떨어지게 되고 시의성을 강조하면 정확성이 낮아지게 된다. 물론 비용의 경우도 마찬가지이다. 따라서 조사계획 시에 적절한 선택이 필요하다.

시의성은 이용자에게 의해서 바로 관측될 수 있다. 중요한 통계자료의 발표는 사전에 공고된 시점에 발표되어야 한다. 이는 이용자들의 사용계획에 도움을 줄 뿐 만 아니라 내부 규율을 제공하게 된다. 통계의 공표시기가 늦어지면, 이들 통계와 연관되는 다른 통계나 이를 바탕으로 한 다른 정책들이 모두 연기되는 등의 심각한 문제가 생긴다.

시의성은 컴퓨터와 같은 새로운 과학기기의 발달과 새로운 기법들의 개발로 많은 진전이 있었다. 하지만 지속적으로 현재의 통계생산 전 과정을 평가하고 새로운 기법들을 도입해야 통계의 품질을 높일 수 있게 된다.

접근성(accessibility)/편리성(convenience)

이용자들이 어떤 통계에 대해서 모르고, 그것을 찾아 낼 수 없고, 접근할 수 없으며, 또한 사용하는 데 비용이 많이 들게 되면, 이미 그 통계는 이용자들에게 가치가 없게 된다. 그러므로 통계의 품질을 평가하는데 있어서 접근성은 이용자들이 그 통계의 존재를 알고, 손쉽게 찾을 수 있고, 그것을 자신의 목적에 이용하기가 쉬운 정도를 나타낸다. 따라서 폭넓은 보급체계와 정보 이용의 편리성 그리고 투명성은 접근성을 결정하는 중요한 요소가 된다.

근래에 들어서는 접근성의 문제는 컴퓨터와 인터넷을 얼마나 잘 활용하는가에 달려 있다. 이용자들의 이용부담을 줄여주는 것 또한 필요하다. 아무리 좋은 통계자료라 해도 이용자들이 손쉽게 사용할 수 없게 되면 사용적합성이 만족된다 할 수 없으므로 통계품질은 떨어지게 된다.

비교성(comparability)

통계자료는 지역 간, 국가 간, 그리고 시간과 같은 전체 시·공간상에서 서로 비교 가능할 때 보다 더 유용하게 된다. 대부분의 국가 통계들은 동일 목적으로 지속적으로 생산되고 있는 관계로 지역 간, 국가 간 그리고 시계열적 비교, 분석이 가능해야 그 활용도가 커진다. 따라서 이러한 성질을 통계 품질을 결정하는 하나의 요인으로 볼 수 있다.

효율성(efficiency)

효율성은 비용(cost)과 정도(precision)의 측면에서 평가될 수 있다. 통계작성에 필요한 예산을 효율적으로 편성하여 통계작성이 원활하게 수행되도록 하는 것이 우선적으로 필요하다. 투입된 비용에 비해 상대적으로 이용자들의 이용 빈도를 높게 하고, 통계의 판매비용을 높이는 것도 효율성을 증대시키는 것이다. 또한 조사방법, 자료수집 및 처리 방법 등의 변경이나 개선으로 인한 효율성의 증대도 통계품질에 영향을 준다.

적절성(relevance)

통계자료의 적절성(관련성이라는 용어로 현재 사용하고 있음)은 이용자들의 실제 요구를 충족시키는 정도를 나타낸다. 통계자료의 유용성은 이용자들의 가장 중요한 문제가 잘 반영되고 있는가에 달려있다. 적절성에 대한 평가는 이용자들의 변화하는 요구에 의존하는 주관적인 문제이다. 그리고 통계에 대한 적절성은 그 통계자료를 사용하는 이용자에게 따라서 달라질 수 있다. 따라서 통계 작성기관은 주어진 자원의 제약조건 하에서 가장 중요한 요구사항을 가능한 한 만족시키는 통계를 만들기 위해서 이용자들의 필요성과 충돌하는 점들을 조정해 주어야 한다. 또한 이용자들의 새로운 필요에 따른 통계(프로그램)들을 개발하고 이

를 위한 예산배정 등의 문제를 다루어야 한다. 그리고 통계작성기관은 현재의 프로그램(통계)들에 대한 적절성(현 통계들이 이용자들의 요구를 충족시키고 있는지 또는 시간의 변화에 따라 충족도가 떨어지는지, 또는 새로운 필요가 생기는지 등)을 지속적으로 모니터링(감시)하는 절차(단계)가 있어야 한다.

그밖에 통계품질을 결정하는 요인들로 국외의 기관들에서는 다음과 같은 것들이 사용되고 있다.

일관성(coherence)

통계자료의 일관성은 그것이 하나의 큰 해석 틀과 시간 내에서 다른 통계자료와 잘 결합되는지의 정도를 나타낸다. 표준적인 개념, 분류, 그리고 목표모집단의 사용은 여러 조사에서 공통된 방법의 사용과 마찬가지로 일관성을 증진시킨다.

일관성에는 같은 시점에 관한 서로 다른 자료들 간의 일관성, 서로 다른 시점에 대한 같은 자료들 간의 일관성, 그리고 국제적인 일관성을 모두 포함한다.

해석가능성(interpretability)

이용자들이 이해할 수 없거나 잘못 이해하기 쉬운 통계자료에는 가치를 부여할 수 없다. 그러므로 이용자들이 통계의 의미를 정확히 이해할 수 있도록 충분한 정보를 제공하는 것이 통계작성기관의 역할이다.

통계자료의 해석가능성은 자료를 적절히 해석하고 이용할 수 있는데 필요한 보조정보와 metadata(통계에 관한 정보, 또는 정보에 관한 정보)의 유용성을 나타낸다. 이러한 정보는 사용되는 근원적인 개념과 변수, 자료수집과 처리 방법, 그리고 통계자료의 정확성의 정도를 통상적으로 제공한다.

각 기관이나 국가마다 품질을 평가하는 차원은 다르지만 대체적으로 공통되는 부분이 많으며, 다른 차원으로 표기되어도 유사한 의미를 갖고 있거나 같은 차원으로 표기되었다 해도 포함하고 있는 내용이 다소 차이가 나는 경우가 있다. 따라서 비교를 위해서 각각의 기관에서 사용하고 있는 단어 표현의 차이를 세세하게 따지지 않고, 전체적인 틀에서 비교하였다. <표 2-1>에서는 국가별, 국제기구별로 통계품질평가에 사용되는 차원들을 비교하였다.

<표 2-1> 통계품질평가 차원별 비교

구분	한국 (i)	캐나다 (ii)	호주 (iii)	영국 (iv)	스웨덴 (v)	덴마크 (vi)	네덜란드 (vii)	Eurostat (viii)	OECD (ix)	IMF (x)
정확성 /신뢰성	○	○	○	○	○	○	○	○	○	○
적절성①	○	○	○	○		○	○	○	○	○
시의성 /정시성	○	○	○	○	○	○	○	○	○	○
접근성 /편리성	○	○	○	○	○	○		○	○	○
비교성	○			○	○			○		
일관성		○	○	○	○	○		○	○	○
효율성/ 비용효과	○						○			
해석가능성		○							○	
기타 1			개편크기 최소화	완결성 ②	내용		응답자 부담	완결성②		정직성③
기타 2			커버리지 포괄성							방법론적 충실성

(주)

① relevance를 관련성 대신 적절성으로 표현

② 완결성(completeness)은 적절성과 연관되고 통계들이 이용자들의 요구와 우선권이 반영될 수 있는 영역

③ integrity는 수집, 편집, 그리고 통계의 보급에 있어서 객관적인 원칙을 엄격히 준수. 여기서는 전문성, 투명성, 그리고 윤리적 기준을 다루므로 integrity를 무결성 대신 정직성으로 표현.

(참고문헌)

(i) 이동명, 김설희. 2002. “국가통계 품질측정을 위한 체계적 접근.” 한국조사연구학회 추계학술대회 발표논문집.

(ii) Gordon Brackstone (1999). Managing Data Quality in a Statistical Agency. *Survey Methodology*, Vol. 25, No. 2, pp139-149.

Statistics Canada(2002), Statistics Canada's Quality Assurance Framework-2002.

(iii) - Australian Bureau of Statistics (1996). Information Paper: Quality of Australian Balance of Payments statistics. Chapter 1. Data quality concepts. Statistics Concepts Library.

- Australian Bureau of Statistics (1998). Balance of Payments and International Investment Position, Australia, Concepts, Sources and Methods. Chapter 15. Data quality. Statistics Concepts Library.

(iv) National Statistics (2003). Draft Quality Measurement and Reporting Framework Consultation Document

[http://www.statistics.gov.uk/about/consultations/ons_consultations/default.asp].

- (v) - Elvers E. and Rosen B. (1999). Quality Concept for Official Statistics. Encyclopedia of Statistical Science, 621-629. New York: John Wiley.
 - International Monetary Fund (2001). Sweden: Report on Observance of Standards and Codes-Data Module; Response by the Authorities; and Detailed Assessments [http://dsbb.imf.org/vgn/images/pdfs/swerosc.pdf].
 - * 스웨덴의 경우, Availability and clarity를 편리성으로, Coherence and comparability를 일관성과 비교성으로 따로 분류하였다. 한편 내용(contents)에는 통계적 목표특성(목적, 모집단, 변수, 통계적 측도, 연구영역, 참조시간(reference time))과 포괄성(comprehensiveness)이 포함되어 있다.
- (vi) Carson S. C. (2000). What is Data Quality? A Distillation of Experience. International Monetary Fund.
- (vii) Willem de Vries and Richard van Brakel (1998). Quality Systems and Statistical Auditing. A Pragmatic Approach to Statistical Quality Management. *Netherlands Official Statistics*.
- (viii) - Werner G. and Hakan Linden (2001). Quality Measurement-Eurostat Experiences, Proceedings of Statistics Canada Symposium.
 - Eurostat (2000). Assessment of the Quality in Statistics.
- (ix) Organization for Economic Cooperation and Development (2002). Quality framework for OECD statistics [http://www.oecd.org/doc/M00029000/ M00029990.doc].
- (x) - International Monetary Fund (2003). Introduction to the data quality reference site [http://dsbb.imf.org/Applications/web/dqrs/dqrsintroduction/#reference].
 - International Monetary Fund (2003). Data quality assessment framework-Generic [http://dsbb.imf.org/vgn/images/pdfs/dqrs_generic.pdf]
 - * IMF는 유용성(serviceability)의 차원 내에 적절성, 적시성, 일관성 등의 부차원들을 포함한다.

2.2 통계품질 평가지표

<표 2-1>은 주요 국가나 국제기구에서 통계품질을 평가하는데 사용하고 있는 요인(차원)들을 보여주고 있다. 국가별, 국제기구별로 통계품질을 평가하는 차원이 약간씩은 다르나 전체적으로 볼 때 큰 차이는 없어 보인다. 다만, 유사한 항목별 구분이 명확하지 않고 중복되어 있어 다소간에 혼란이 있음을 알 수 있다.

본 절에서는 통계청에서 현재 사용하고 있는 통계품질평가지표를 절차별, 차원별로 살펴 보았다. 이는 통계품질을 측정하는 데 있어서 구체적으로 사용되는 세부 항목들을 검토하고 (중복, 누락, 또는 분류가 적절치 못한 것), 우리나라에서 사용하고 있는 지표들과 외국의 경우(4장 참조)를 비교하여 현재의 세부평가지표들이 적절한지를 분석하기 위한 것이다. 또한 이를 바탕으로 5장에서 새로운 통계품질평가지표를 개발하는데 참고하고자 한다.

<표 2-2> 대한민국 통계청(절차별)

절 차	세 부 지 표
1. 조사기획	1-1. 조사목적 명확성 1-2. 문서화의 충실성 1-3. 국제적 정의 준수 1-4. 적정 예산 확보 1-5. 이용자 요구사항 파악 및 반영 1-6. 기획자의 직무 교육 1-7. 현장조사 업무량 파악 1-8. 개편 작업의 시의성 1-9. 개편 작업의 공개성
2. 모집단 및 표본	2-1. 모집단의 명확한 정의 2-2. 적절한 표본 추출 2-3. 표본의 일관성 유지 2-4. 적절한 표본 규모 2-5. 조사대상처 명부 관리 2-6. 표본오차 공개 2-7. 표본추출정보 공개
3. 조사표 설계	3-1. 조사표 표현 및 흐름의 적절성 3-2. 충분한 용어 설명 3-3. 사전조사 실시 3-4. 조사표 및 입력프로그램의 친화성 3-5. 지침서 충실성
4. 조사직원 관리	4-1. 체계적인 조사직원 교육프로그램 4-2. 주기적인 조사직원 교육 실시 4-3. 조사직원의 업무전문지식정도 파악
5. 조사실시	5-1. 체계적인 현장관리 5-2. 실제조사방법 파악 5-3. 조사시 문제점 해결 및 해결방법 공유 5-4. 현장조사에 대한 점검 실시 5-5. 현장조사의 무응답 처리 지침
6. 자료처리 및 집계	6-1. 조사 무응답 실태 파악 및 분석 6-2. 입력자료의 내용검사 시스템 적절성 6-3. 모수 추정 방법의 적절성 6-4. 다른 통계와의 비교 분석
7. 자료공표	7-1. 간행물 사전 점검 체계 7-2. 설명자료에 대한 충실성 7-3. 자료 공표 시의성 7-4. 자료 접근 동시성 7-5. 공표 사전예고 7-6. 사전예고 준수
8. 자료이용	8-1. 자료 제공매체 다양성 8-2. DB 구축 8-3. 공표시기로부터 DB개재시기까지의 소요시간 8-4. 조사후 공표되지 않는 항목 비율 8-5. 자료의 국제적 비교성 8-6. 자료의 시계열 비교성 8-7. 원시자료 이용 8-8. 응답자 비밀 보호 8-9. 자료 문의처 제공

<표 2-3> 대한민국 통계청(차원별)

차원	세부지표
정확성	1-2. 문서화의 충실성 1-3. 국제적 정의 준수 1-7. 현장조사 업무량 파악 2-1. 모집단의 명확한 정의 2-2. 적절한 표본 추출 2-3. 표본의 일관성 유지 2-5. 조사대상처 명부 관리 3-1. 조사표 표현 및 흐름의 적절성 3-2. 충분한 용어 설명 3-3. 사전조사 실시 3-4. 조사표 및 입력프로그램의 친화성 3-5. 지침서 충실성 4-1. 체계적인 조사직원 교육프로그램 4-2. 주기적인 조사직원 교육 실시 4-3. 조사직원의 업무전문지식정도 파악 5-1. 체계적인 현장관리 5-2. 실제조사방법 파악 5-3. 조사시 문제점 해결 및 해결방법 공유 5-4. 현장조사에 대한 점검 실시 5-5. 현장조사의 무응답 처리 지침 6-1. 조사 무응답 실태 파악 및 분석 6-2. 입력자료의 내용검사 시스템 적절성 6-3. 모수 추정 방법의 적절성 7-1. 간행물 사전 점검 체계 8-8. 응답자 비밀 보호
적절성 (관련성)	1-1. 조사목적 명확성 1-5. 이용자 요구사항 파악 및 반영 1-6. 기획자의 직무 교육
시의성	1-8. 개편 작업의 시의성 7-3. 자료 공표 시의성 7-5. 공표 사전 예고 7-6. 사전 예고 준수
접근성	1-9. 개편 작업의 공개성 2-6. 표본오차 공개 2-7. 표본추출 정보 공개 7-2. 설명자료에 대한 충실성 7-4. 자료접근 동시성 8-1. 자료제공매체 다양성 8-2. DB 구축 8-3. 공표시기로부터 DB 게재시기까지의 소요시간 8-7. 원시자료 이용 8-9. 자료 문의처 제공
비교성	6-4. 다른 통계와의 비교 분석 8-5. 자료의 국제적 비교성 8-6. 자료의 시계열 비교성
효율성	1-4. 적정예산확보 2-4. 적절한 표본 규모 8-4. 조사후 공표되지 않는 항목 비율

제 3장 2002년도 통계품질평가결과 분석

이 장에서 다룬 통계품질평가결과는 2002년도에 실시된 12개 통계 전체에 대한 분석이다. 통계별 분석을 위해서는 다수의 평가자들에 의한 자료가 요구되나 2002년도에는 통계별로 외부와 내부 평가자가 각 1인이었기 때문에 유의한 통계적 분석은 곤란하여 개별 통계에 대한 결과분석은 실시하지 않았다. 결과분석은 48개 지표 중 <8-3.조사결과 자료공표 시기로부터 DB계재시기까지의 소요시간>과 <8-4. 조사 후 공표하지 않은 항목 비율>을 제외한 46개 지표에 대한 것이다.

3.1 자체 및 외부 평가결과 분석

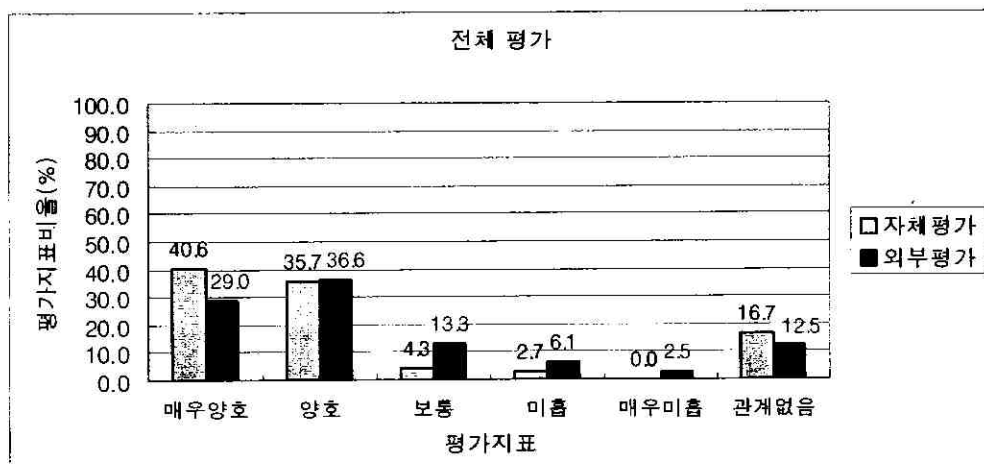
3.1.1 결과 요약

(1) 외부평가와 자체평가 사이에 절차별 혹은 차원별 평균점수에 있어서 통계적으로 유의한 차이가 존재한다.

(2) 자체평가자와 외부평가자 간에는 근본적인 평가의 시각차가 존재한다. 예를 들어, 자체평가에서는 현장조사 관련 부분에서 상대적으로 낮은 점수를 부여한 반면 외부평가자는 높은 점수를 부여하고, 조사기획과 설계 등의 절차에 낮은 평가를 하였다. 또한 '매우 미흡'에 대한 판정에서 내·외부평가자의 차이 역시 큰 것으로 나타났다. 일부 항목에서는 '관계없음'에 대한 응답에도 큰 차이를 보이고 있다.

(3) 품질지표(설문 문항)가 절차별로 심각하게 편중되어 있고, 일반적으로 잘 알려진 응답율, 대체율 등의 지표가 전혀 반영되어 있지 않다.

<그림 3-1> 12개 통계에 대한 평가방법에 따른 응답 현황



3.1.2 제안

(1) 현재의 평가시스템으로는 정확한 평가와 품질의 개선을 도모하기 어렵다. 장기적으로는 통계별로 내부평가시스템을 강화하여 지속적인 절차 및 차원별 품질평가를 시행하는 것이 바람직하다. 예를 들어, 외부평가와 내부평가에서는 '공표사전예고'와 같이 명백한 문항에서도 응답의 결과가 큰 차이(자체평가 100% '매우 양호', 외부평가는 '매우 양호'가 45.5% 그리고 '양호'가 18.2%)를 보이고 있음은 평가의 신뢰성을 의심케 한다.

(2) 기본적으로 평가는 개별 통계들의 품질개선 성과를 기준으로 하는 것이 바람직하다. 즉, 통계별로 개선이 필요한 항목을 선정하여 개선 목표를 설정하고, 집중적으로 품질개선 노력을 실행 한 후에, 그 성과에 대한 평가를 실시할 것을 제안한다. 각 통계를 비교 평가할 때는 개별 통계의 특성을 반영하여 차원별 혹은 품질평가문항별로 가중치를 부여할 수도 있다.

(3) 현재 사용하고 있는 차원과 절차에 따른 품질지표(문항)의 배분과 품질지표(문항)에 대한 수정 및 추가 그리고 변경이 필요하다. 즉, 차원에서는 '일관성(coherence)'의 추가를 검토해볼 필요가 있고, 절차에서는 8개 절차를 재조정하는 것이 바람직하다('조사직원관리와 조사실시' 절차를 '통계조사의 시행과 자료수집', '자료공표와 자료이용' 절차를 '통계이용자서비스'로 통합하고 '사후관리 및 품질평가' 절차를 새로이 추가한다. 덧붙여서 '모집단과 표본' 절차를 '표본설계'로, '자료처리 및 집계' 절차를 '통계조사결과의 분석' 절차로 명칭을 변경한다). 또한 품질지표(문항)는 중복되지 않고 명확하게 그리고 세분화하는 것이 효과적이다.

3.2 작성절차에 대한 분석

3.2.1 전체적인 분석 결과

아래 <표 3-1>로부터 <표 3-7>까지는 2002년 통계품질평가 자료를 근거로 얻어진 결과이다. 전체적으로 자체평가와 외부평가에는 작성절차별로 차이가 존재하는 것으로 나타났다. 구체적인 내용은 다음과 같다.

(1) 자체평가 결과 <자료공표>와 <자료이용> 그리고 <조사표 설계>가 상대적으로 높은 점수를 받았고, <조사실시>와 <조사직원 관리>가 낮은 점수가 나온 반면, 외부평가에서는 <조사실시>와 <조사직원 관리>가 높은 점수를 받았다.

(2) 특히 <조사기획>, <모집단과 표본> 그리고 <조사표 설계>는 자체평가에서는 상대적

으로 높은 점수를 받은 반면 외부평가에서는 낮은 점수를 받았다. 외부평가에서 <조사기획>, <모집단 및 표본> 그리고 <조사표설계>가 가장 낮은 평가를 받았다.

(3) 외부평가 결과 '매우 미흡' 항목이 존재하는 절차는 <조사기획>, <모집단 및 표본>, <조사표 설계>, <자료처리 및 집계>로 나타났다. 반면에 외부평가 결과가 자체평가 결과보다 높은 점수를 받은 항목은 주로 <조사직원 관리>와 <조사실시> 절차에서 나타났고, <자료처리 및 집계>에서도 두 항목이 높게 평가되었다.

<표 3-1> 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소	비율 (양호/미흡)
자체평가	4.4	0.28	0.063	7->8->3->1=2=5->6->4	4.8/3.9	76.3/2.7
외부평가	4.0	0.18		5->4=7->8->6->1=2=3	4.3/3.8	65.6/8.6

- * 상관계수는 자체평가와 외부평가의 절차에 따른 연관정도를 의미한다.
- * 순위는 절차별 평균점수의 상대적 비교이다.
- * 최대/최소는 각 평가에서 절차에 따른 평균점수의 최대값과 최소값을 의미한다.
- * 비율(양호/미흡)은 각 평가별 5점 척도('매우 양호', '양호', '보통', '미흡', '매우 미흡')를 3점 척도로 분류했을 때의 '양호'와 '미흡'을 의미한다. 즉, 이 때 '양호'는 5점 척도에서는 '매우 양호' + '양호'를 그리고 '미흡'은 5점 척도에서는 '매우 미흡' + '미흡'을 의미한다.

<표 3-2> 자체평가와 외부평가 평균점수 간의
유의성 검정(t-test에 의한 p-값)

	대응비교
단측검정	0.01088
양측검정	0.02176

<표 3-3> 3점 척도로 고려할 때, '미흡' 판정의 비율(%)

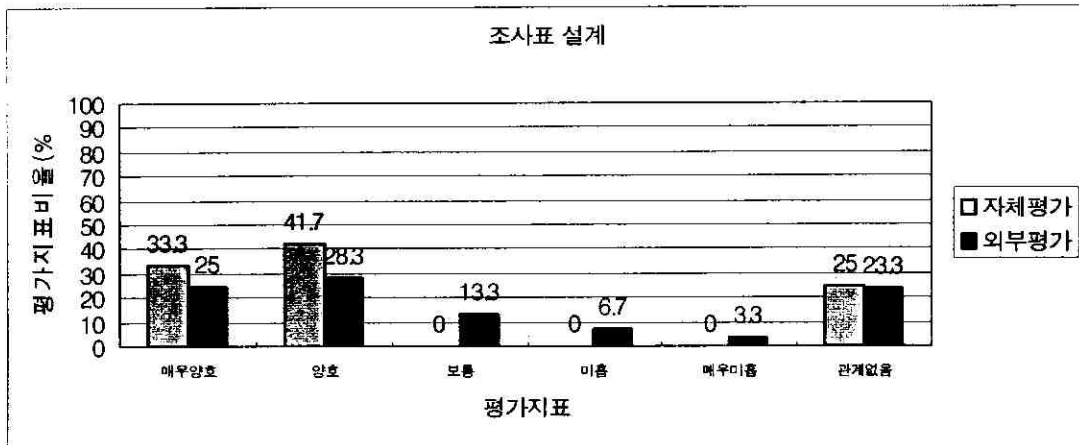
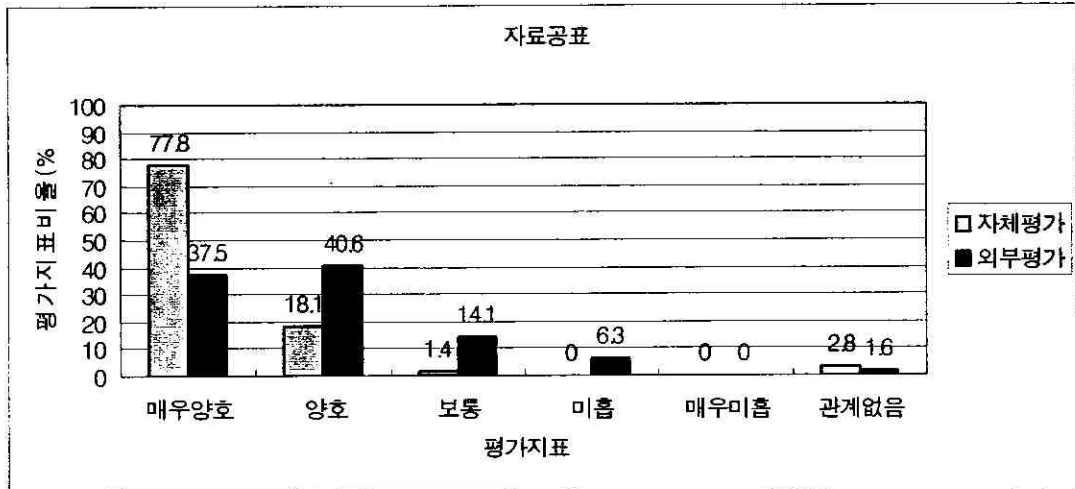
	1.조사 기획	2.모집단 및 표본	3.조사표 설계	4.조사직 원 관리	5.조사 실시	6.자료처리 및 집계	7.자료 공표	8.자료 이용
자체평가	2.8	7.1	0.0	0.0	0.0	6.3	0.0	3.6
외부평가	13	9.2	10.0	2.9	1.7	11.1	6.3	10.1

<표 3-4> 절차별 평균점수 차이(크기 순)

절차	7. 자료 공표	3. 조사표 설계	1. 조사 기획	2. 모집단 및 표본	8. 자료 이용	4. 조사원 관리	6. 자료처리 및 집계	5. 조사 실시
차이	0.7	0.6	0.5	0.5	0.5	0.2	0.1	0.0

* 평균점수 차이 = 자체평가 평균점수 - 외부평가 평균점수

<그림 3-2> 자료공표 절차와 조사표설계 절차에서의 평가방법에 따른 응답 현황



<표 3-5> 외부 평가자가 지적한 '매우 미흡' 과 '미흡' 항목

	'매우 미흡'	'미흡'
1. 조사기획	1-4. 적정 예산 확보 1-5. 이용자 요구사항 파악 및 반영 1-8. 개편 작업의 시의성 1-9. 개편 작업의 공개성	1-1. 조사목적 명확성 1-2. 문서화의 충실성 1-6. 기획자의 직무 교육 1-7. 현장조사 업무량 파악
2.모집단 및 표본	2-6. 표본오차 공개 2-7. 표본추출정보 공개	2-4. 적절한 표본 규모
3. 조사표 설계	3-2. 충분한 용어 설명 3-4. 조사표 및 입력프로그램의 친화성	3-1. 조사표 표현 및 흐름의 적절성 3-3. 사전조사 실시
4. 조사직원 관리	4-3. 조사직원의 업무 전문지식정도 파악	
5. 조사실시		5-3. 조사시 문제점 해결 및 해결방법 공유
6.자료처리 및 집계	6-1. 조사 무응답 실태 파악 및 분석 6-2. 입력자료의 내용검사 시스템 적절성 6-3. 모수 추정 방법의 적절성 6-4. 다른 통계와의 비교 분석	
7. 자료공표		7-2. 설명자료에 대한 충실성 7-5. 공표 사전예고
8. 자료이용		8-5. 자료의 국제적 비교성 8-6. 자료의 시계열 비교성 8-7. 원시자료 이용

<표 3-6> 외부평가가 자체평가보다 점수가 높은 항목

절차	항목
1. 조사기획	1-6. 기획자의 직무 교육
4. 조사직원관리	4-1. 체계적인 조사직원 교육프로그램 4-2. 주기적인 조사직원 교육 실시 4-3. 조사직원의 업무 전문지식정도 파악
5. 조사실시	5-3. 조사시 문제점 해결 및 해결방법 공유 5-4. 현장조사에 대한 점검 실시 5-5. 현장조사의 무응답 처리 지침
6. 자료처리 및 집계	6-1. 조사 무응답 실태 파악 및 분석 6-3. 모수 추정 방법의 적절성

<표 3-7> 평가별 평균점수 차이가 큰 항목(0.5이상)

차이	도수	문항
1.4	1	2-7.표본추출 정보 공개
1.2	1	7-5.공표 사전예고
1.1	1	1-4.적정예산확보
1.0	1	3-4.조사표 입력 프로그램의 친화성
0.9	2	1-8.개편작업의 시의성, 3-3.사전조사 실시
0.8	5	1-1.조사목적의 명확성, 1-2.문서와의 충실성, 3-2.충분한 용어 설명, 7-2.설명자료에 대한 충실성, 7-6.사전예고 준수
0.7	4	1-7.현장조사 업무량 파악, 8-5.자료의 국제적 비교성, 8-6.자료의 시계열 비교성, 8-8.원시자료의 이용
0.6	3	2-5.조사 대상처 명부관리, 7-3.자료공표 시의성, 8-8.응답자 비밀보호
0.5	2	1-5.이용자 요구사항 파악 및 반영, 1-9.개편작업의 공개성

[참고]모든 문항에서 자체평가 점수가 외부평가 점수보다 높다.

3.2.2 차이 발생의 원인

(1) 자체 평가자와 외부 평가자 간의 인식의 차이로 생각된다. 즉, 외부 평가자는 대개의 경우 학자들로서 기획, 설계, 추출방법 등 비교적 이론적인 부분에 치중하고, 이 부분에서 자체평가와 큰 차이를 보이고 있다. 조사실시와 조사원 관리는 크게 비중을 두지 않는 절차이기 때문으로 생각된다. 특히 이 절차들에 대한 평가는 전적으로 매뉴얼에 의한 평가가 될 수밖에 없기 때문에 실무자들의 자체평가 결과와는 당연히 많은 차이가 날 수 있다고 생각된다.

(2) 자체평가에서 상대적으로 높은 점수를 받은 절차와 낮은 점수를 받은 절차의 차이는 '관심'과 '의지' 그리고 '용기'에 의한 차이일 수 있다. 즉, 자체평가에서 상대적으로 높은 점수를 받은 절차는 주로 관례적으로 이미 틀이 잡혀 전해지고 있는 절차로서, 큰 문제가 없는 한 전임자가 만들고 무난히 수행해 온 절차를 특별히 개편할 이유를 찾지 못하기 때문으로 생각된다. 그러나 낮은 점수를 받은 절차는 직접적으로 조사가 이루어지는 절차로 항상 개선의 여지를 느끼고 아쉬워 할만한 부분이 많은 절차가 아닌가 생각된다.

3.3 차원별 분석

3.3.1 전체적인 분석결과

<표 3-8>과 <표 3-9>는 각각 차원별 문항분포의 내역과 차원 내에서의 문항별 평균점수의 차이를 보여 주고 있으며, 다음의 결과를 보여 준다.

(1) 차원별 문항수의 차이가 극심하고, 차원과 문항(지표)의 재검토가 필요하다.

정확성(25문항)>접근성(9문항)>시의성(4문항)>관련성=비교성(3문항)>효율성(2문항)

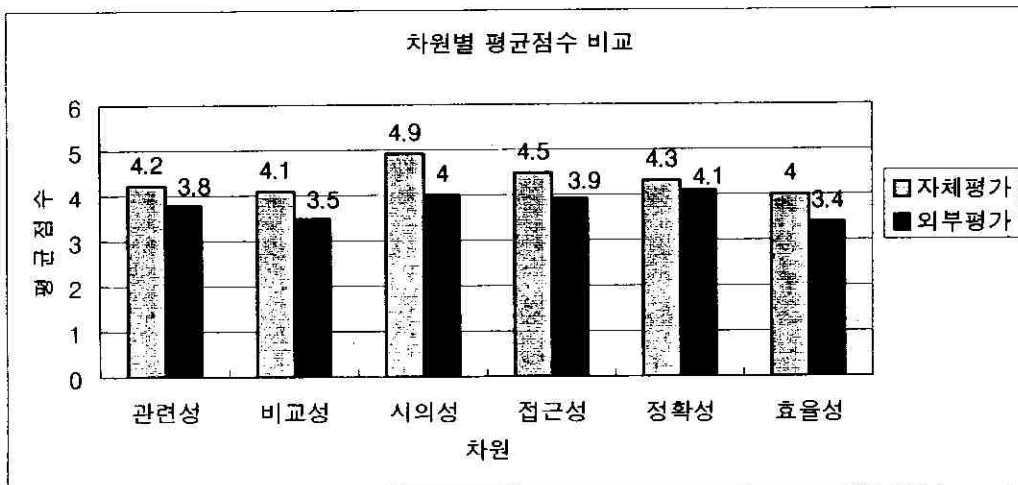
(2) 외부평가 결과 <관련성>, <비교성>, <효율성>은 모든 항목에서 '미흡' 또는 '매우 미흡'이 나타나고 있다. 특히 <효율성>관련 항목은 자체평가에서도 모두 '미흡'으로 나타났다.

(3) t-검정(대응비교-단측과 양측) 결과 p-value는 각각 0.00112와 0.00224로 외부평가의 평점과 자체평가의 평점간에는 지극히 유의한 차이를 보이고 있다.

(4) 평균점수는 자체평가는 모든 차원이 4.0이상이지만, 외부평가에서는 모든 차원의 점수가 4.1이하이다. 순위는 다음과 같다.

자체평가	시의성>접근성>정확성>관련성>비교성>효율성
외부평가	정확성>시의성>접근성>관련성>비교성>효율성

<그림 3-3> 차원 및 평가방법별 평균점수



(5) 평균점수의 차는 작게는 0.2점에서부터 0.9점까지 나타나고 있다. 평균점수 차이의 순서는 다음과 같다. (자체평가 평균점수 - 외부평가 평균점수)

시의성(4.9-4.0) > 접근성(4.5-3.9) = 효율성(4.0-3.4)
 > 관련성(4.2-3.8) = 비교성(4.1-3.5) > 정확성(4.3-4.1)

<표 3-8> 차원별 문항 분포 내역

차원	문항수 (미흡-매우 미흡)	내 용
정확성	25 (5-6)	<조사기획>1-2. 문서화의 충실성 1-3. 국제적 정의 준수 1-7. 현장조사 업무량 파악 <모집단과 표본>2-1. 모집단의 명확한 정의 2-2. 적절한 표본 추출 2-3. 표본의 일관성 유지 2-5. 조사대상처 명부 관리 <조사표 설계>3-1. 조사표 표현 및 흐름의 적절성 3-2. 충분한 용어 설명 3-3. 사전조사 실시 3-4. 조사표 및 입력프로그램의 친화성 3-5. 지침서 충실성 <조사원 관리>4-1. 체계적인 조사직원 교육프로그램 4-2. 주기적인 조사직원 교육 실시 4-3. 조사직원의 업무전문지식정도 파악 <조사실시>5-1. 체계적인 현장관리 5-2. 실제조사방법 파악 5-3. 조사시 문제점 해결 및 해결방법 공유 5-4. 현장조사에 대한 점검 실시 5-5. 현장조사의 무응답 처리 지침 <자료처리 및 집계>6-1. 조사 무응답 실태 파악 및 분석 6-2. 입력자료의 내용 검사 시스템 적절성 6-3. 모수 추정 방법의 적절성 <자료공표>7-1. 간행물 사전 점검 체계 <자료이용>8-8. 응답자 비밀 보호
접근성	9 (2-3)	<조사기획>1-9. 개편 작업의 공개성 <모집단과 표본>2-6. 표본오차 공개 2-7. 표본추출 정보 공개 <자료공표>7-2. 설명자료에 대한 충실성 7-4. 자료 접근 동시성 <자료이용>8-1. 자료 제공매체 다양성 8-2. DB 구축 8-7. 원시자료 이용 8-9. 자료 문의처 제공
시의성	4 (1-1)	<조사기획>1-8. 개편 작업의 시의성 <자료공표>7-3. 자료 공표 시의성 7-5. 공표 사전 예고 7-6. 사전 예고 준수
관련성	3 (3-0)	<조사기획>1-1. 조사목적 명확성 1-5. 이용자 요구사항 파악 및 반영 1-6. 기획자의 직무 교육
비교성	3 (2-1)	<자료처리 및 집계>6-4. 다른 통계와의 비교 분석 <자료이용>8-5. 자료의 국제적 비교성 8-6. 자료의 시계열 비교성
효율성	2 (1-1)	<조사기획>1-4. 적정 예산 확보 <모집단과 표본>2-4. 적절한 표본 규모

[설명] 1. 이탤릭체-외부평가 결과 “미흡”이 있는 문항, “매우 미흡”이 있는 문항은 이탤릭+ 밑줄
 2. 음영-자체평가 결과 “미흡”이 있는 문항

<표 3-9> 평가점수에 의한 비교

①정확성(25문항)

문	항	외부평가	자체평가	차이
소	계	4.1	4.3	0.2
<조사기획>				
1-2.	문서화의 충실성	3.9	4.7	0.8
1-3.	국제적 정의 준수	3.9	4.1	0.2
1-7.	현장조사 업무량 파악	3.9	4.6	0.7
<모집단 및 표본>				
2-1.	모집단의 명확한 정의	4.6	4.8	0.2
2-2.	적정한 표본 추출	4.1	4.4	0.3
2-3.	표본의 일관성 유지	3.9	4.2	0.3
2-5.	조사대상처 명부 관리	4.0	4.6	0.6
<조사표 설계>				
3-1.	조사표 표현 및 흐름의 적절성	3.9	4.3	0.4
3-2.	충분한 용어 설명	3.4	4.2	0.8
3-3.	사전조사 실시	3.5	4.4	0.9
3-4.	조사표 및 입력프로그램의 친화성	3.6	4.6	1.0
3-5.	지침서 충실성	4.7	4.7	0.0
<조사원 관리>				
4-1.	체계적인 조사직원 교육프로그램	4.2	3.9	0.3*
4-2.	주기적인 조사직원 교육 실시	4.3	4.2	0.1*
4-3.	조사직원의 업무전문지식정도 파악	3.9	3.6	0.3*
<조사실시>				
5-1.	체계적인 현장관리	4.3	4.3	0.0
5-2.	실제조사방법 파악	4.1	4.4	0.3
5-3.	조사시 문제점 해결 및 해결방법 공유	4.4	4.3	0.1*
5-4.	현장조사에 대한 점검 실시	4.4	4.3	0.1*
5-5.	현장조사의 무응답 처리 지침	4.2	4.0	0.2*
<자료처리 및 집계>				
6-1.	조사 무응답 실태 파악 및 분석	3.8	3.5	0.3*
6-2.	입력자료의 내용 검사 시스템 적절성	4.2	4.6	0.4
6-3.	모수 추정 방법의 적절성	3.8	3.7	0.1*
<자료공표>				
7-1.	간행물 사전 점검 체계	4.1	4.3	0.2
<자료이용>				
8-8.	응답자 비밀 보호	4.4	5.0	0.6

* 표시는 외부평가점수가 자체평가점수보다 높게 나타난 항목을 의미

②접근성(9문항)

문항	외부평가	자체평가	차이
소 계	3.9	4.5	0.6
<조사기획>			
1-9. 개편 작업의 공개성	3.5	4.0	0.5
<모집단 및 표본>			
2-6. 표본오차 공개	3.2	3.3	0.1
2-7. 표본추출 정보 공개	3.4	4.8	1.4
<자료공표>			
7-2. 설명자료에 대한 충실성	3.9	4.7	0.8
7-4. 자료 접근 동시성	4.5	4.9	0.4
<자료이용>			
8-1. 자료 제공매체 다양성	4.3	4.7	0.4
8-2. DB 구축	4.3	4.5	0.2
8-7. 원시자료 이용	3.2	3.9	0.7
8-9. 자료 문의처 제공	4.7	4.9	0.2

③시의성(4문항)

문항	외부평가	자체평가	차이
소 계	4.0	4.9	0.9
<조사기획>			
1-8. 개편 작업의 시의성	3.9	4.8	0.9
<자료공표>			
7-3. 자료 공표 시의성	4.3	4.8	0.5
7-5. 공표 사전 예고	3.8	5.0	1.2
7-6. 사전 예고 준수	4.1	4.9	0.8

④관련성(3문항)

문항	외부평가	자체평가	차이
소 계	3.8	4.2	0.4
<조사기획>			
1-1. 조사목적 명확성	4.0	4.8	0.8
1-5. 이용자 요구사항 파악 및 반영	3.5	4.0	0.5
1-6. 기획자의 직무 교육	3.9	3.7	0.2

⑤비교성(3문항)

문항	외부평가	자체평가	차이
소계	3.5	4.1	0.6
<자료처리 및 집계>			
6-4. 다른 통계와의 비교 분석	3.7	4.1	0.4
<자료이용>			
8-5. 자료의 국제적 비교성	3.8	3.8	0.0
8-6. 자료의 시계열 비교성	4.4	4.4	0.0

⑥효율성(2문항)

문항	외부평가	자체평가	차이
소계	3.4	4.0	0.6
<조사기획>			
1-4. 적정 예산 확보	3.2	4.3	1.1
<모집단과 표본>			
2-4. 적정한 표본 규모	3.7	3.7	0.0

3.4 기타 분석 결과

이 장에 수록되지 못한 분석 결과는 다음과 같으며, 이 내용은 [부록]에 첨부하였다.

- ① 절차별 세부 분석 결과
- ② 각 문항 및 평가방법에 따른 비교
- ③ 평가방법에 따른 평균 점수 차이에 대한 분석
: <자체평가 점수 - 외부평가 점수>에 따른 분석
- ④ 차원별 분석 결과(그래프 모음)
- ⑤ 기타 : '관계없음'에 대한 도수 차이

제 4장 외국 사례연구

4.1 국제통화기구(International Monetary Fund ; IMF)

국제통화기구는 국민계정(national accounts), 국제수지(balance of payments), 정부재정 통계(government finance statistics), 통화통계(monetary statistics), 소비자가격지수(consumer price index), 생산자가격지수(producer price index) 등의 거시 경제적 분석에 사용되는 자료들의 품질평가를 위해 자료품질평가체제(Data Quality Assessment Framework ; DQAF)를 마련하였다. 이 체계는 제도적 환경(institutional environments), 통계적 과정(statistical processes)과 통계 생산품(statistical products)의 특성을 파악하기 위하여 다음 6개 차원들을 이용하여 자료의 품질을 평가하도록 되어있다.

차원(Dimensions)	요소(Elements)	지표(Indicators)
0. 품질의 필수조건 (Prerequisites of quality)	0.1 합법적이고 제도적인 환경 (Legal and institutional environment) -통계에 대한 협력적인 환경	0.1.1 통계 수집, 처리와 배포에 대한 책임이 명시됨 0.1.2 자료생산기관 간의 자료 공유와 협력이 적절함 0.1.3 응답자 자료는 비밀이 유지되며 통계적 목적만을 위하여 사용됨 0.1.4 응답을 높이기 위해 통계적 보고는 합법적인 명령이나 조치를 통해 이루어짐
	0.2 자원(Resources) -재원들이 통계적 프로그램의 요구에 적절함	0.2.1 직원과 각종 자원은 통계적 프로그램에 적절함 0.2.2 자원의 효율적인 사용을 위한 조치들이 취해짐
	0.3 품질 인식 (Quality awareness) -품질이 통계적 작업의 초석임	0.3.1 통계 생산과정들은 품질에 초점을 맞추는데 적절함 0.3.2 통계 생산과정들은 통계의 수집, 처리와 배포에 관한 품질을 관리하는데 적절함 0.3.3 통계 생산과정들은 품질에 관한 다양한 요소들의 균형을 맞추는 것을 포함한 여러 가지 품질 문제들을 다루고 현재 또는 미래의 요구에 대한 계획 수립에 적절함
1. 정직성(Integrity) -통계의 수집, 처리와 배포에 있어 객관성의 원칙이	1.1 전문성(Professionalism) -통계적 정책과 실행은 전문적인 원칙을 따라야 함	1.1.1 통계는 편견 없이 만들어짐 1.1.2 소스들과 통계적 기법의 선택은 통계적 고려에 의해서만 이루어짐

만드시 지켜져야 함		1.1.3 적절한 통계는 잘못된 해석과 오용에 관한 설명이 첨부됨
	1.2 투명성(Transparency) -통계적 정책과 실행은 투명해야 함	1.2.1 통계가 수집되고 처리되고 배포될 때의 용어들과 조건들이 공개됨 1.2.2 통계 공개 이전에 내부적인 정치적 접근이 공개적으로 확인됨 1.2.3 통계 기관의 생산품이 위와 같이 명확히 확인됨 1.2.4 방법론, 소스 자료와 통계적 기법에 대한 주요 변동이 미리 예고됨
	1.3 도덕적 기준 (Ethical standards) -정책과 실행은 도덕적 기준을 따라야 함	1.3.1 직원 행동에 대한 지침이 적절하고 직원들에게 잘 알려짐
2. 방법론적 건전성 (Methodological soundness) - 통계에 대한 방법론적 기초는 국제적으로 인정되는 기준, 지침, 또는 좋은 방법을 따름	2.1 개념과 정의 (Concepts and definitions) -사용되는 개념과 정의는 국제적으로 인정되는 통계적 체제와 조화됨	2.1.1 개념과 정의에 관한 구성은 국제적으로 인정되는 기준, 지침 또는 좋은 방법을 따름
	2.2 범위(Scope) -범위는 국제적으로 인정되는 기준, 지침, 또는 좋은 방법과 조화됨	2.2.1 범위는 국제적으로 인정되는 기준, 지침 또는 좋은 방법과 광범위하게 일관성이 있음
	2.3 분류(Classification)/ 부문 (Sectorization) -분류와 부문체계는 국제적으로 인정되는 기준, 지침, 또는 좋은 방법과 조화됨	2.3.1 분류와 부문 체계는 국제적으로 인정되는 기준, 지침 또는 좋은 방법과 광범위하게 일관성이 있음
	2.4 기록을 위한 기본 (Basis for recording) -유출(flows)과 재고(stocks)는 국제적으로 인정되는 기준, 지침, 또는 좋은 방법에 따라 평가되고 기록됨	2.4.1 시장 가격은 유출과 재고를 평가하는데 사용됨 2.4.2 기록은 자연 증가식으로 이루어짐 2.4.3 총액식/순액식 절차(grossing/netting procedures)는 국제적으로 인정되는 기준, 지침 또는 좋은 방법과 광범위하게 일관성이 있음
3. 정확성과 신뢰성 (Accuracy and reliability) -원자료(source data)와 통계적 기법은 건전하며, 통계적 결과물들은 충분히 사실대로 기술	3.1 원자료 -원자료는 통계를 만들기 위한 적절한 기초를 제공함	3.1.1 원자료는 국가의 특정 조건들을 고려하는 포괄적인 자료 수집 프로그램으로부터 얻어짐 3.1.2 원자료는 요구되는 정의, 범위, 분류, 평가와 기록 시간에 합리적으로 근사함 3.1.3 원자료는 적시적임
	3.2 통계적 기법 -사용되는 통계적 기법은 건전한	3.2.1 자료 편집은 건전한 통계적 기법들을 사용

	통계적 절차를 따름	3.2.2 자료 조정, 변환과 통계적 분석과 같은 통계적 방법은 건전한 통계적 기법들을 사용
	3.3 원자료의 평가와 확인 -원자료는 정기적으로 평가되고 확인됨	3.3.1 센서스, 표본조사와 행정기록을 포함한 원자료는 포함범위(coverage), 표본오차, 응답오차, 비표본오차에 대해 정기적으로 평가됨. 평가 결과는 모니터되며 계획 수립에 이용됨
	3.4 중간자료(Intermediate data)와 통계적 결과물(statistical outputs)의 평가와 확인 -중간 결과와 통계적 결과물들은 정기적으로 평가되고 확인됨	3.4.1 주요 중간 자료는 적용 가능한 다른 정보를 가지고 확인됨 3.4.2 중간 자료 내에서 통계적 모순이 평가되고 조사됨 3.4.3 통계적 결과물 내에서 통계적 모순과 다른 가능한 문제들에 관한 지표들이 조사됨
	3.5 개정 연구(Revision studies) -신뢰성의 척도	3.5.1 개정에 대한 연구와 분석은 정기적으로 행해지고 통계적 과정들을 알리는데 사용됨
4.유용성(Serviceability) -통계는 적절하고 적시적이고 일관성이 있어야 하며 예상할 수 있는 개정 정책을 따름	4.1 적절성(Relevance) -통계는 연구 분야에 관한 적절한 정보를 다룸	4.1.1 이용자들의 요구에 부합하는 통계의 적절성과 실용성이 모니터 됨
	4.2 시의성(Timeliness)과 주기성(Periodicity) -시의성과 주기성은 국제적으로 승인되는 배포 기준을 따름	4.2.1 시의성은 배포 기준을 따름 4.2.2 주기성은 배포 기준을 따름
	4.3 일관성(Consistency) -통계는 데이터 세트 내에서, 시간에 따라 그리고 주요 데이터 세트와 일관성이 있음	4.3.1 통계는 데이터 세트 내에서 일관성이 있음 4.3.2 통계는 적절한 기간에 걸쳐 일관성이 있음 4.3.3 통계는 다른 자료 소스들이나 통계적 체계를 통해 얻어진 통계들과 일관성이 있음
	4.4 개정 정책과 실행(Revision policy and practice) -자료 개정은 규칙적이고 공표된 절차를 따라야 함	4.4.1 개정은 정기적이고 체계적이며 투명한 계획을 따름 4.4.2 예비 자료는 명확하게 확인됨 4.4.3 개정에 대한 연구와 분석은 공개됨
5.접근성(Accessibility) -자료와 메타데이터(metadata)는 쉽게 이용할 수 있어야 하고 이용자들에게 대한 지원이 적절함	5.1 자료 접근성(Data accessibility) -통계는 명료하고 이해할 수 있는 방법으로 나타내야 하고 배포의 형식은 적절하며 통계는 편견 없이 이용될 수 있음	5.1.1 통계는 적절한 해석과 의미 있는 비교가 용이함(본문, 표와 차트의 페이지 배정과 명확성) 5.1.2 배포 매체와 형식은 적절함 5.1.3 통계는 미리 공고된 계획대로 공개됨 5.1.4 통계는 모든 이용자들이 동시에 이용 가능

		5.1.5 공개될 수 있으나 공표되지 않은 통계는 요구하면 이용 가능함
	5.2 메타데이터 접근성 (Metadata accessibility) -최신의 적절한 메타데이터가 이용 가능함	5.2.1 공개되는 통계에 관한 개념, 범위, 분류, 기록의 기초, 자료의 소스와 통계적 기법 등을 소개하는 문서인 메타데이터가 이용 가능하며 국제적으로 인정되는 기준, 지침 또는 좋은 방법들과의 차이점들이 설명됨
	5.3 이용자들에 대한 지원 (Assistance to users) -즉각적이고 다양한 지원 서비스가 가능함	5.3.1 각 분야에 대한 담당자들이 공개됨 5.3.2 비용에 관한 정보를 포함하여 출판물, 문서와 다른 서비스들에 관한 카탈로그가 널리 이용 가능함

(참고문헌)

1. International Monetary Fund (2003). Introduction to the data quality reference site [<http://dsbb.imf.org/Applications/web/dqrs/dqrsintroduction/#reference>].
2. International Monetary Fund (2003). Data quality assessment framework-Generic [http://dsbb.imf.org/vgn/images/pdfs/dqrs_generic.pdf].

4.2 경제협력개발기구

(Organization for Economic Cooperation and Development ; OECD)

경제협력개발기구 통계의 품질 개선은 경제협력개발기구 통계 전략(OECD Statistics Strategy ; OSS)의 주요 사업들 중의 하나이다. 경제협력개발기구에서는 품질에 대한 정해진 틀은 없으나 기구 내의 통계전문가들은 현재 개인적인 차원에서 품질 개선을 위한 많은 노력을 하고 있다. 전사적 품질관리체제(Total Quality Management Framework; TQMF)에서 사용되고 있는 여러 가지 기법들이 이미 사용되고 있으며 몇 개의 프로젝트들이 2001년에 시작이 되었다. 또한 2002년 기구 내에서 진행된 품질개선 활동들에는 다음과 같은 것들이 있다.

- 긴급 품질 문제를 다루는 전문 회의 조직
- 자료 품질 문제와 방법에 관한 직원 세미나, 강습, 실연 등의 프로그램 구축
- 자료 수집, 관리와 배포에 대한 최적 방법의 선택과 개선
- 통계 생산품의 카탈로그 개발 및 품질 개선을 위한 통계적 지침의 시행

- 참고 시리즈(reference series)를 포함하는 경제협력개발기구의 데이터 웨어하우스 구축
- 경제협력개발기구 인트라넷과 인터넷에 '자료 품질' 사이트 개설

한편 경제협력개발기구는 통계를 체계적으로 평가하고 비교하며 개선시킬 수 있는 공통적인 품질체제(quality framework)를 갖추고 있다. 품질은 이용자 요구에 관한 '사용 적합성(fitness for use)'으로 정의되며 이 정의는 과거 통계 품질의 정확성(accuracy)보다 더 넓은 개념이다. 경제협력개발기구는 통계 품질을 다음 8개의 차원으로 분류한다.

차원	내용	OECD 환경
1. 적절성(Relevance) -자료에 대한 '가치'의 질적인 평가	<ul style="list-style-type: none"> •'가치'란 이용자들이 추구하는 목적에 있어 자료가 기여하는 정도임 •'가치'는 규정되는 주제들의 적용범위와 적절한 개념의 사용 여부에 의해 좌우됨 •적절성을 측정하기 위해서는 이용자 그룹과 그들의 요구를 확인해야 함 	<ul style="list-style-type: none"> •이용자들은 사무국, 위원회, 회원 정부와 다른 외부 이용자들을 포함함 •사무국과 위원회가 주요 이용자들임
2. 정확성(Accuracy) -조사하고자 하는 수량이나 특성들을 자료를 사용하여 올바르게 추정하거나 기술하는 정도	<ul style="list-style-type: none"> •정확성은 추정되는 값과 참값간의 근사성을 말함 •정확성은 많은 속성들을 가지며 이들 속성들은 개별적인 주요 원인들을 통해서 생기는 오차들에 관하여 측정되거나 기술됨 •표본조사를 기반으로 하는 추정치에 있어서 오차의 주요 원인들은 포함범위(coverage), 표본추출, 무응답, 응답, 처리 등이 있음 	<ul style="list-style-type: none"> •발행되는 자료의 정확성은 참여 기관들로부터 받은 자료의 정확성에 의해 주로 결정됨 •사무국의 여러 가지 활동들이 발행되는 자료의 전반적인 정확성에 영향을 미칠 수 있음
3. 확실성(Credibility) -단순히 자료 생산자의 이미지, 즉 브랜드 이미지를 근거로 이용자들이 갖는 신뢰	<ul style="list-style-type: none"> •이용자가 갖는 신뢰는 시간이 흐르면서 형성됨 •신뢰에 대한 한 가지 중요한 측면은 자료의 객관성이며 이는 자료가 적절한 통계적인 기준에 따라 전문적으로 생산되며 정책과 실행이 투명하다는 것을 의미함 •확실성은 생산과정의 정직성(integrity)에 의해 부분적으로 결정됨 	<ul style="list-style-type: none"> •회원국들로부터 받은 자료의 질이 낮아 경제협력개발기구의 전반적인 신뢰성에 영향을 미칠 것으로 판단되는 경우 사무국은 그 자료의 발행을 거절해야 함 •특정 자료의 수집에 관해 사무국과 회원국 간의 협정이 맺어진 경우 정치적 압력에 의해 자료 수집이 철회될 수 없음
4. 시의성(Timeliness) -통계 생산품이 실제로 가치가	<ul style="list-style-type: none"> •단기간 자료 또는 구조적 자료에 동등하게 적용되는 개념 	<ul style="list-style-type: none"> •경제협력개발기구에 의해 발행되는 자료의 시의성은 참여 기관

있고 적용될 수 있는 기간을 반영		<p>으로부터 받는 자료의 시의성에 의해 결정됨</p> <ul style="list-style-type: none"> •사무국 자체가 자료의 수집, 처리와 배포과정에서 발생할 수 있는 지체의 원인이 될 수 있음
<p>5. 정시성(Punctuality)</p> <p>-발행 계획에 맞추어 자료가 발행이 되는 정도를 반영</p>	<ul style="list-style-type: none"> •발행계획은 목표 발행 일자 또는 규정된 기간 내에 자료를 공개하는 것을 포함 	<ul style="list-style-type: none"> •발행 일정은 외부 이용자들이 통계를 적시에 사용할 수 있도록 함 •발행 일정은 내부 이용자들이 발행 일자에 근거하여 자신들의 일을 계획할 수 있는 능력을 높이게 함 •발행 일정은 사무국이 정치적인 이유로 발행 일정을 간섭하려는 압력에 굴하지 않는 역량을 높이게 함
<p>6. 접근성(Accessibility)</p> <p>-자료에 얼마나 쉽게 접근할 수 있는냐를 반영</p>	<ul style="list-style-type: none"> •접근성은 이용할 수 있는 자료 형식의 적합성, 보급 매체, 메타데이터(metadata)의 유용성, 이용자 지원 서비스 등을 포함함 	<ul style="list-style-type: none"> •내부 이용자와 외부 이용자 모두 접근 방법의 차이로 인하여 접근성에 대한 매우 다른 인식을 할 수 있음
<p>7. 해석가능성(Interpretability)</p> <p>-이용자가 자료를 이해하고 적절히 사용하며 분석하는데 있어서의 용이성을 반영</p>	<ul style="list-style-type: none"> •해석가능성의 정도는 자료와 관련한 개념, 목표모집단, 변수, 방법론에 대한 정의와 자료의 제약성을 기술하는 정보의 적합성에 의해 좌우됨 •정의와 절차에 관련한 메타데이터가 해석가능성을 높인데 도움이 되며 이러한 메타데이터의 일관성이 해석가능성의 한 측면이 됨 	<ul style="list-style-type: none"> •통계적 처리가 분산적인 모형을 따르는 상황을 고려하고 이용자들의 혼동을 막기 위해 회원국 간의 서로 다른 공표 메커니즘이 최소화되어야 함 •서로 다른 용도로 이용할 수 있는 정의들이 있는 경우 사무국은 이용자들이 자신들의 용도에 가장 적절한 것을 선택할 수 있도록 도와야 함
<p>8. 일관성(Coherence)</p> <p>-자료들이 어느 정도 논리적으로 연관되어 있고 상호 일관성이 있는지 정도를 반영</p>	<ul style="list-style-type: none"> •데이터세트 내에서의 일관성 -기본적인 자료 항목들은 모순이 되지 않는 개념, 정의, 분류들에 근거하며 의미 있게 결합될 수 있음 •데이터세트들 간의 일관성 -자료들은 공통적인 개념, 정의, 분류에 근거해야 하며 차이가 있을 경우 설명이 됨 •시간에 따른 일관성 -자료들은 시간에 따라 공통적인 개념, 정의, 분류에 근거해야 하며 차이가 있을 경우 설명이 됨 •국가간의 일관성 -자료들은 국가간에 공통적인 개 	<ul style="list-style-type: none"> •회원국 간의 일관성이 중요함 •시간에 따른 또는 국가 간의 개념 및 방법론에 있어서의 변동을 설명하는데 있어 메타데이터의 역할은 매우 기본적인 것임 •데이터세트 간 일관성의 결여가 설명되지 않는 경우 통계의 해석가능성과 확실성이 크게 낮아질 수 있음

	념, 정의, 분류에 근거해야 하며 차이가 있을 경우 설명이 됨	
--	------------------------------------	--

(참고문헌)

1. Organization for Economic Cooperation and Development (2002). Quality framework for OECD statistics [http://www.oecd.org/doc/M00029000/M0002990.doc].

4.3 Eurostat

1994년 Eurostat은 전 세계 통계 조직들의 경험과 유럽 연합(European Union)의 여러 국가 통계 기관들의 많은 지원을 바탕으로 품질 측정(quality measurement)에 대한 접근을 시작하였다. 현재 Eurostat은 통계품질관리를 위한 기본적인 구조를 가지고 있으며 통계품질의 정의 및 품질 보고서(quality reports)에 대한 통일된 기준체제를 마련하였다.

특히 Eurostat의 품질 보고서는 Eurostat의 품질 개념을 따르는 표준적인 보고 형식을 갖고 있다. 다음 표는 Eurostat이 채택한 품질 개념의 7가지 주요 구성 요소와 내부 품질 보고서(Internal Quality Reports; IQR)에 대한 품질 지표이다.

품질 구성 요소	내용	내부 품질 보고서에 대한 품질 지표
1. 통계적 개념의 적절성 (relevance of statistical concept) : 통계적 개념들이 이용자들의 요구를 만족시키는 경우 그 개념들은 적절한 것이므로 이용자들이 기대하는 개념들을 확인하는 것이 필수적임	•통계적 척도, 변수, 모집단, 단위, 도메인, 기준 시간 등과 같은 통계적 개념들은 이용자들이 필요로 하는 것이어야 함	•이용자 만족을 측정하는 통계적 생산품의 수와 비율
2. 추정값 정확성 (accuracy of estimates) : 추정값과 미지의 실제 모집단값 간의 근사성으로 정의	•정확성의 개념은 많은 통계기관들과 학계에서 광범위하게 연구되고 있는 품질의 개념임 •정확성은 표본 오차와 비표본 오차로 나누어지는 오차들에 대한 분석에 초점을 맞추는 것임 •비표본오차는 다음과 같이 분류됨	•변이계수를 계산하는 통계 생산품에 근거한 표본 조사의 수와 비율 •표본 조사에 근거한 통계에 대한 1개 중요 변수의 변이계수 •단위 무응답에 대한 정보를 제공하는 통계적 생산품의 수와 비율

	<ul style="list-style-type: none"> - 포함오차 (coverage errors) - 측정오차 (measurement errors) - 처리오차 (processing errors) - 무응답 오차 (non-response error) - 모형 가정 오차 (model assumption error) 	<ul style="list-style-type: none"> •단위 무응답율 •1개 중요 변수에 대한 항목 무응답율 •통계가 근거로 하는 서로 다른 소스(sources)의 수와 형태 •잠정결과와 확정결과의 차 •비표본 오차에 대한 정보를 제공하는 통계적 생산품의 수와 비율
<p>3. 결과물 배포의 시의성과 정시성 (timeliness and punctuality in disseminating results)</p> <p>: 대부분의 이용자들은 공개 횟수가 많으며 예정된 시간에 발행되는 최신 수치를 원함</p>	<ul style="list-style-type: none"> •이용자들은 일반적으로 통계 정보가 최단 기간 내에 생산되어 가능한 한 빨리 그 정보를 이용할 수 있기를 요구함 •통계 처리 기간을 최소화하기 위해서는 자료 수집, 편집, 교체, 추정과 배포 등의 전반적인 과정을 효율적으로 관리해야 함 •하나의 방안으로서 응답자들의 일부 표본들에 근거해서 잠정 추정치를 제시한 후에 수정된 결과를 공급할 수 있음 •정시성과 정확성은 서로 상반될 수 있는 개념으로서 통계 기관들은 두 가지 품질 요소에 균형을 맞추기 위해 노력해야 함 	<ul style="list-style-type: none"> •잠정결과물과 확정결과물과의 생산 시간 간격 •기준 기간(reference period)의 끝과 최초 발행 일자 간의 평균 시간 •적시에 공개되는 발행물의 비율
<p>4. 정보의 접근 가능성과 투명성 (accessibility and clarity of the information)</p> <p>: 통계적 자료는 이용자들이 쉽게 접근할 수 있고 이용자들이 바라는 형식으로 되어 있으며 적절하게 문서화되어 있을 때 가장 가치가 있음.</p>	<ul style="list-style-type: none"> •어떤 종류의 통계를 이용할 수 있는지 이용자들이 쉽게 알 수 있어야 함 •통계에 대한 접근이 편리해야 함 •통계에는 개념과 방법에 관한 필요한 정보들이 함께 제공되어야 함 •통계에 대한 분석에서는 통계적 생산품의 서비스 차원이 강조되어야 함 	<ul style="list-style-type: none"> •통계 배포 수단의 개수와 형태 •메타데이터 또는 방법론적 참조 사항들을 공개하는 통계적 생산품들의 비율
<p>5. 통계의 비교성 (comparability of statistics)</p> <p>: 주어진 특성에 대한 통계는 공간과 시간에 걸쳐 비교 할 수 있을 때 가장 유용함</p>	<ul style="list-style-type: none"> •유럽 수준에서의 통합적인 통계를 평가하기 위해서는 회원국 간의 동일 통계에 대한 비교가 가능해야 함 •통계 비교를 위하여 기준 개념과 측정 과정의 변경에 기인한 변화는 문서화되어야 하며 이들 변화의 효과가 평가되어야 함 	<ul style="list-style-type: none"> •시간에 따른 개념과 측정의 변동을 설명하고 그 변동 효과에 대한 양적인 추정치를 제공하는 통계적 생산품의 수와 비율 •개념과 측정에 있어서 유럽 표준과의 차이를 설명하고 그 변동 효과에 대한 양적인 추정치를 제공하는 통계적 생산품의 수

		와 비율 •지리적 비교에 따른 차이점 •도메인(부모집단) 비교에 따른 차이점
6. 일관성(Coherence) : 서로 다른 조사들로부터 얻어지는 통계들이라 할지라도 공통적인 정의, 분류, 방법 등을 사용하는 경우 일관성이 있어 이용자들이 통계들을 명확하게 관련시켜 이해할 수 있음	•다양한 조사들 또는 행정 자료들로부터 얻어지는 유사한 통계는 가능하면 그 차이점을 확인하여 수량화해야 함	•다른 도메인(부모집단)으로부터 얻어지는 통계와의 일관성에 대해 기술적이거나 양적인 정보를 제공하는 통계적 생산품의 수와 비율
7. 완전성(completeness) : 통계가 이용 가능한 도메인(부모집단)에 대해서는 유럽 통계 시스템의 이용자에게 의해 제기되는 요구가 반영되어야 함	•정치적 관여자들과 대중 매체들과 같은 전략적 이용자들은 자신들의 요구와 우선적인 일들을 정의하고, 유럽 수준에서 측정되어야 하는 개념들을 채택하거나 만들어 냄	•요구되는 통계 대비 제공되는 통계의 비율

(참고문헌)

1. Werner G. and Hakan Linden (2001). Quality Measurement-Eurostat Experiences, Proceedings of Statistics Canada Symposium.
2. Eurostat (2000). Assessment of the Quality in Statistics.

4.4 캐나다(Statistics Canada)

정보의 품질에 대한 신뢰는 통계 기관에 있어서 생존의 문제이다. 만약 그 정보가 의심을 받게 된다면 그 기관에 대한 신용도 의심을 받게 되고 신뢰할 수 있는 정보를 제공하는 독립적이고 객관적인 기관으로서의 평판은 손상된다. 그러므로 품질에 대한 관심은 국가통계 기관(National Statistical Office)을 운영하는데 있어 무엇보다도 중요하다. 또한 품질이란 익숙한 용어이지만 정의하기 쉽지 않는 개념이기도 하다. 고객의 입장에서 통계적 결과물에 대한 사용 적합성을 반영하는 여러 가지 측면들을 포함시키기 위해서 품질이 정의되어야 하며 이를 위하여 통계정보의 품질에 대한 다음과 같은 6가지 차원들을 설정한다.

차원	주요 내용	절차 및 방법
1. 적절성 (Relevance) : 통계 정보가 고객의 실제 요구에 맞는 정	•정보는 이용자들에게 있어 가장 중요한 문제들을 다루어야 함 •적절성에 대한 평가는 이용자들	•고객의 요구 파악 : 국가통계기관은 주요 이용자들의 현재와 미래 요구에 뒤지지 않도록 다음 메커니즘을

<p>도를 반영</p>	<p>의 다양한 요구에 의존하는 주관적인 것임</p> <ul style="list-style-type: none"> •국가통계기관은 주어진 인적, 물적 등의 제약 내에서 가장 중요한 요구와 이용자들의 만족을 극대화시키는 프로그램을 만들기 위하여 서로 다른 이용자들의 상반되는 요구들에 대해 비교 검토를 하고 균형을 맞추어야 함 	<p>구성</p> <ul style="list-style-type: none"> -통계적 프로그램의 정책과 우선 과제에 관한 조언을 제공하는 국가통계자문회 -주요 주제 영역에 있어서의 전문가 자문위원회 -주요 연방정부 부처간 특별 상호협정 -차관들 간의 정책 및 프로그램 결정에 수석 통계학자(Chief Statistician) 참여 -통계 정책에 관한 연방-지방 자문위원회와 지방 및 지역 정부들의 통계적 요구를 인지하기 위한 특별 주제들에 관한 보조위원회 -통계적 개발을 다루는 교육, 보건, 법무 영역에서의 특별 연방-지방 협정 -주요 산업 및 소규모 기업체 협회와의 회의 -개별 이용자 및 이용자 질의를 통한 피드백 <p>•프로그램 재고</p> <p>:다루어지는 주제들뿐만 아니라 생산되는 정보의 정확성 (accuracy)과 시의성 (timeliness)에 관하여 현재의 프로그램이 이용자들의 요구를 만족시키는지 평가하기 위하여 주기적인 프로그램의 검토가 필요하며 다음 평가 방법들이 사용될 수 있음</p> <ul style="list-style-type: none"> -위탁을 받은 독립적인 전문가가 이용자와 상담하고 프로그램에 관한 변경을 권고 -프로그램 자체적으로 피드백 정보를 주기적으로 수집하고 평가 -국가통계기관은 중앙에서 통계 프로그램의 다양한 요소를 포함하는 이용자 만족도 조사를 행하고 통계적 생산물의 판매액 또는 사용량과 인터넷 접속 횟수를 포함한 무료 통계와 서비스의 이용현황 파악 <p>•우선적 정보 결정</p> <p>: 이용자들의 모든 요구를 수용할 경우 항상 예산범위를 초과하므로, 서로 다른 이용자들 간의 다양한 요구를 저울질하기 위한 주요 방법으로 정규연간계획주기(regular annual planning cycle)를 활용</p>
<p>2. 정확성(Accuracy)</p> <p>-통계적 정보가 조사 가 계획된 현상들을 정확하게 기술하는</p>	<p>•정확성은 통계적 추정치와 관련하여 기술</p> <p>•정확성은 전통적으로 편향(bias)와 분산(variance)의 두 가지</p>	<p>•조사 설계(survey design)</p> <p>:예산과 응답자 부담(respondent burdens)과 같은 제약 하에서 정확성과 시의성 (timeliness)이 가장 적절하게 균형을 이루</p>

<p>정도를 말함</p>	<p>요소로 구분됨</p> <ul style="list-style-type: none"> •정확성은 포함범위(coverage), 표본추출, 무응답, 응답과 같은 오차의 원인들에 관하여 설명이 될 수 있음 	<p>도록 하는 것이 목적이며 조사 설계의 다음 8가지 주요 측면들이 고려되어야 함</p> <ul style="list-style-type: none"> -정확성, 비용, 시의성과 응답자 부담들이 전반적으로 균형을 이룰 수 있도록 하고 이러한 것들이 실제 확인될 수 있도록 함 -새로운 자료 수집을 최소화하기 위하여 현존 자료 또는 행정기록과 같은 대체자료의 확인 -질문지와 조사표의 적절한 사전조사 -사용하는 추출틀에 의한 목표모집단의 포함범위(coverage) 평가 -전반적으로 균형을 맞추되 표본추출 및 추정 방법과 정확성, 시의성, 비용, 응답자 부담 및 시간에 따른 자료의 비교 등에 관한 이들 방법들의 영향을 적절히 고려 -응답률을 높이고 무응답에 대해서는 재조사를 행하며 결측자료(missing data)를 다루는 적절한 방법들 사용 -자료를 수집하고 처리하는 모든 단계에 대해 품질을 확보하기 위한 방법 고려 -수정 또는 조정되는 자료의 내외부적 일관성 점검 <p>•조사설계 시행</p> <p>:좋은 조사설계는(예를 들어 품질 보증 과정을 통해) 실행 오차를 줄이기 위한 것이지만 이것이 항상 잘못될 수 있으므로 다음과 같은 두 가지 정보가 실행 단계에서 필요함</p> <ul style="list-style-type: none"> -시행 중 발생하는 모든 문제들을 실시간으로 확인하고 수정하기 위한 정보로서 이를 위해서는 조사가 진행되는 동안 문제들을 조정하고 수정하는데 필요한 정보를 관리자에게 제공하는 적시적인 정보시스템이 필요함 <p>-조사 후에 설계가 계획대로 진행이 되었는지, 설계 중 어떤 점들에 문제가 있었는지, 앞으로의 설계에 도움이 될 수 있는 어떤 교훈들이 있었는지 평가하는 정보</p> <p>•정확성 평가</p> <p>: 정확성을 측정하는데 있어서 흔히 조사가 진행되는 과정에서 기록되어야 하는 정보가 있으므로 조사 설계 단계에서 그에 대한 측정이 고려되어야 하며 모든 조사에서 반드시 고려되어야 하는 정확성 측정에 관한 4가지</p>
---------------	--	--

		<p>주요 영역은 다음과 같음</p> <ul style="list-style-type: none"> -전체 모집단 또는 주요 부모집단에 대한 목표모집단의 다음과 같은 포함범위(coverage) 평가 <ul style="list-style-type: none"> ◦ 목록틀(list frame)의 포함범위 ◦ 모집단의 목록을 만들기 위한 센서스의 포함범위 ◦ 지역표본조사(area sample survey)의 포함범위 -사용된 표본추출방법에 대한 표본오차의 평가 -무응답률 또는 대체된 추정치의 백분율 -조사 결과들에 대한 다른 중요한 정확성 또는 일관성 문제들
<p>3. 시의성 (timeliness)</p> <p>-통계정보에 대한 기준 시점(reference point)과 통계적 정보를 공개하는 일자 간의 간격을 말함</p>	<ul style="list-style-type: none"> •시의성은 전형적으로 정확성과 상반됨 •통계정보의 시의성은 적절성에 영향을 줌 	<ul style="list-style-type: none"> •시의성을 유지하기 위한 방법들 -주요 정보는 미리 알려진 일자에 공개 -보다 적시에 자료를 만들기 위해 일부 통계 프로그램에 대해서는 잠정자료를 우선 공개한 다음에 수정된 확정자료를 공개 -특별 조사와 새로운 조사에 대한 시의성은 조사 착수와 공개 일자 간의 경과 시간이 판단 기준이 됨 -주문형 자료를 제공하는 프로그램에 대한 시의성의 적절한 척도는 주문과 배달 간의 경과 시간임
<p>4. 접근성 (Accessibility)</p> <p>-통계정보가 국가통계기관으로부터 얻어질 수 있는 용이성을 말함</p>	<ul style="list-style-type: none"> •접근성은 통계정보에 접근할 수 있는 형식이나 매체의 적합성뿐만 아니라 정보의 존재 여부를 확인하는데 있어 용이성도 포함함 •정보에 대한 사용 비용 역시 일부 이용자들에게 접근성에 대한 하나의 관점이 될 수 있음 	<ul style="list-style-type: none"> •접근성 관리 : 이용자들이 알지 못하거나 접근할 수 없는 통계정보는 큰 가치가 없는 것이므로 이용자들이 정보에 접근이 용이하도록 다음 4가지 주요 측면들을 고려해야 함 -이용자들에게 어떠한 정보를 이용할 수 있고 그 정보를 얻기 위해서는 어떻게 해야 하는지를 알려주는 잘 색인화된 통합된 카탈로그 시스템이 필요 -보급 경로를 통해 이용자들에게 맞는 형식으로 정보를 제공하는 통합된 배달시스템이 필요 -각 통계 프로그램에 의해 통합적인 카탈로그 시스템 내에 있는 개별 프로그램의 통계정보의 범위가 정해지고 적절하게 배달 시스템이 사용될 수 있도록 함 -카탈로그와 배달시스템에 대한 유용성 여부와 이용자 만족도를 얻을 수 있는 수단이 있어야 함

		<p>•카탈로그 시스템 : 이용자들이 보다 효과적으로 시스템을 이용할 수 있도록 다음과 같은 요소들을 고려</p> <ul style="list-style-type: none"> -전통적인 인쇄형 카탈로그는 대부분 내용들이 최신 것들이 아니므로 통계 상품에 대한 온라인 카탈로그 시스템이 사용되어야 하며 통계정보의 특성을 알 수 있게 하는 메타데이터 베이스가 이 시스템에 연결되어야 함 -카탈로그 시스템에는 이용자들이 통계기관에서 사용되는 용어를 모르는 경우에도 정보를 검색할 수 있도록 컴퓨터의 기억정보 색인을 구성 <p>•국가통계기관 이외의 다른 조직들의 역할 : 공공기관 이나 국가통계기관 이외의 다른 조직들은 다음과 같은 측면에서 통계정보의 접근성을 개선시키는데 중요한 역할을 하며 이러한 조직들과 제휴할 수 있는 기회를 열어놓아야 함</p> <ul style="list-style-type: none"> -통계정보의 단순한 배포자 -국가통계기관의 정보와 다른 기관들의 정보를 결합시켜 정보의 가치를 높임 <p>•정보 가격 정책 : 통계정보를 이용하는데 드는 비용은 접근성의 측면에서 중요하므로 다음을 고려</p> <ul style="list-style-type: none"> -기본정보는 무료로 이용할 수 있도록 하는 반면 특정 생산품, 즉 특정 생산품보다 자세한 정보를 생산하는데 드는 비용은 회수되어야 함 <p>•체계적인 이용자 피드백 : 이용자들은 접근성 판단의 중요한 기준이며 다음과 같은 방법을 통해 피드백</p> <ul style="list-style-type: none"> -카탈로그와 배달 시스템의 다양한 요소들에 대한 자동화된 통계 이용 -특정 생산품, 서비스, 배달시스템에 대한 이용자 만족도 조사 -제안, 불평, 칭찬 등의 자발적인 이용자 피드백
<p>5. 해석가능성 (Interpretability) -통계정보를 적절하게 해석하고 이용하</p>	<p>•부가정보와 메타데이터는 개념, 변수, 분류 방법, 자료 수집 방법과 통계정보의 정확성에 대한 지표 등을 포함</p>	<p>•해석가능성 관리 : 이용자들이 이해할 수 없거나 쉽게 오해할 수 있는 통계정보는 가치가 없는 것임. 이용자들이 적절하게 통계정보를 해석할 수</p>

<p>는데 필요한 부가정보와 메타데이터(metadata)의 유용성을 반영함</p>		<p>있도록 충분한 정보를 제공하는 것이 국가통계기관의 책임이며 이를 위해 정보에 대한 정보로 불리는 메타정보 또는 메타데이터의 제공이 해석가능성 관리에 주요 관심사가 됨. 통계정보를 이해하는데 필요한 정보는 다음 3가지로 구분할 수 있음</p> <ul style="list-style-type: none"> -자료와 관련한 개념과 분류 -자료를 수집하고 만드는데 사용되는 방법 -자료의 정확성의 척도 <p>해석가능성을 관리하기 위해 필요한 3가지 요소는 다음과 같음</p> <ul style="list-style-type: none"> -이용자들에게 자료를 해석하는데 필요한 기본적인 정보를 제공하기 위한 정책 -국가통계기관의 자료 각각을 기술하는데 필요한 정보를 포함하는 통합적인 메타데이터베이스 -국가통계기관에 의한 자료의 직접적인 해석
<p>6. 일관성 (Coherence) -광범위한 분석적 체계 내에서 시간에 따른 다른 통계정보들과 함께 성공적으로 사용될 수 있는 정도를 반영함</p>	<ul style="list-style-type: none"> •공통적인 조사방법론, 표준적인 개념, 분류와 목표모집단을 사용함으로써 일관성을 높일 수 있음 •일관성이 반드시 완전한 수치적인 일치를 말하는 것은 아님 	<ul style="list-style-type: none"> •일관성 관리 : 통계자료의 일관성은 다음 세 가지를 말함 -동일 시점에서 서로 다른 자료 항목들 간의 일관성 -서로 다른 시점에 동일한 자료 항목 간의 일관성 -국제적인 일관성 <p>국가통계기관 내에서 일관성을 다루는 방법은 다음 3가지를 들 수 있음</p> <ul style="list-style-type: none"> -국가통계기관이 측정하는 모든 주제들에 대해 표준적인 체계, 개념, 변수와 분류의 개발과 이용 -측정되는 수량이 일관성 있는 방법으로 정의될 때에도 측정 과정에서 서로 다른 소스들 간에 모순이 없도록 함 -서로 다른 소스들로부터 얻어지는 자료들에 대한 비교와 통합

Note. 정확성, 시의성과 해석가능성은 각 통계 결과물의 특성들로 간주될 수 있는 반면에 적절성, 접근성과 일관성은 국가통계기관의 전체 결과물들에 대해 고려될 수 있는 특성들임.

(참고문헌)

1. Gordon Brackstone (1999). Managing Data Quality in a Statistical Agency. Survey Methodology, Vol. 25, No. 2, pp139-149.

4.5 영국(Office for National Statistics ; ONS)

공식통계(official statistics)의 이용자들이 공식통계의 장단점을 이해하고 그 통계들을 적절하게 이용하는 방법을 알기 위해서는 품질척도(quality measures)와 지표(indicators)를 살펴볼 필요가 있다. 또한 통계 생산자인 영국통계청(ONS) 스텝들과 국가통계기관들에서 일하는 다른 스텝들이 품질 개선을 관리하고 지속적인 품질 개선을 하도록 하기 위해서는 품질 측정(quality measurement)이 필요하다.

"체계적인 품질측정과 보고(systematic quality measurement and reporting)"는 영국통계청 품질전략 및 계획(ONS Quality Strategy and Plan)의 7가지 중점사항들 중 하나이며, 품질은 국가통계시행령(National Statistics Code of Practice)의 8개 원칙들 중 하나이다. 이 시행령은 국가통계기관의 지식(knowledge)은 정기적으로 발행되는 품질 척도에 의해 유지된다는 점을 명확히 밝히고 있다.

4.5.1 체제의 목적과 실행

모든 영국통계청 결과물들에 대한 체계적인 품질 측정 및 보고는 큰 사업으로서 단계적으로 행해질 것이다. 이를 위해서는 체계적인 품질 측정과 보고에 관한 체제(framework)를 우선적으로 마련해야 하며, 그 목적은 영국통계청 결과물들의 품질을 측정하고 보고하기 위한 기준을 제공하는 것이다. 이 체제는 영국통계청 내에서 기존에 행해진 사업과 품질에 관한 EUROSTAT의 사업을 참조하여 개발되었으며 최종적인 체제는 다음 내용들을 포함할 것이다.

- ① 품질과 그 차원들에 대한 개요
- ② 각 품질차원에 영향을 주는 주요 과정들에 대한 점검목록
- ③ 각 품질차원에 대한 품질척도와 지표

초안 체제(draft framework)에서는 일반적인 이용자 필요성에 중점을 둔 품질 척도와 지표가 상술되며 최종 체제는 영국통계청 결과물들 전반에 적합할 것이다. 앞으로 특정 결과물에 대한 품질 척도와 지표가 함께 개발될 것이다. 품질척도와 지표들 중 일부는 개별 추정치들에 대해서 사용될 것이며 다른 척도와 지표들은 결과물 수준에 적합할 것이다.

또한 이 초안 체제는 결과물들에 대한 품질 보고서(Quality Reports)를 개발하기 위한 기초를 마련하기 위하여 사용될 것이며 이들 보고서들은 사용하기 편리한 형식으로 품질에 관한 요약물 담을 것이다

결과물들 전반에 품질측정과 보고에 대한 이러한 체계적인 접근 방법을 적용시키는 일은

영국통계청 현대화 프로그램(ONS Modernization Program)의 일부로서 단계적으로 행해질 것이다. 또한 이 초안 체제의 다음 단계는 품질척도와 지표 각각에 대해 중요도(priority)를 부여하는 일이 될 것이다.

4.5.2 품질의 의미

품질의 개념은 흔히 이용자 요구에 관한 목적의 적합성(fitness of purpose)으로 설명된다. 과거 품질은 정확성(accuracy)과 시의성(timeliness)의 개념으로 생각했지만 현재의 품질은 훨씬 더 넓은 정의를 갖는 것으로 인식되고 있다. 통계가 정확하고 적시적이라 할지라도 이용자들이 쉽게 접근할 수 없고 적절한 개념에 근거하지 않는다면 좋은 품질을 갖는다고 생각할 수 없다. 영국통계청은 유럽통계시스템(European Statistical System)에서 정의되는 다음과 같은 품질차원을 채택하였다.

- ① 적절성(relevance)
- ② 정확성(accuracy)
- ③ 시의성(timeliness)과 정시성(punctuality)
- ④ 접근성(accessibility)과 투명성(clarity)
- ⑤ 비교성(comparability)
- ⑥ 일관성(coherence)
- ⑦ 완전성(completeness)

각 차원에 대한 적절한 부차원들(sub-dimensions)이 있으며 이 부차원들은 비즈니스 조사(business survey)를 위한 유럽통계 표준품질보고서(Eurostat Standard Quality Report)에 기초한 것이다.

영국통계청에서 마련한 차원, 부차원, 과정, 품질척도와 지표들은 다음과 같다.

차원과 부차원	적절한 과정	품질척도	품질지표
1. 적절성 (relevance) 1.1 목표와 목적 1.2 소스와 정의 1.3 결과물	<ul style="list-style-type: none"> •고객과 투자자 상담 또는 피드백 메커니즘 - 이용자 요구 이해, 적합한 발행물 개발, 이용자 요구에 맞는 개념과 측정을 정의, 계절적 조정과 다른 표준화 적용 •이용자 만족도 평가 	1.1 목표와 목적 <ul style="list-style-type: none"> •측정되는 통계적 개념과 이용자들이 관심을 갖는 개념 간의 차이에 대한 평가 	1.1 목적과 목표 <ul style="list-style-type: none"> •결과물 주요 이용자들의 기술 및 분류 •주요 이용자들의 요구와 결과물의 용도 기술 •이용자 만족도 조사 결과 기술 •적절성 결여에 대한 원인들 기술 •고객 피드백 또는 만족도 조사를 기초로 적절성을 개선시키기 위한

	<p>와 조사로부터의 피드백</p> <ul style="list-style-type: none"> •결과물과 소스의 검토 •결과물과 소스의 분석 •사업 계획과 우선 순위 결정 		<p>조치와 행동 기술</p> <p>1.2 소스와 정의</p> <ul style="list-style-type: none"> •주요 통계 개념들(통계적 측도, 변수, 모집단, 단위, 도메인, 기준 시간) 기술. 이들에 대한 자세한 내용을 볼 수 있도록 참고표가 달려 있고 이용 가능함 •자료 수집 목적, 자료의 장점과 단점 등을 포함한 주요 소스들 기술 <p>1.3 결과물</p> <ul style="list-style-type: none"> •계절적 조정과 연령/성별 표준화와 같이 결과물에 적용되는 표준화 기술
<p>2. 정확성 (accuracy)</p> <p>2.1 표본오차</p> <p>2.2 비표본오차</p> <p>2.2.1 포함오차</p> <p>2.2.2 무응답오차</p> <p>2.2.3 측정오차</p> <p>2.2.4 처리오차</p> <p>2.2.5 모형가정오차</p>	<p>2.1 표본 오차</p> <ul style="list-style-type: none"> •표본설계 •표본설계 시행 <p>2.2 비표본오차</p> <p>2.2.1 포함 오차</p> <ul style="list-style-type: none"> •표본추출틀 작성과 유지 <p>2.2.2 무응답오차</p> <ul style="list-style-type: none"> •응답 최대화 방법 •결측 또는 타당하지 않거나 일관성이 없는 응답에 대한 대체 (imputation) •단위 무응답에 대한 가중치 조정 <p>2.2.3 측정오차</p> <ul style="list-style-type: none"> •설문지 설계와 검증 •조사원 교육 •자료수집절차 검증 <p>2.2.4 처리오차</p> <ul style="list-style-type: none"> •설계 중 시스템의 점검과 검증 •자료포착시스템 (data capture system) •코딩 시스템 •편집과 확인 	<p>2.1 표본오차</p> <ul style="list-style-type: none"> •표본오차 <p>2.2 비표본오차</p> <p>2.2.1 조사범위 오차</p> <ul style="list-style-type: none"> •포함 오차 (coverage error) <p>2.2.2 무응답오차</p> <ul style="list-style-type: none"> •주요 추정치에 대한 무응답 편향(bias) <p>2.2.3 측정오차</p> <ul style="list-style-type: none"> •측정오차에 기인한 분산과 편향 <p>2.2.4 처리오차</p> <ul style="list-style-type: none"> •처리편향과 분산 <p>2.2.5 모형가정오차</p> <ul style="list-style-type: none"> •모형에 의한 편향과 분산 	<p>2.1 표본 오차</p> <ul style="list-style-type: none"> •주요 추정치의 표준오차 추정치 (또는 변이계수) •주요 추정치에 대한 설계효과 (design effects) •추정치에 표본오차가 있다는 언급 •오분류(misclassification), 무응답 등과 같은 요소를 포함한 분산 추정 방법의 기술 •정확한 표준오차가 제공되지 않는 추정치에 대한 표준오차 유도 또는 근사적 방법 기술 •표준오차 추정치에 대한 참고문헌 또는 링크 <p>2.2 비표본 오차</p> <ul style="list-style-type: none"> •추정치가 비표본 오차를 포함한다는 언급과 비표본 오차의 주요 원인 기술 •주요 추정치들에 대한 대체값들의 전체 기여 •주요 항목 대체율 <p>2.2.1 포함오차</p> <ul style="list-style-type: none"> •포함률(coverage rate)의 정의 및 오차의 원인 확인 •전체 모집단과 부차 모집단에 대한 목표 모집단과 연구모집단 정의 및 비교 •소스, 작성 일자와 마지막 갱신 일자를 포함한 표본추출틀에 관한

<ul style="list-style-type: none"> •품질 관리 •표준화 시스템 <p>2.2.5 모형 가정 오차</p> <ul style="list-style-type: none"> •추정과 가중치 부여 •이상치 확인과 처리 •특별 조정 		<p>사항들의 확인 및 기술</p> <ul style="list-style-type: none"> •과소포함(undercoverage), 중복(duplications), 부적격(ineligibility)과 오분류에 대한 비율 추정치 •주요 추정치에 대한 포함오차의 예상효과를 평가 •포함오차 문제들을 다루는 방법 기술 <p>2.2.2 무응답 오차</p> <ul style="list-style-type: none"> •오차의 원인으로 무응답 오차 정의 및 확인 •부차그룹(sub-groups)에 의한 전체 단위 응답률 •주요 항목 무응답률 •응답자와 응답거부자의 차이 기술 및 모집단 특성에 의한 예상 무응답 편향에 대한 평가 언급됨 •최종 추정치에 대한 무응답/대체의 예상효과 평가 <p>2.2.3 측정 오차</p> <ul style="list-style-type: none"> •측정 오차 정의 및 그 오차의 원인 확인 •설문지 개발과 조사원 훈련을 포함한 측정 오차를 줄이기 위한 과정 기술 •서로 다른 자료 수집 방법에 기인한 차이점들 평가 •편집오류율(edit failure rate) •편집률(editing rate) •주요 추정치들에 대한 편집된 값들의 전체 기여 •설문에 답하는데 어려움이 있는 응답자들의 비율 •비표준적인 기준시간 때문에 이용할 수 없는 자료에 기인하여 조정을 해야 하는 응답 비율 •대리응답률 •조사원 분산(interviewer variance) •방법에 기인한 차이점들의 평가 <p>2.2.4 처리 오차</p> <ul style="list-style-type: none"> •처리 오차의 정의 및 원인 확인
---	--	---

			<ul style="list-style-type: none"> •처리시스템과 품질관리 기술 •스캐닝과 입력 오차율 •코더들 간의 분산 (corder variance) <p>2.2.5 모형 가정 오차</p> <ul style="list-style-type: none"> • 사용된 모형과 추출 절차 기술
<p>3. 시의성 (timeliness)과 정시성 (punctuality)</p> <p>3.1 생산시간</p> <p>3.2 주기성</p> <p>3.3 정시성</p>		<ul style="list-style-type: none"> •시간 지체: 잠정 결과물의 발행 기준 일자/기간 •시간 지체: 최종 결과물의 발행 기준 일자/기간 •시간 지체: 자료 수집의 끝과 최초 결과의 발행 •시간 지체: 실제 발행일과 예정 발행일 •새로운 조사와 특별 조사에 대한 시간 지체: 조사 착수와 발행일 •현존하는 소스를 사용하는 주문형 자료 요청에 대한 시간 지체: 자료 요구일과 배달일 	<ul style="list-style-type: none"> •자료의 시의성에 대한 주요 이용자 필요성에 대한 기술 •발행 횟수 •발행 일자에 대한 계획표
<p>4. 접근성 (accessibility)과 정보의 투명성 (clarity)</p> <p>4.1 보급</p> <p>4.2 분석자의 요구</p> <p>4.3 정보를 찾기 위한 지원</p> <p>4.4 투명성</p>	<ul style="list-style-type: none"> •이용가능성 확보 •접근로의 개선과 홍보 		<p>4.1 보급</p> <ul style="list-style-type: none"> •통계는 Web 또는 다른 매체를 통하여 이용 가능함 •통계는 모든 이용자들에게 동시에 사전에 예고된 계획표대로 공개됨 <p>4.2 분석자의 요구</p> <ul style="list-style-type: none"> •비용을 포함한 일반인이 이용할 수 있는 마이크로데이터에 대한 자세한 이용 절차가 언급됨 •미공표 자료에 대한 단순 테이블을 얻기 위한 비용 및 평균 소요 시간을 포함한 자세한 이용 절차가 기술됨 <p>4.3 정보를 찾기 위한 지원</p> <ul style="list-style-type: none"> •Web 기반 자료 내에서 이용 가

			<p>능한 탐색 도구가 있음</p> <ul style="list-style-type: none"> •카탈로그나 다른 적절한 문서들에 결과물들에 관한 사항들이 표시됨 •결과물들은 검색 엔진을 이용하여 찾을 수 있음 •기술적인 정보를 포함한 다른 정보들이 이용 가능함 <p>4.4 투명성</p> <ul style="list-style-type: none"> •개념과 분류가 기술됨 •자료를 수집하고 결과물을 만드는 데 사용되는 방법론이 기술됨 •자료의 정확성의 측도가 이용 가능함 •원문, 그래프, 지도 등을 포함하는 자료의 설명이 있음
<p>5. 비교성 (comparability)</p> <p>5.1 시간에 따른 비교성</p> <p>5.2 지리적 비교성 : 국내적 국제적 비교</p> <p>5.3 도메인(부모 집단) 간의 비교성</p>	<p>5.2 지리적 비교성</p> <ul style="list-style-type: none"> •국제적으로 그리고 국내적으로 일치하는 정의와 방법 사용 	<p>5.1 시간에 따른 비교성</p> <ul style="list-style-type: none"> •변화에 기인한 불연속성의 추정 	<p>5.1 시간에 따른 연속성</p> <ul style="list-style-type: none"> •연속 발행물 내에서 불연속 횡수 •모든 불연속 발행은 예고됨 •시간에 따른 개념, 정의, 분류, 방법의 변동의 영향에 대한 기술과 평가 <p>5.2 지리적 비교성</p> <ul style="list-style-type: none"> •국내적/국제적으로 일치하는 정의와 기준을 기술 •적용하는 국제적 규약 기술 •국내적/국제적 정의에 벗어나는 것과 그 이유를 기술 <p>5.3도메인 간의 비교성</p> <ul style="list-style-type: none"> •도메인 간의 차이 기술
<p>6. 일관성 (coherence)</p> <p>6.1 잠정 통계와 최종 통계 간의 일관성</p> <p>6.2 서로 다른 소스들로부터 얻어진 유사한 추정</p>	<ul style="list-style-type: none"> •일관성 있는 시스템의 사용 •조화된 정의와 질문의 사용 •방법론적 기준 사용 	<p>6.1 잠정통계와 확정통계 간의 일관성</p> <ul style="list-style-type: none"> •규칙적으로 수정되는 결과물들에 대해서 잠정추정치와 확정추정치 간의 변동 추정치가 주어진(변동의 원인도 포함됨) 	<p>6.1 잠정통계와 확정통계 간의 일관성</p> <ul style="list-style-type: none"> •수정에 대한 계획표 •수정을 조건으로 하는 자료와 수정된 자료가 공개됨 •자세한 수정 분석에 대한 참고문/링크가 이용 가능함 <p>6.2 서로 다른 소스들로부터 얻어진 유사한 추정치들 간의 일관성</p> <ul style="list-style-type: none"> •다른 소스들로부터의 유사한 추

치들 간의 일관성 6.3 동일한 사회-경제적 도메인 내의 통계 간의 일관성 6.4 소스들과 결과물들 간의 일관성		정치에 대한 기술 •다른 소스들로부터의 추정치들의 비교 6.3 동일한 사회-경제적 도메인 내의 통계 간의 일관성 •일관성 있는 기반에서 얻어진 통계들에 관해 이용자들에게 조언 •일관성 없는 기반에서 얻어진 통계들에 관해 이용자들에게 조언 6.4 소스들과 결과물들 간의 일관성 •다른 결과물들에 대한 소스로서 사용되는 추정치들에 대해 추정치들을 비교하고 차이점들에 대한 이유 기술
7. 완결성 (completeness)		•포함범위와 자세한 사항, 그리고 확인되는 현재 자료에 관해 주요 사용자 요구들 간 차이 기술 •완결성 결여의 이유 기술 •요구에 맞는 계획 기술

Note. 위의 표는 이용자에게 중점을 두는 모든 결과물들에 대해 사용될 수 있는 품질 척도와 지표들을 나타낸다. 특정 결과물들에 대한 품질 척도와 지표, 그리고 지속적인 품질 개선을 시키고자 하는 생산자들에 적합한 품질 지표들은 현재 개발 중에 있다.

(참고문헌)

1. National Statistics (2003). Draft Quality Measurement and Reporting Framework Consultation Document [http://www.statistics.gov.uk/about/consultations/ons_consultations/default.asp].

4.6 스웨덴(Statistics Sweden)

품질 문제에 대한 현재의 지배적인 접근 방법은 다음과 같은 요소들을 갖는 전사적 품질(total quality)의 개념에 근거한다.

- 생산품의 품질은 생산품 이용자들의 의견과 그들의 이용 목적에 대한 적합성에 의해

결정된다.

- 품질의 개념은 생산품이 얼마나 이용자들의 요구와 기대에 부합하는지에 관한 모든 면들을 반영해야 한다.

품질 평가는 주관적인 견해를 가진 이용자들이 내리는 것이다. 그들의 평가는 생산품에만 의해 좌우되는 것이 아니라 생산품과 사용 목적의 결합에 의해 결정된다. 통계의 품질이란 공개되었을 때 통계가 얼마나 이용자들의 요구와 기대에 부응하는지에 대한 모든 관점들을 일컫는 것이다.

통계를 적절히 사용할 수 있도록 하기 위해서 이용자들은 통계의 특성들에 관한 정보가 필요하다. 이러한 목적을 위해 생산자는 흔히 '품질서(quality declaration)'라 불리는 중립적이고 기술적인 정보를 제공해 주어야 한다.

품질 개념은 통계적 조사계획(statistical survey planning)에 영향을 주는 이용자뿐만 아니라 생산자가 고려해야 할 품질 점검목록(checklist)을 제공해준다.

다음은 스웨덴 통계청의 공식통계에 대한 품질 개념이다.

개 념	분 류
1. 보고서의 내용(contents of reports)	<ul style="list-style-type: none"> • 통계적 목표 특성(statistical target characteristics) <ul style="list-style-type: none"> -단위(units)와 모집단(population) -변수(variables) -통계적 척도(statistical measures) -연구 도메인(study domains) -기준 시간(reference time) • 포괄성(comprehensiveness) <p>:특정 영역(예: 국가 통계시스템으로부터 얻어지는 경제 통계 전반)에 대한 통계 시스템을 말하며 많은 이용자들은 모든 중요한 측면들에 관한 정보를 제공하는 통계 시스템을 원함</p>
2. 정확성(accuracy)	<ul style="list-style-type: none"> • 전반적인 정확성(overall accuracy) <p>:전체 오차(total error)의 크기</p> • 부정확성의 원인(sources of inaccuracy) <ul style="list-style-type: none"> -표본추출(sampling) -포함범위(coverage) -측정(measurement) -무응답 자료(nonresponse data) -자료 처리(data processing) <p>: 코딩, 편집, 추정</p>

	<p>-모형 가정(model assumptions)</p> <ul style="list-style-type: none"> • 정확성 척도의 제시 (presentation of accuracy measures) : 정확성이 부족한 통계를 비판 없이 사용할 경우 잘못된 결론을 내릴 수 있음. 식견 있는 이용자들은 적절한 정확성의 척도가 제시되는 경우 오류를 피할 수 있음
3. 시의성(timeliness)	<ul style="list-style-type: none"> • 주기성(periodicity) : 반복적인 조사로부터 얻어지는 통계는 통상적으로 월, 분기, 년 등의 정기적인 계획에 따라 생산되므로 다음과 같은 주기성(또는 도수)이 다루어짐 <p>-생산자의 자료 수집 주기성 -발행되는 통계의 기준 시간(reference times)의 주기성 -통계가 공개되는 주기성</p> <ul style="list-style-type: none"> • 생산 시간(production time) : 기준 시점(기준 기간의 끝)과 통계 공표 시점 간의 차로서 생산 시간이 짧을수록 좋음 • 정시성(punctuality) : 공표 약속 시간과 실제 공표 시간의 일치를 말함
4. 일관성(coherence)과 비교성(comparability)	<ul style="list-style-type: none"> • 시간에 따른 비교성(comparability over time) : 시간에 걸쳐 반복되는 조사는 시계열(time series) 자료를 만들게 되므로 동일 시계열 자료 내에 있는 통계들이 어느 정도 동일한 방법으로 동일한 것을 추정하는지에 관심을 가짐. 예를 들어 조사하고자 하는 특성들에 대한 정의와 조사방법이 시간에 따라 변화가 없는 경우 비교가 용이함 • 공간에 따른 비교성(comparability over space) : 통계의 공통적인 용이성은 서로 다른 지역에 대해 비교가 가능하다는 것임. 또한 비지리적인 특성에 따른 비교, 예를 들어 자녀의 수에 따른 평균 가구 수입의 비교도 가능함. 시간에 따른 비교성과 마찬가지로 목표 특성들에 관한 정의들과 조사방법이 유사해야 함 • 일반적인 일관성(coherence in general) : 일관성은 비교 목적만이 아니고 서로 다른 소스들, 예를 들어 서로 다른 조사들로부터 얻어진 통계들을 결합하여 사용할 수 있다는 것을 말함. 이를 위해서는 단위, 모집단, 도메인, 변수, 기준 시간 등에 관한 정의가 일치해야 함
5. 이용가능성(availability)과 투명성(clarity)	<ul style="list-style-type: none"> • 배포 방식(forms of dissemination) : 어떠한 배포 매체(책자, 디스켓, CD-Rom 등)와 어떤 배포 경로가 사용되는가를 말함

	<ul style="list-style-type: none"> • 제시(presentation) : 통계가 어떠한 형식으로 배포 매체에 실리는 가를 말하는 것으로서 지면 배경과 원문, 그래프, 다른 수치들의 명료성 등을 포함함 • 문서화(documentation) : 발행된 통계에 관련한 문서 자료를 획득하는 사용자들의 능력을 말하며 대부분의 사용자들은 쉽게 읽을 수 있는 '품질서'를 원함. 또한 고급 사용자들은 개인적인 용도로 마이크로 데이터를 이용하고자 할 때 통계 생산 과정에 대한 정확한 문서 자료에 흔히 관심을 가짐 • 마이크로데이터에 대한 접근 (access to microdata) • 정보 서비스(information services) : 이용자 자신이 얻고자 하는 정보가 있을 때, 어떠한 도움을 받을 수 있는지 그리고 공표된 통계에 대한 자세한 질문들에 대해 답을 얻을 수 있는 지를 말함
--	---

Note. 다른 국가통계기관들에서는 품질 개념 중 스웨덴 통계청이 사용하는 '내용(contents)' 대신 적절성(relevance)이라는 용어를 사용한다. 적절성은 평가의 측면에만 의존하는 용어이다.

(참조)

스웨덴 통계청은 1994년, 이 기관의 모든 통계 생산 과정에 적용시킬 수 있는 품질 향상의 원칙을 채택했다. 이러한 결정은 전사적품질관리(Total Quality Management; TQM)에 관한 서적, 논문, 교육 과정과 TQM 원칙을 자신들의 일에 도입하거나 시도한 다른 통계기관들로부터의 경험에 관한 연구에 근거한 것이었다. 미국의 민간통계기관인 WESTAT은 청내의 최고 관리자들과 직원들에 대한 TQM 개념의 교육과 훈련을 맡았다. 그 이후로 프로젝트 내에서 TQM 팀을 감독하기 위한 70여명의 지도자들이 교육을 받았고, 1997년까지 100개 이상의 프로젝트들이 착수되었다. 약 70개의 프로젝트가 수행되어 비용 절감, 시의성의 향상과 생산과정에 대한 보다 깊은 지식 습득 등의 결과들을 얻었다. 직원들에 대한 조사에서 약 60퍼센트가 TQM이 스웨덴 통계청에 '매우 중요하다' 또는 '중요하다'고 답하였다.

품질개선에 관한 WESTAT의 주요 관점은 통계기관을 최종 생산품들을 생산하는 과정들로 이루어진 하나의 시스템으로 보는 것이다. 또한 TQM 개념에 있어서 한 가지 중요한 것은 통계 생산의 주요 과정들에 대해서 '현행최적방법들(Current Best Methods; CBM)'을 개발하고 사용하는 것이다. CBM은 점검목록(check lists)을 사용하는데 이는 반복적인 과정에 대해서는 정확한 진행 순서를 따르도록 하기 위해서이고 특이하고 창조적인 복잡한 과정에 대해서는 주요한 요소들을 상기시키도록 하기 위함이다.

(참고문헌)

1. Elvers E. and Rosen B. (1999). Quality Concept for Official Statistics. Encyclopedia of Statistical Science, 621-629. New York: John Wiley.
2. International Monetary Fund (2001). Sweden: Report on Observance of Standards and Codes-Data Module; Response by the Authorities; and Detailed Assessments[<http://dsbb.imf.org/vgn/images/pdfs/swerosc.pdf>].

4.7 네덜란드(Statistics Netherlands)

네덜란드 통계청은 1996년 'CBS 2000'이라는 종합적인 품질 프로그램을 채택했다. 이 사업 계획의 목적들 중 하나는 모든 통계 부서에 품질 시스템을 도입하는 것이었다. 이 시스템의 표준적인 모형이 준비되었고 품질 시스템에 대한 잠정적 가이드라인은 1997년 완성되었다. 또한 통계 감사 시스템(system of statistical auditing)을 도입하였는데 그 목적은 다음 세 가지로 나눌 수 있다.

- 통계 부서들이 품질 관리에 관하여 어떠한 일들을 하고 있는지 실제로 파악
- 어떻게 품질 관리를 해야 하는지에 관한 제안을 도출
- 무엇이 최적 방법인지 파악하여 이들을 품질 시스템을 위한 가이드라인에 포함시킴

공식 통계의 품질에 대한 여러 가지 측면이 있지만 네덜란드 통계청은 다음과 같은 품질 측면들을 고려한다.

- 적절성(relevance)
- 시의성(timeliness)
- 정확성(accuracy)
- 비용 효율성(cost-effectiveness)
- 응답자 부담(burden for data providers)

예를 들어, 네덜란드 통계청은 1996년 응답자 부담을 파악하기 위하여 '응답부담미터(response burden meter)'를 도입하였다. 이 지표는 1994년을 100을 기준으로 하여 사용

하며 매년 어느 정도 응답자 부담이 줄었는지 의회에 보고하고 있다.

(참고문헌)

1. Willem de Vries and Richard van Brakel (1998). Quality Systems and Statistical Auditing. A Pragmatic Approach to Statistical Quality Management. *Netherlands Official Statistics.*

4.8 미국(U.S. Census Bureau)

미국 통계국은 다른 국가 통계 기관들과 달리 통계 품질 차원을 따로 구분해서 채택을 하고 있지는 않으나, 2002년부터 배포되는 통계정보에 적용되는 품질 가이드라인(Census Bureau Section 515 Information Quality Guidelines)에는 기본적인 품질 기준들을 명시하고 있다. 이 기준들은 다음 표에서 설명되는 바와 같이 실용성(utility), 객관성(objectivity)과 정직성(integrity) 등 세 가지이며 이들 기준들은 세계 주요 국가 통계 기관들 및 국제 통계기관들에서 채택하고 있는 적절성(relevance), 시의성(timeliness), 접근성(accessibility), 정확성(accuracy) 등의 품질 차원 요소들을 포함하고 있다. 다음 표는 미국통계국이 채택하고 있는 품질 기준들이다.

품질 기준	내 용
1. 실용성(Utility) : 통계국은 일반에 배포되는 정보가 이용자들에게 유용하도록 하며 보다 유용한 자료를 계속해서 개발하기 위해 주력해야 함	<ul style="list-style-type: none"> •정책과 요구에 부응하는 적절한 객관적인 정보를 제공할 목적으로 통계 프로그램들에 대한 우선적인 사업들을 결정하기 위해 통계국은 자료 이용자들이 대한 지속적인 조사와 생산 검토뿐만 아니라 행정부, 의회, 관심 있는 비정부인들과 단체, 자문위원회와 조사 후원자들과 긴밀하게 협조함 •관리예산국(Office of Management and Budget)과 의회가 센서스와 같은 일부 주요 프로그램들에 대한 자료 내용 요구를 검토함 •적절성(relevance)이란 정보 생산품이 현재의 요구와 예상되는 미래의 요구 모두에 유용한 정보를 제공하는 정도를 말함. 통계국은 제공하는 정보가 계속해서 적절하도록 광범위한 이용자와 접촉을 유지함. 통계국이 수집하는 정보는 적절한 척도를 제공하기 위하여 설계되며 이들 척도들은 공식통계로서 일반에게 공개됨 •통계국은 적시에 통계정보가 일반에게 배포되도록 함. 시의성(timeliness)은 자료의 기준기간(reference period)과 자료 공개 시점 간의 근접성뿐만 아니라 자료 배포의 횟수를 포함하는 개념임. 일반에게 공개되는 자료가 높은 품질을 유지하도록 적시에 자료가 수집되고 발행되도록 노력해야함 •접근성(accessibility)이란 고객들이 통계적 자료, 생산품 또는 서비스를 얻는데 필요한 접근의 용이성 또는 노력을 의미함. 통계국은 자료 유용성과 이해성

	<p>에 관해 서로 다른 요구를 하는 다양한 범위의 자료 이용자들이 접근할 수 있도록 통계적 정보 생산품을 배포함. 통계국은 자료 이용자들에 대한 접근성과 이해성을 확보하기 위하여 유용성 테스트를 행함</p> <ul style="list-style-type: none"> •통계국은 접근의 편이, 간단한 인터페이스 메커니즘과 서로 다른 자료 소스들 간의 비교성에 대한 고객들의 기대에 부응하도록 계속적인 개선을 위해 노력함
<p>2.객관성 (Objectivity) : 통계국은 정확하고 신뢰할 수 있으며 편향 없는 정보를 제공하고 정보 생산품들이 정확하고 분명하고 완전하며 편향 없는 방법으로 제공되도록 함. 이러한 객관성은 신뢰할 수 있는 자료 소스들과 건전한 분석 기법들을 사용하고 세밀히 검토된 통계 생산품을 만들기 위하여 좋은 자격을 갖춘 사람들을 씬으로서 달성될 수 있음</p>	<ul style="list-style-type: none"> •객관성은 통계적 추정치들에 대한 오차 발생이 불가피하다는 것을 인식하는 것을 말함. 오차들은 표본오차(sampling errors)와 비표본오차(nonsampling errors)로 구분됨. 통계국은 이들 오차들에 대한 크기를 수량화한 정보를 제공함 •통계국은 객관성에 관련한 품질 개선을 위하여 다음을 실행함 <ul style="list-style-type: none"> -모든 전문적인 작업에 현대 통계적 이론과 방법 이용 -우리의 사명에 적절한 직원과 전문 기술의 개발 -자료 편집, 분석 과정의 개선을 포함한 자료의 타당성과 신뢰성을 개선시키기 위한 지속적 품질 보증 프로그램의 실행 -통계 분야와 적절한 주제 영역에 있는 전문 기관들과의 긴밀하고 지속적인 관계의 개발 •신뢰할 수 있는 자료 소스들의 사용 <ul style="list-style-type: none"> -통계국의 정보 생산품은 신뢰성과 정확성이 입증된 자료에 기초하여 만들어 짐 <ul style="list-style-type: none"> ◦통계국은 행정 기록과 다른 자료 소스들을 포함한 자료 소스, 측정 방법과 응답자 부담(respondent burden)을 최소화시키는 센서스와 표본조사를 위한 자료 수집과 처리 방법 등을 결정함 ◦이를 위해 통계국은 조사 요구와 목적, 정도(precision)의 수준, 지리적 범위, 자료수집방법과 응답자, 표본추출틀, 표본설계, 추정, 분산추정과 다른 측정 내용들을 결정함 -통계국은 자료 수집 과정에 품질 측정, 과정 관리 등을 넣음으로서 통계의 품질과 객관성이 명백해지도록 함 -통계국은 과정들의 일관성을 확보하기 위하여 자료 수집 과정의 모든 요소들을 포괄적으로 문서화함 •건전한 분석 기법들의 사용 <ul style="list-style-type: none"> -통계국은 통계적 생산품의 객관성을 확보하기 위하여 건전한 분석 기법들을 사용함 <ul style="list-style-type: none"> ◦통계국은 적절한 분석 방법들을 결정하고 사용할 책임을 가짐 ◦통계국은 자료처리를 분석하는데 사용하는 기법들을 평가하고, 보다 효과적이고 정확하며 신뢰할 수 있는 분석 도구들을 연구함 ◦통계국은 분석의 품질을 평가하고 보고함 -통계국은 적절한 통계적 검정들을 행하고 공식적인 생산품의 표본설계 특성들을 제시하며 표본오차들의 척도들을 문서화함 <ul style="list-style-type: none"> ◦결측자료, 측정오차, 처리오차 등을 포함한 비표본오차를 확인하기 위하여 예비품질검사 및 탐색적자료기법을 이용함 ◦분석을 위해 사용된 자료, 계산, 분석 결과를 보고하기 위하여 사용된 본문, 표와 숫자들을 포함하여 분석에서 발생할 수 있는 오차를 막기 위하여 추가적인 품질 검사를 행함 •자료 공개를 위한 배포 전 검토 <ul style="list-style-type: none"> -통계국에 의해 공개되는 모든 문서들은 내용, 통계적 조사 방법론과 문서의

	<p>정책 포함 등이 광범위하게 검토됨</p> <p>•자료 품질과 방법론을 이용자들에게 알림 -통계국은 자료의 수집과 처리에 사용되는 개념과 방법론들, 생산되는 자료의 품질, 자료의 이용과 해석에 영향을 미칠 수 있는 자료의 다른 특징들을 이용자들에게 알림 -통계적 생산품에는 사용된 방법과 절차와 이에 대한 참고문헌, 그리고 자료 사용에 영향을 미칠 수 있는 다른 정보들이 함께 포함됨 -이용자들에게 제공되는 문서에는 자료 품질에 관한 유용한 정보가 실리도록 하고 품질에 관한 인식을 불러일으키도록 함 -통계국은 일반에게 개념과 방법에 대한 정의와 기술과 함께 배포되는 통계적 정보의 품질에 대한 지표를 제공함 -방법론에 관한 정보는 이용자들로 하여금 자료가 자신들이 측정하고자하는 것에 적절히 근사하는지 그리고 자신들이 사용하는 추정치들이 자신들의 목적을 위해 허용할 수 있는 오차범위에서 생산되었는지 결정할 수 있도록 해야 함 -자료품질 또는 방법론에 관한 문서에서 제공되는 정보의 수준은 자료수집의 형태, 자료소스와 분석, 법령에 의한 비밀보호, 배포매체, 자료이용의 범위와 효과 그리고 통계 프로그램의 총예산 등에 의해 좌우됨</p> <p>•오차 수정을 위한 방침 -통계국 정보 품질 가이드라인을 따르지 않는 조직들에 의해 배포된 정보의 수정을 요구할 기회를 일반에게 부여하는 메커니즘을 마련함</p>
<p>3. 정직성 (Integrity) : 통계국에 의해 배포되는 정보는 부적절한 접근, 변경과 파괴로부터 보호됨</p>	<p>•통계국은 정보와 보조적인 정보를 포함한 배포된 정보가 손실, 오용, 불법 접근과 변경으로부터 보호되도록 함</p>

Note. 미국 통계국의 정보 품질 가이드라인에는 투명성(transparency)과 재생산성(reproducibility)을 추가적으로 언급하고 있으며 특히 정보의 폭 넓은 배포를 통하여 투명성을 확보하고자 함. 또한 이 품질 가이드라인 내의 “통계국 활동을 위한 시행 원칙(Performance Principles for Census Bureau Activities)”에는 통계국의 자료와 정보 생산품의 품질을 유지하기 위한 활동들에 관한 기본 원칙들을 절차별로 상세히 기술하고 있음.

(참고문헌)

1. U.S. Census Bureau (2003). Census Bureau Section 515 Information Quality Guidelines[<http://www.census.gov/qdocs/www/>].

5장 새로운 통계품질평가지표

2002년도에 실시한 통계품질평가결과를 분석하고 외국의 사례와 현행 사용하고 있는 통계품질평가지표를 검토해서 새로운 평가지표를 개발하였다. 새로운 평가지표는 현행 6개 차원(적절성, 정확성, 시의성, 접근성, 비교성, 효율성)에 일관성을 더하여 총 7개 차원으로 구성하였다. 7개 차원에서 81개 평가지표를 정하고 기존의 4개 평가지표는 재검토하는 것이 바람직하다. 81개 평가지표는 차원별로, 적절성 9개, 정확성 37개, 시의성 8개, 접근성 11개, 비교성 6개, 일관성 6개, 그리고 효율성 4개로 구성되었다. 81개 평가지표 중에서 우선적으로 사용해야 할 필수 평가지표 48개를 분류하였고, 그 외의 평가지표들은 통계의 품질을 향상시키기 위해서 점진적으로 다양하게 사용하는 것이 바람직하다. 물론 통계에 따라서는 사용할 지표가 다를 수 있다.

새로운 평가지표를 개발하고 사용하는데 있어서는 통계품질 향상에 필요한 적절한 인원과 예산이 마련되어야 함은 물론이고, 이들을 실행하는데 필요한 여건이 조성되어 있는 것을 기본 전제조건으로 한다. 또한 많은 사례에서도 입증된 바와 같이 기관장의 통계품질향상을 위한 의지는 통계의 품질향상에 절대적이다.

5.1 차원별 품질평가지표

<표 5-1>에서는 새로 구성된 81개의 평가지표들을 7개 차원별로 분류하였고, 각 차원을 몇 개의 부차원(sub-dimension)으로 세분화하였다. 81개의 평가지표들 중에서 통계품질을 평가하는데 반드시 필요한 48개의 필수 평가지표를 선정하였다. 또한 기존에 사용한 평가지표를 비교에 넣어 상호 비교가 용이하게 하였다.

<표 5-1> 차원별 품질평가지표(81개 항목)

차원	부차원	평가지표(항목)	비고	구분
1. 적절성(9)	1.1 이용자 요구	◦ 이용자들에 대한 체계적인 모니터링 프로그램이 있다.		필수
		◦ 이용자들의 요구에 대한 사업계획과 우선순위를 정한다.		
		◦ 이용자(만족도) 조사결과를 반영한다.	관련,1-5	필수
	1.2 통계개념 정의	◦ 이용자 요구에 맞는 통계 개념을 사용한다.	관련,1-1	필수
		◦ 개념간의 차이를 기술하고 있다.		
	1.3 설문지작성	◦ 설문지의 적합성을 평가하는 시스템을 갖추고 있다.	정확,3-1 정확,3-2 정확,3-4	필수
1.4 결과물	◦ 결과물의 주요 이용자들을 분류하고 있다.			

2. 정확성 (37)

		◦ 결과물에 대한 이용자 만족도 조사를 한다.			
		◦ 결과물에 대한 이용자들의 의견을 반영하고 있다.		필수	
2.1 일반사항		◦ 통계자료에 대한 오차(표본, 비표본)를 파악하고 있다.		필수	
		◦ 응답자들의 응답부담을 측정하고 줄이는 조치를 취하고 있다.			
		◦ 응답자들의 비밀보호를 위해 적극 노력하고 있다.	정확,8-8	필수	
		◦ 공표하기 전에 현장조사부터 집계분석 단계까지 적절한 내용검토 절차가 있다.	정확,6-2		
		◦ 간행물 발간 이전에, 간행물에 대한 사전점검체계가 있다.	정확,7-1	필수	
	2.2 표본오차		◦ 조사를 위한 표본설계를 한다.	정확,2-1 정확,2-2 정확,2-3 정확,6-3 효율,2-4	필수
			◦ 표본설계에 따른 진행과정을 점검하고 문제점등을 파악하여 조치하는 시스템을 운영하고 있다.		필수
			◦ 주요 추정치들에 대한 표본분산, 분산추정치, 그리고 변이계수 등을 계산한다.		필수
	2.3 비표본오차		◦ 설계효과를 계산한다.		
			◦ 비표본오차의 주요 원인들을 파악하여 조치해 주고 있다.		필수
			◦ 체계적인 현장조사 관리시스템이 구축되어있다.	정확,5-1 정확,5-2 정확,5-3 정확,5-4	필수
			◦ 주요 항목들에 대한 대체율을 계산한다.		
	2.3.1 포함오차		◦ 목표(연구)모집단과 조사모집단을 정의하고 비교하고 있다.		필수
			◦ 표본추출틀에 대한 내용을 설명하고 변동사항을 반영해 준다.	정확,2-5	필수
			◦ 포함오차(과소, 과대, 중복 등)를 계산한다.		
		◦ 포함오차의 영향과 문제 해결 방안을 제시하고 있다.			
2.3.2 무응답오차		◦ 응답거절, 결측 및 부적격 응답에 대한 교체방법을 제시하고 있다.	정확,5-5	필수	
		◦ 무응답오차의 원인을 파악하여 조치한다.	정확,6-1	필수	
		◦ 무응답률(단위, 항목)을 파악하고 무응답에 따른 가중치를 조정해주고 있다.			
		◦ 응답집단과 무응답집단의 특성을 비교·분석한다.			
		◦ 무응답 및 대체효과를 평가한다.			
2.3.3 측정오차		◦ 측정오차의 원인을 파악하여 조치한다.		필수	
		◦ 사전조사를 실시한다.	정확,3-3	필수	
		◦ 다양한 자료수집 방법들을 비교·분석 한다.			
		◦ 조사원에 대한 교육을 실시한다.	정확,4-1 정확,4-2 정확,3-5	필수	
		◦ 조사 직원의 조사관련 전문지식 숙지 여부를 파악하고, 이를 평가하는 제도적 장치가 있다.	정확,4-3	필수	
		◦ 대리응답률을 측정한다.			
		◦ 조사원간의 차의 특성을 측정한다.			

	2.3.4 처리오차	◦ 처리오차의 원인을 파악하여 조치한다.		필수	
		◦ 자료처리 시스템을 운영한다.		필수	
		◦ 편집오류율을 측정한다.			
		◦ 입력오류율을 계산한다.			
	2.3.5 모형오차	◦ 입력자간의 차의 특성을 파악한다.			
		◦ 사용한 모형과 추출절차를 명시하고 있다.		필수	
		◦ 이상치에 대한 조치를 취하고 있다.			
3. 시의성 (8)	3.1 주기성	◦ 자료수집이 주기적으로 이루어지고 있다.		필수	
		◦ 개편 작업이 정기적으로 이루어지고 있다.	시의,1-8	필수	
		◦ 통계자료의 공표를 주기적으로 한다.		필수	
	3.2 생산시간	◦ 통계작성시의 기준시점과 공표시점과의 차이는 얼마나 되는가?(잠정결과물과 최종결과물 모두에 대해)		필수	
		◦ 주문형 통계의 경우 요청시점과 전달시점과의 차이는 얼마나 되는가?			
		◦ 새로운(또는 특별)조사의 착수시점과 공표시점간의 차이는 얼마나 되는가?			
		◦ 조사결과 자료 공표시기로부터 DB계재 시기까지의 소요 시간은?	접근,8-3		
	3.3 정시성	◦ 통계자료의 공표계획 시점과 공표시점과의 차이는 얼마나 되는가?	시의,7-5 시의,7-6	필수	
	4. 접근성 (11)	4.1 보급	◦ 온라인상으로 통계자료에 대한 정보를 제공한다.	접근,7-4	필수
			◦ 통계자료와 관련 내용들을 여러 매체(책자, 디스켓, CD-ROM 등)로 제공한다.	접근,8-1	필수
4.2 정보이용 및 검색지원		◦ 원시자료를 포함한 마이크로자료에 대한 자세한 사용방법을 기술한다.	접근,8-7	필수	
		◦ 미공표(non-published)자료에 대한 자세한 사용방법을 기술			
		◦ 결과물을 검색엔진을 통해 검색할 수 있다.		필수	
		◦ 조사결과 자료에 대한 DB가 구축되어 신속하고 편리한 자료 이용이 가능하다.	접근,8-2	필수	
		◦ 관련 용어, 개념, 주석 등을 간행물에 충실하게 수록하여 이용자들의 이해를 돕는다.	접근,7-2	필수	
4.3 투명성		◦ 기술적 정보를 포함한 다양한 정보를 이용할 수 있다.			
		◦ 통계를 이해하고 사용하는데 도움을 주는 메타데이터를 제공한다.		필수	
		◦ 통계생산에 관련된 사항들을 공개한다.	접근,1-9 접근,2-6 접근,2-7 효율,8-4	필수	
		◦ 통계와 관계되는 모든 사항에 대한 질의 응답체계가 구축되어 있다.	접근,8-9	필수	

5. 비교성 (6)	5.1 시계열 비교	◦ 시계열 자료의 불연속 횟수는?	비교,8-6	필수
		◦ 시계열 변동의 영향을 평가한다.		
	5.2 국내외 비교	◦ 사용하고 있는 개념, 정의, 체계 등은 국내(국제)적 기준 또는 지침에 따르고 있다.	정확,1-3	필수
		◦ 국내(국제)적 기준 또는 지침에 벗어나는 이유를 파악한다.		필수
	◦ 자료에 대한 국제적 비교가 가능하다	비교,8-5		
	5.3 부차모집단 비교	◦ 부차모집단(도메인 포함)을 비교하고 있다.		
6. 일관성 (6)	6.1 통계작성의 일관성	◦ 통계조사시스템을 일관성 있게 유지하고 있다.		필수
		◦ 표준적인 목표 특성치들을 일관성 있게 사용한다.		필수
	6.2 통계들 간의 일관성	◦ 공표하고 있는 통계를 관련된 유사통계와 비교분석한다.	비교,6-4	필수
		◦ 서로 다른 통계(또는 추정치)들을 결합하는 방법 등을 사용하고 있다.		
		◦ 잠정통계에 대한 수정계획이 마련되어 있다.		
	◦ 잠정통계와 확정통계간의 차이를 설명하고 있다.		필수	
7. 효율성 (4)	7.1 기획과 생산과정	◦ 통계작성을 위한 효율적인 예산산정이 이루어지고 있다.	효율,1-4	필수
		◦ 조사방법, 자료처리, 그리고 추정방법 등의 변경(또는 개선)에 따른 효율성을 비교분석한다.		필수
	7.2 결과물 이용	◦ 투입된 비용에 따른 통계자료의 이용 빈도를 측정한다.		필수
		◦ 투입된 비용에 따른 통계자료의 판매비용을 측정한다.		
개선사항 (현행지표)		◦ 기획자는 직무교육, 회의참가 등을 통하여 업무와 관련된 전문지식을 습득하고 있다 ◦ 통계작성 관련 자료가 문서화되어, 일관성 있는 업무 추진이 가능하다. ◦ 조사기획자는 현장조사에 대한 업무량을 파악하고 있다. ◦ 조사 실시 주기에 비춰 결과자료는 시기적절하게 공표된다.	관련,1-6 정확,1-2 정확,1-7 시의,7-3	

※ 비교는 2002년 품질평가지표

5.2 품질평가지표의 필요성과 평가내용

<표 5-2>에서는 차원별로 평가지표를 사용하게 되는 필요성과 해당 평가지표에서 평가하는 내용들을 살펴보았다. 실제 평가지표를 선정해서 사용하는 경우는 품질관리팀이 현장 상황을 고려하고, 전문가들과의 협의를 거쳐서 선정하는 것이 바람직하다. 이들 평가지표들은 시대적 변화와 이용자들의 요구에 따라 수정보완이 되어야 하고, 통계별로 사용될 평가지표(항목)는 달라질 수 있다. 이들 평가지표들을 사용할 때는 통계품질 매뉴얼(가이드라인)을 함께 사용하여 도움을 받도록 한다.

<표 5-2> 품질평가지표의 필요성과 평가내용

○ 적절성

부차원	평가지표(항목)	필요성 및 평가내용	비고
1.1 이용자 요구	◦ 이용자들이 대한 모니터링 프로그램이 있다.	◦ 필요성; 통계품질의 목표는 이용자들의 사용적합성을 얼마나 충족시키는가에 있으므로 이용자들의 요구와 의견을 상시 파악하고 이들을 반영해야 한다. ◦ 평가내용; 통계 정보에 대한 이용자들의 의견과 현재와 미래에 대한 요구를 파악하기 위해 정부 또는 외부의 주요 이용자에게 대한 체계적인 모니터링 프로그램을 유지하고 있는지를 평가.	
	◦ 사업계획과 우선순위를 정한다.	◦ 필요성; 통계의 이용자와 이들의 요구는 매우 다양하다. 통계생산부서에서는 요구를 충족시킬 사업의 시행을 위해서 사업계획에 따라 사업의 우선순위를 정해야 한다. ◦ 평가내용; 이용자들의 요구 사항을 정리하고, 이들에 대한 사업계획수립과 사업의 중요도에 따른 우선순위를 정하고 있는지를 평가. 여기서는 개별 통계에 대한 우선순위를 말함.	
	◦ 이용자(만족도) 조사결과를 반영한다.	◦ 필요성; 이용자들의 사용적합성을 만족시키는 통계를 생산하기 위해서는 이용자들의 의견과 요구를 파악하는 것이 1차적이지만 궁극적으로는 이들이 충분히 반영되어야 한다. 여기서의 만족도 조사는 전체적인 이용자들의 만족도 조사이다. ◦ 평가내용; 이용자만족도 조사결과로부터 반영된 사항과 미반영 사항을 구분하고 사유를 구체적으로 적시하고 있는지, 그리고 반영 정도가 어느 정도인지를 평가.	관련,1-5
1.2 통계 개념 정의	◦ 이용자 요구에 맞는 통계 개념을 사용한다.	◦ 필요성; 생산된 통계의 개념이 이용자들이 요구하는 개념과 일치하지 않을 경우 이용의 혼선은 물론 잘못된 개념의 사용으로 그릇된 결과를 초래할 수 있다. ◦ 평가내용; 이용자들이 요구하는 통계의 개념이 무엇인지를 파악하고, 이들에 적합한 개념을 사용하고 있는지를 평가. 예를 들면, 요구하는 척도, 변수, 추출단위, 도메인, 기준기간 등의 적절성.	관련,1-1
	◦ 개념간의 차이를 기술하고 있다.	◦ 필요성; 생산된 통계의 개념이 이용자들이 요구하는 개념과 일치하지 않을 경우 그 차이를 상세히 설명하여 이용의 편리성을 제공함은 물론 이용의 오류를 방지한다. ◦ 평가내용; 차이가 나는 개념들, 개념간의 차이점, 그리고 개념간의 차이로 인해 생기는 문제점등을 구체적으로 설명하고 있는지를 평가.	
1.3 설문지 작성	◦ 설문지의 적합성을 평가하는 시스템을 갖추고 있다.	◦ 필요성; 설문지는 측정하고자 하는 것을 제대로 측정할 수 있도록 설계되어야 한다. ◦ 평가내용; 각각의 설문 문항들이 측정하고자 하는 것을 정확	정확,3-1 정확,3-2 정확,3-4

		히 측정하도록 작성되고 있음은 물론 응답자들의 입장을 잘 반영하고 있는지를 평가한다 또한, 설문 작성에 전문가 집단이 참여하고, 설문지의 신뢰성과 타당성을 측정할 수 있는 팀이 구성되어 있는지를 평가.	
1.4 결과물	◦ 결과물의 주요 이용자들을 분류하고 있다.	◦ 필요성; 통계 결과의 주요 이용자 그룹을 분류해서 그룹별 이용상태와 그룹별 특성(요구 사항과 의견 등의 차이) 비교가 필요하다. ◦ 평가내용; 주요 이용자 그룹의 분류기준, 그룹별 특성비교, 그룹별 관리상태 등을 평가.	
	◦ 결과물에 대한 이용자 만족도 조사를 한다.	◦ 필요성; 전체적인 이용자 만족도 조사와는 별개로 개별 통계에 대한 이용자 만족도 조사를 한다. 이유는 통계별로 이용자 그룹과 이용 특성 등이 다르게 되므로 이에 따른 만족도 조사를 실시해서 한다. ◦ 평가내용; 통계별 이용자 그룹이 분리되어 있는지, 통계별 요구 사항과 의견들을 수합하여 비교 분석되고 있는지를 평가	
	◦ 결과물에 대한 이용자들의 의견을 반영하고 있다.	◦ 필요성; 생산된 통계의 품질을 높이는데 이용자들에 대한 만족도 조사결과를 반영해야 한다. 따라서 이들의 반영 정도를 평가하는 것은 필요하다. ◦ 평가내용; 이용자로부터의 피드백 또는 이용자 만족도 조사에 근거하여 개선하기 위한 적절한 조치를 취하고 있는지를 평가.	

○ 정확성

부차원	평가지표(항목)	필요성 및 평가내용	비고
2.1 일반 사항	◦ 통계자료에 대한 오차(표본, 비표본)를 파악하고 있다.	◦ 필요성; 생산된 통계의 질(정확성)을 평가하는 가장 기본적인 잣대는 오차이다. ◦ 평가내용; 각종 조사에 의해 얻어지는 통계자료에 대해 정기적으로 표본오차(전수조사의 경우는 표본오차가 없으므로 제외)와 비표본오차를 측정하고 관리하고 있는 지를 평가.	
	◦ 응답자들의 응답부담을 측정하고 줄이는 조치를 취하고 있다.	◦ 필요성; 최근 조사에서 응답자들의 응답거부나 무응답이 증가하고 있는데, 그 이유 중의 하나는 응답자들에게 지나친 응답부담을 주기 때문이다. 따라서 응답자들의 응답부담을 측정하고 줄이는 적절한 조치를 취하는 것이 필요하다. ◦ 평가내용; 응답자들의 응답 부담을 수량화하여 주기적으로 측정하고 있는지, 응답자들의 응답부담을 덜어주기 위해서 이미 존재하는 행정적 자료를 이용하여 응답자들의 부담을 최소화하고, 불필요한 통계자료의 수집을 피하고 있는지, 그리고 응답부담을 덜어주는 조사방법을 사용하고 있는지를 평가.	

	<ul style="list-style-type: none"> ◦ 응답자들의 비밀보호를 위해 적극 노력하고 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 정확한 통계는 응답자들의 신분(비밀)이 보장되는 경우에 이루어질 수 있다. 그러므로 응답자들에 대한 비밀은 철저하게 보장되어야 한다. ◦ 평가내용; 응답자들의 비밀을 보장해주는 통계법 13조에 대한 공지여부와 비밀보호의 필요성, 보호할 내용, 보호방법, 보호의무 등에 대한 내용을 응답자들이 충분히 숙지할 수 있도록 조치하고 있는지를 평가. 	정확,8-8
	<ul style="list-style-type: none"> ◦ 공표하기 전에 현장조사부터 집계분석 단계까지 적절한 내용검토 절차가 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 통계가 공표된 후 변경이나 수정 등은 통계의 신뢰도를 하락시키고, 통계생산에 관계되는 모든 부서 및 종사자들에 대한 불신을 증가시키게 된다. 따라서 공표 전 점검작업이 필요하다. ◦ 평가내용; 통계작성의 전 과정을 충실히 수행하였는지, 점검시스템(점검인력, 점검내용 및 지침서 등)이 갖추어져 있는지를 평가. 	정확,6-2
	<ul style="list-style-type: none"> ◦ 간행물 발간 이전에, 간행물에 대한 사전점검체계가 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 이용자들은 통계생산에 관계되는 모든 결과를 수록하고 있는 간행물을 이용하게 된다. 따라서 이용자의 이용 편리성을 높이고 이용 시 오류를 최대한 줄이기 위해서 간행물에 대해서도 사전점검이 필요하다. ◦ 평가내용; 통계 이용자들이 이용하고 있는 간행물, 보도자료, 보고서, 월보 등에 대해 간행물 사전점검절차를 마련하여 수행하고 있는 지, 그리고 통계자료 점검지침을 준수하고 있는지를 평가. 	정확,7-1
2.2 표본 오차	<ul style="list-style-type: none"> ◦ 조사를 위한 표본설계를 한다. 	<ul style="list-style-type: none"> ◦ 필요성; 건물이나 공장을 건설할 때 설계가 필요하듯이 표본조사를 통해서 통계를 생산하는 경우에도 표본설계가 필요하다. 표본설계를 하지 않으면 계획적인 조사가 이루어지기 어렵고 조사된 자료로부터 적합한 통계생산과 추정에 문제가 발생할 가능성이 높다. ◦ 평가내용; 모집단이 정확히 정의되고 있는지, 표본추출틀, 표본추출방법과 표본크기는 적절한지, 그리고 추정치와 추정치의 분산계산 공식들이 유도되었는지 등을 평가. 	정확,2-1 정확,2-2 정확,2-3 정확,6-3 효율,2-4
	<ul style="list-style-type: none"> ◦ 표본설계에 따른 진행과정을 점검하고 문제점등을 파악하여 조치하는 시스템을 운영하고 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 실제 조사는 표본설계에 따라서 진행되어야 하므로 수시로 진행과정을 점검하여야 한다. ◦ 평가내용; 추출틀 사용에 문제는 없는지, 표본크기는 제대로 유지되는지, 표본교체(substitution)는 지정된 방법대로 진행되고 있는지, 추정방법의 적용에는 문제가 없는지 등 표본설계에 의해 조사가 진행되는 과정에서 발생하는 문제들을 파악하여 수정하고 조치하는 가를 평가. 	
	<ul style="list-style-type: none"> ◦ 주요 추정치들에 대한 표본 분산, 분산추정치, 그리고 변이계수 등을 계산한다. 	<ul style="list-style-type: none"> ◦ 필요성; 통계조사의 목적 중에 하나는 조사대상 모집단의 특성을 파악하는 것이므로 조사 자료를 바탕으로 모집단 특성값을 추정한다. ◦ 평가내용; 모집단의 특성을 나타내는 추정치를 구하고 추정 	

		<p>치의 분산추정값과 변이계수를 계산한다. 무응답 등으로 인한 정확한 분산추정값 계산이 곤란한 경우에는 근사계산을 한다.</p>	
	<p>◦ 설계효과를 계산한다.</p>	<p>◦ 필요성; 기존의 방법과 다른 설계를 한 경우(예를 들어, 추출방법, 표본 수, 그리고 추정방법 등을 달리한 경우) 그 효과를 계산해야 한다. ◦ 평가내용; 기존의 방법을 변경하거나 새로운 방법의 사용 시 주요 변수들에 대한 분산(또는 분산추정치)을 구해서 기존 방법들의 분산과 비교한다.</p>	
2.3 비 표 본 오 차	<p>◦ 비표본오차의 주요 원인들을 파악하여 조치해 주고 있다.</p>	<p>◦ 필요성; 조사 환경의 변화로 비표본오차가 증가하고 있는 추세이다. 따라서 이들의 원인을 파악해서 적절한 조치를 취하는 것이 요구된다. ◦ 평가내용; 주요 비표본오차의 원인들이 포함오차, 무응답오차, 측정오차, 처리오차 등으로 분리되어 파악되고 있으며, 이에 대해 적절한 조치들이 취해지고 있는지를 평가.</p>	
	<p>◦ 체계적인 현장조사 관리시스템이 구축되어있다.</p>	<p>◦ 필요성; 생산된 통계의 품질은 조사 현장에서의 작업이 얼마나 체계적이고 효율적으로 운영되느냐에 달려있다. ◦ 평가내용; 조사 직원의 조사방법 숙지 정도, 업무수행을 관리하기 위한 점검체계, 현장조사 시 발생하는 문제의 처리 방법, 그리고 이러한 해결 방법의 공유 정도 등을 평가.</p>	<p>정확,5-1 정확,5-2 정확,5-3 정확,5-4</p>
	<p>◦ 주요 항목들에 대한 대체율을 계산한다.</p>	<p>◦ 필요성; 우선적으로 주요 조사 항목들에 대한 무응답발생 정도를 파악해야 항목 무응답으로 인한 오류를 줄일 수 있다. ◦ 평가내용; 주요 항목들에 대한 무응답 대체(imputation)방법과 대체율을 구하고 있는지를 평가</p>	
2.3.1	<p>◦ 목표(연구)모집단과 조사모집단을 정의하고 비교하고 있다.</p>	<p>◦ 필요성; 통상적으로 목표모집단과 조사모집단간에 차이가 있으므로 실제 조사대상모집단을 정의하고 목표모집단과의 차이를 비교하는 것이 필요하다. ◦ 평가내용; 목표모집단을 대신할 조사모집단이 무엇이고, 목표모집단과 차이 나는 점(기준시점의 차이, 누락 또는 제외 사항)들을 상세히 설명하고 문서화 되어 있는지를 평가</p>	
포함 오차	<p>◦ 표본추출틀에 대한 내용을 설명하고 변동사항을 반영해 준다.</p>	<p>◦ 필요성; 조사대상 표본은 표본추출틀을 이용해서 선정된다. 표본추출틀은 시간에 따라 변하므로 표본추출틀에 대한 전반적인 사항들을 파악하고 변동사항을 바로 반영하여야 한다. ◦ 평가내용; 원자료의 작성일자과 최종 갱신일자를 포함한 표본추출틀에 대한 내용을 파악하고 변동사항이 반영되고 있는지를 평가.</p>	<p>정확,2-5</p>

	<ul style="list-style-type: none"> ◦ 포함오차(과소, 과대, 중복 등)를 계산한다. 	<ul style="list-style-type: none"> ◦ 필요성; 완전한 추출틀은 현실적으로 불가능하므로 포함오차가 생기게 된다. 이들 포함오차들을 파악하고 관리해야 한다. ◦ 평가내용; 추출틀의 과대, 과소, 중복, 부적격과 오분류 정도가 어떻게 측정되고 있는지를 평가. 	
	<ul style="list-style-type: none"> ◦ 포함오차의 영향과 문제 해결 방안을 제시하고 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 포함오차의 정도에 따라 추정에 오류가 생기므로 이의 영향을 측정해서 해결방안을 마련해야 한다. ◦ 평가내용; 포함오차(과소, 과대, 중복 등)에 따라 추정에서 발생하는 문제, 그리고 이들을 해소하기 위해서 사용되는 방법, 나아가 포함오차를 줄이는 방안들이 마련되고 있는지를 평가. 	
	<ul style="list-style-type: none"> ◦ 응답거절, 결측 및 부적격 응답에 대한 교체방법을 제시하고 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 최근에는 표본으로 선정된 응답자들의 접촉이 어려워지고, 응답자들의 응답거부가 증가하고 있으므로 표본교체를 통해 당초 계획한 표본크기를 유지할 필요가 있다. ◦ 평가내용; 응답자 결측, 응답거부, 부적격 응답자들에 대한 적절한 교체(substitution)방법을 사용하고 있는지를 평가. 	정확,5-5
	<ul style="list-style-type: none"> ◦ 무응답오차의 원인을 파악하여 조치한다. 	<ul style="list-style-type: none"> ◦ 필요성; 무응답의 원인이 다양하므로 그 원인을 파악해서 무응답을 줄이는 것이 필요하다. ◦ 평가내용; 무응답오차의 원인인 응답부재, 응답거부, 응답불능, 부적격 등의 비율이 파악되고 이에 대한 조치가 취해지고 있는지를 평가. 	정확,6-1
2.3.2 무응답 오차	<ul style="list-style-type: none"> ◦ 무응답률(단위, 항목)을 파악하고 무응답에 따른 가중치를 조정해주고 있다. 	<ul style="list-style-type: none"> ◦ 필요성; 무응답이 발생하면 관측값이 줄어들고 편향이 생기므로 표본단위들에 대한 가중치조정이 필요하다. ◦ 평가내용; 단위무응답률과 항목무응답률을 파악하고 무응답으로 인한 가중치를 조정해서 추정에 반영해주고 있는지를 평가. 	
	<ul style="list-style-type: none"> ◦ 응답집단과 무응답집단의 특성을 비교 분석한다. 	<ul style="list-style-type: none"> ◦ 필요성; 조사자료는 응답집단으로부터 얻어지므로 응답집단의 특성만이 반영될 경우 추정이 정확치 못하게 될 수 있다. 따라서 무응답집단과 응답집단의 특성을 비교해서 차이를 반영해주어야 한다. ◦ 평가내용; 응답집단과 무응답집단 간의 인구학적 특성과 그룹 간 특성치를 비교하고, 모집단 특성에 의한 예상 무응답 편향 등을 계산하여 반영하고 있는지를 평가. 	
	<ul style="list-style-type: none"> ◦ 무응답 및 대체효과를 평가한다. 	<ul style="list-style-type: none"> ◦ 필요성; 무응답이 추정치에 주는 영향과 무응답을 다른 값으로 대체해 준 경우의 효과를 측정해야 한다. ◦ 평가내용; 무응답이 모집단총계나 평균 등의 추정에 주는 효과와 무응답을 적절한 방법을 사용해 다른 값으로 대체해줄 경우의 대체효과를 측정하고 있는지를 평가. 	

2.3.3 측정 오차	<ul style="list-style-type: none"> • 측정오차의 원인을 파악하여 조치한다. 	<ul style="list-style-type: none"> • 필요성; 측정오차를 줄이기 위해서는 원인을 파악해서 그에 맞는 조치를 취해주어야 한다. • 평가내용; 측정오차는 측정대상, 측정방법, 측정절차 등에 따라 달라지므로 이들을 정의하고 측정된 오차를 원인별로 분류하며 그에 대한 적절한 조치를 취하고 있는지를 평가 	
	<ul style="list-style-type: none"> • 사전조사를 실시한다. 	<ul style="list-style-type: none"> • 필요성; 조사를 실시하기 전에 조사도구인 설문지 작성이 적절한지를 점검해야 한다. • 평가내용; 사전조사를 실시하고 조사된 결과를 설문지 수정 및 보완에 반영하고 있는지를 평가. 	정확,3-3
	<ul style="list-style-type: none"> • 다양한 자료수집 방법들을 비교·분석 한다. 	<ul style="list-style-type: none"> • 필요성; 조사에 적합한 자료수집방법을 선택하기 위해 다양한 자료수집 방법들을 비교분석한다. • 평가내용; 조사에 가장 적합한 방법을 찾기 위해서 면접조사, 우편조사, 전화조사, 인터넷조사 등 다양한 자료수집 방법들의 장·단점을 비교·분석하여 반영하고 있는지를 평가. 	
	<ul style="list-style-type: none"> • 조사원에 대한 교육을 실시한다. 	<ul style="list-style-type: none"> • 필요성; 통계자료의 질은 조사원들에 의해 크게 영향을 받으므로 조사원들에 대한 교육은 절대적으로 필요하다. • 평가내용; 조사지침서, 조사표 작성방법, 조사직원 수행사항, 용어해설 등 조사원 교육을 위한 자료가 충실히 준비되어 있고, 교육이 체계적인 프로그램에 의해 실시되고 있는지를 평가. 	정확,4-1 정확,4-2 정확,3-5
	<ul style="list-style-type: none"> • 조사 직원의 조사관련 전문 지식 숙지 여부를 파악하고, 이를 평가하는 제도적 장치가 있다. 	<ul style="list-style-type: none"> • 필요성; 조사 직원들에 대한 전문지식 습득여부를 평가하고 이를 반영해 주어야 조사 직원들의 전문성이 향상된다. • 평가내용; 조사 직원들의 교육내용 습득정도와 조사관련 전문지식의 숙지여부를 측정해서 반영해주는 제도적 장치가 있는가를 평가. 	정확,4-3
	<ul style="list-style-type: none"> • 대리응답률을 측정한다. 	<ul style="list-style-type: none"> • 필요성; 대리응답이 자료의 질적 수준에 영향을 주므로 이를 파악하여 조치할 필요가 있다. • 평가내용; 대리응답률을 측정하고 대리응답의 차이점 등을 분석해서 조사나 추정에 반영하고 있는지를 평가. 	
	<ul style="list-style-type: none"> • 조사원간의 차의 특성을 측정한다. 	<ul style="list-style-type: none"> • 필요성; 조사원간에 측정오차의 차이가 있으므로 이를 파악해서 조사원간의 수준차이로 인한 오류를 줄여야 한다. • 평가내용; 조사원간 특성(평균 측정오류 수, 오류율, 오류에 대한 분산 등)을 측정하여 조사원 교육, 보상, 그리고 조사 전반에 반영하고 있는지를 평가. 	
2.3.4 처리	<ul style="list-style-type: none"> • 처리오차의 원인을 파악하여 조치한다. 	<ul style="list-style-type: none"> • 필요성; 자료처리 과정에서 발생하는 오차를 원인별로 파악해서 오차를 줄이도록 한다. 	

오차		<ul style="list-style-type: none"> 평가내용; 자료의 부호화과정, 입력방법 등에 따른 오차를 측정하고 오차를 줄이는데 반영하고 있는지를 평가 	
	<ul style="list-style-type: none"> 자료처리 시스템을 운영한다. 	<ul style="list-style-type: none"> 필요성; 자료처리 과정은 복잡하고 새로운 기술이 지속적으로 도입되므로 체계적인 시스템 운영이 필요하다. 평가내용; 자료처리과정, 전산화, 품질관리방법들이 어떻게 적용되고 있는지를 평가. 	
	<ul style="list-style-type: none"> 편집오류율을 측정한다. 	<ul style="list-style-type: none"> 필요성; 비용을 줄이고, 정확하고 신뢰있는 통계를 만들기 위해 편집오류율을 측정한다. 평가내용; 편집의 기준, 방법, 절차 등이 마련되고 편집의 오류가 측정되고 있는지를 평가. 	
	<ul style="list-style-type: none"> 입력오류율을 계산한다. 	<ul style="list-style-type: none"> 필요성; 완전한 data set을 만들기 위해서는 입력오류를 줄이는 것이 무엇보다 중요하다. 평가내용; 사람과 기계에 의한 입력오류율을 측정하고 조치방안이 마련되어 있는지를 평가. 	
	<ul style="list-style-type: none"> 입력자간의 차의 특성을 파악한다. 	<ul style="list-style-type: none"> 필요성; 자료를 입력하는 사람에 따라 입력의 차이가 생기므로 입력자간 특성 차이를 분석해서 입력오류를 줄인다. 평가내용; 입력자간 특성(입력오류 수에 대한 평균, 분산, 비율 등)을 측정하여 입력오류를 줄이는데 반영하고 있는지를 평가. 	
2.3.5 모형 오차	<ul style="list-style-type: none"> 사용한 모형과 추출절차를 명시하고 있다. 	<ul style="list-style-type: none"> 필요성; 모형기반에서는 사용한 모형에 대한 설명과 표본추출의 경우 추출절차가 명확해야 한다. 평가내용; 모형 사용을 위한 가정, 사용된 모형의 특성, 그리고 이들 모형으로부터의 추출절차의 적합성을 평가. 	
	<ul style="list-style-type: none"> 이상치에 대한 조치를 취하고 있다. 	<ul style="list-style-type: none"> 필요성; 이상치는 통계분석에 있어서 큰 영향을 미치므로 이상치에 대한 적절한 조치가 필요하다. 평가내용; 이상치를 확인하고 발생원인을 찾아, 이상치의 효과를 분석해서 조치해주고 있는지를 평가. 	
	<ul style="list-style-type: none"> 모형에 의한 오차(편향, 분산)를 계산한다. 	<ul style="list-style-type: none"> 필요성; 모형의 타당성을 평가하기 위해서는 모형에 의한 오차를 계산해야 한다. 평가내용; 사용한 모형에 의한 편향과 분산 등을 계산하여 모형의 타당성을 분석하고 있는지를 평가. 	
	<ul style="list-style-type: none"> 추정, 가중치부여, 추정방법 등에 대한 적용과정을 비교한다. 	<ul style="list-style-type: none"> 필요성; 새로운 모형과 추정기법들을 개발하기 위해서는 사용한 모형으로부터 추정에 관련된 제반사항들을 비교해야 한다. 평가내용; 가중치계산, 이용한 추정방법에 대한 상세한 분석과 야기되는 문제들을 파악하여 반영하고 있는지를 평가. 	

○ 시의성

부차원	평가지표(항목)	필요성 및 평가내용	비고
3.1 주기성	<ul style="list-style-type: none"> 자료수집이 주기적으로 이루어지고 있다. 	<ul style="list-style-type: none"> 필요성; 시계열 자료를 생산하고, 자료수집시기 차이로 인한 오차를 줄이기 위해서는 자료수집이 주기적이어야 한다. 평가내용; 자료 수집이 주기적으로 이루어지고 있는지, 만약 주기적이지 못하다면 그 이유가 명확히 설명되고 비주기적으로 인한 문제점들이 보완되고 있는지를 평가. 	
	<ul style="list-style-type: none"> 개편 작업이 정기적으로 이루어지고 있다. 	<ul style="list-style-type: none"> 필요성; 조사 관련 개편작업이 정기적으로 이루어져야 자료의 노후화와 오류를 방지할 수 있다. 평가내용; 전체적으로 개편작업이 정기적으로 이루어지고 있으며, 개편 시 변경 사항들에 대한 비교연구가 이루어지고 있는지를 평가. 	시의,1-8
	<ul style="list-style-type: none"> 통계자료의 공표를 주기적으로 한다. 	<ul style="list-style-type: none"> 필요성; 통계자료를 주기적으로 공표해야 이용자들의 이용 편리성을 유지할 수 있고 시계열자료의 이용이 가능하게 된다. 평가내용; 통계자료의 공표가 주기적으로 이루어지고 있는지, 만약 주기적이지 못하다면 그 이유가 명확히 설명되고 비주기적으로 인한 문제점들이 보완되고 있는지를 평가. 	
3.2 생산 시간	<ul style="list-style-type: none"> 통계작성시의 기준시점과 공표시점과의 차이는 얼마나 되는가?(잠정결과물과 최종결과물 모두에 대해) 	<ul style="list-style-type: none"> 필요성; 통계 이용자들에게 가능한 한 최근의 자료를 제공해야 한다. 평가내용; 통계작성시의 기준시점과 공표시점의 차이를 측정하고 차이 발생의 원인을 파악하여 조치하고 있는지를 평가. 잠정결과물이 있는 경우도 마찬가지이다. 	
	<ul style="list-style-type: none"> 주문형 통계의 경우 요청시점과 전달시점과의 차이는 얼마나 되는가? 	<ul style="list-style-type: none"> 필요성; 이용자들이 요청한 주문 통계를 최대한 빠른 시일 내에 제공해야 한다. 평가내용; 주문통계의 경우, 요청시점과 제공시점과의 차이를 측정하고 차이가 발생하는 원인을 파악하여 이용자들이 통계를 시의 적절하게 사용할 수 있도록 조치하였는지를 평가. 	
	<ul style="list-style-type: none"> 새로운(또는 특별)조사의 착수시점과 공표시점간의 차이는 얼마나 되는가? 	<ul style="list-style-type: none"> 필요성; 조사 시작부터 공표 시까지의 소요시간을 측정해서 최소화하는 방법을 개발한다. 평가내용; 조사 착수시점과 공표시점간의 차이를 측정하고 과정별 소요시간, 시간소요의 원인과 주요 지연 사항들을 파악하고 조치하고 있는지를 평가. 	
	<ul style="list-style-type: none"> 조사결과 자료 공표시기로부터 DB수록 시기까지의 소요 시간 	<ul style="list-style-type: none"> 필요성; 자료공표와 함께 DB수록하여 이용자들이 편리하고 신속하게 통계자료를 이용토록 한다. 	접근,8-3

		<ul style="list-style-type: none"> 평가내용; 최초의 공표시점과 DB수록 시점간의 차이를 계산하고 차이가 기준치보다 클 경우 지연의 원인과 차이를 줄이는 방안이 있는지를 평가. 	
3.3 정시성	<ul style="list-style-type: none"> 통계자료의 공표계획 시점과 공표시점과의 차이는 얼마나 되는가? 	<ul style="list-style-type: none"> 필요성; 통계자료는 계획대로 공표되어야 이용자들의 이용적합성을 만족시키고, 공표된 통계에 대한 신뢰를 유지할 수 있다. 평가내용; 공표계획 시점과 공표시점과의 차이를 계산하고 차이가 발생한 경우 그 원인을 파악해서 조치를 취했는지를 평가. 	시의,7-5 시의,7-6

○ 접근성

부차원	평가지표(항목)	필요성 및 평가내용	비고
4.1 보급	<ul style="list-style-type: none"> 온라인 상으로 통계자료에 대한 정보를 제공한다. 	<ul style="list-style-type: none"> 필요성; 정보통신기술과 컴퓨터의 발달로 통계자료를 온라인상에서 이용하려는 요구가 급증하고 있다. 평가내용; 통계와 각종 문서, 그리고 보고서 등을 적절한 형식(한글, Excel, MS word, PDF 등)으로 Web상에 제공하고, 이를 이용자들이 동시에 신속하게 사용할 수 있는지를 평가. 	접근,7-4
	<ul style="list-style-type: none"> 통계자료와 관련 내용들을 여러 매체(책자, 디스켓, CD-ROM 등)로 제공한다. 	<ul style="list-style-type: none"> 필요성; 통계 이용자들의 편리성을 최대한 제공하기 위해서 다양한 매체를 사용해야 한다. 평가내용; 통계와 각종 문서, 그리고 보고서 등을 책자, 디스켓, CD-ROM 등으로 제공하고 있으며, 이들을 손쉽게 획득할 수 있는 방법들이 알려져 있는지를 평가. 	접근,8-1
4.2 정보 이용 및 검색 지원	<ul style="list-style-type: none"> 원시자료를 포함한 마이크로자료에 대한 자세한 사용방법을 알려준다. 	<ul style="list-style-type: none"> 필요성; 원시자료를 포함한 마이크로자료에 대한 이용이 편리하도록 해야 한다. 평가내용; on-line과 off-line 상에서 원시자료를 포함한 마이크로자료에 대한 이용방법과 이용비용 등을 자세히 알려주고 있는지를 평가. 	접근,8-7
	<ul style="list-style-type: none"> 미공표(non-published)자료에 대한 자세한 사용방법을 알려준다. 	<ul style="list-style-type: none"> 필요성; 미공표자료 중에서 사용 가능한 경우와 범위를 상세히 알려주어야 한다. 평가내용; 미공표자료의 사용조건(사용가능 경우와 사용가능 정보, 그리고 정보 사용에 소요되는 비용과 시간 등)을 상세히 알려주고 있는지를 평가. 	
	<ul style="list-style-type: none"> 결과물을 검색엔진을 통해 검색할 수 있다. 	<ul style="list-style-type: none"> 필요성; 결과물을 Web상에서 이용하는 사례가 증가하고 있으므로 이용의 편리성을 제공해주어야 한다. 평가내용; 결과물을 검색할 수 있는 검색엔진의 종류, 특성, 그리고 사용방법 등을 상세히 설명하고 있는지를 평가. 	
	<ul style="list-style-type: none"> 조사결과 자료에 대한 DB가 구축되어 신속하고 편리한 자료 이용이 가능하다. 	<ul style="list-style-type: none"> 필요성; 통계를 신속하고 편리하게 이용할 수 있는 DB체계가 구축되어 있어야 한다. 	접근,8-2

		<ul style="list-style-type: none"> 평가내용; 개별통계에 대한 DB가 구축되어있고 DB가 구축되어 있는 통계목록을 제공하고, DB이용에 대한 상세한 설명이 되어 있는지를 평가. 	
	<ul style="list-style-type: none"> 관련 용어, 개념, 주석 등을 간행물에 충실하게 수록하여 이용자들의 이해를 돕는다. 	<ul style="list-style-type: none"> 필요성; 간행물 이용에 불편이 없도록 각종 정보(용어, 개념, 주석 등)를 간행물 내에 수록해주어야 한다. 평가내용; 보도자료, 보고서, 월보 등의 간행물 이용에 불편이 없도록 자료에 용어, 개념, 주석 등을 충실하게 수록하고 있는지를 평가. 	접근,7-2
	<ul style="list-style-type: none"> 기술적 정보를 포함한 다양한 정보를 이용할 수 있다. 	<ul style="list-style-type: none"> 필요성; 가급적 전문적이고 기술적인 다양한 정보의 이용이 가능해야 한다. 평가내용; 이용 가능한 전문적이고 기술적인 다양한 정보의 목록과 이들을 이용하는데 관련된 사항들이 상세히 설명되고 있는지를 평가. 	
4.3 투명성	<ul style="list-style-type: none"> 통계를 이해하고 사용하는데 도움을 주는 메타데이터를 제공한다. 	<ul style="list-style-type: none"> 필요성; 생산되고 있는 통계의 종류가 많고 그 내용 또한 방대하므로 이들 통계에 대한 정보인 메타데이터의 제공은 필수적이다. 평가내용; 통계 자료를 이해하는데 도움이 되는 메타데이터가 구축되어 있고, 메타데이터의 구성과 내용 및 사용방법 등이 구체적으로 마련되어 있는지를 평가. 	
	<ul style="list-style-type: none"> 통계생산에 관련된 사항들을 공개한다. 	<ul style="list-style-type: none"> 필요성; 통계정보 공개를 통해서 통계자료에 대한 사용자들의 이해와 생산된 통계에 대한 신뢰를 높여야 한다. 평가내용; 통계작성의 전체적인 과정(표본크기, 추출방법, 추정기법 등)과 생산된 통계의 특성치(표준오차, 변이계수, 신뢰계수 등)를 어떻게 공개하고 있는지를 평가. 	접근,1-9 접근,2-6 접근,2-7 효율,8-4
	<ul style="list-style-type: none"> 통계와 관계되는 모든 사항에 대한 질의 응답체계가 구축되어 있다. 	<ul style="list-style-type: none"> 필요성; 이용의 편리성을 높이기 위해서는 사용자들의 궁금한 사항이나 불편사항들이 즉시에 해소될 수 있어야 한다. 평가내용; 통계와 관련된 모든 사항에 대해 신속하고 편리한 응답체계가 구축되어 실시되고 있는지를 평가. 	접근,8-9

○ 비교성

부차원	평가지표(항목)	필요성 및 평가내용	비고
5.1 시계열 비교	<ul style="list-style-type: none"> 시계열 자료의 불연속 횟수는? 	<ul style="list-style-type: none"> 필요성; 시계열자료는 연속성과 일관성이 필수적이다. 평가내용; 시계열 자료의 불연속 횟수를 표시하고, 개념, 정의, 방법, 기준 등이 동일하게 적용되지 않은 자료에 대한 이용방법이 마련되어 있는지를 평가. 	비교,8-6
	<ul style="list-style-type: none"> 시계열 변동의 영향을 평가한다. 	<ul style="list-style-type: none"> 필요성; 시계열자료의 변동에 대한 영향을 평가해서 이를 반영해야 한다. 	

		<ul style="list-style-type: none"> 평가내용; 시계열 변동에 따른 개념, 정의, 분류, 방법론 등의 영향을 평가해서 반영하는 조치를 취하고 있는지를 평가. 	
5.2 국내외 비교	<ul style="list-style-type: none"> 사용하고 있는 개념, 정의, 체계 등은 국내(국제)적 기준 또는 지침을 따르고 있다. 	<ul style="list-style-type: none"> 필요성; 생산된 통계의 국내(국제) 비교연구를 위해서는 통계작성에 관련된 개념, 정의 등의 기준이 동일해야 한다. 평가내용; 통계생산에 사용되고 있는 개념과 정의, 방법들이 국내(국제)적으로 공통된 것인지, 관련된 국제적 규약을 언급하고 있는지를 평가. 	정확,1-3
	<ul style="list-style-type: none"> 국내(국제)적 기준 또는 지침에 벗어나는 이유를 파악한다. 	<ul style="list-style-type: none"> 필요성; 국내(국제)적 기준이나 지침에서 벗어나는 것들을 파악해서 이용에 오류가 없도록 한다. 평가내용; 국내(국제)적 기준이나 지침에 벗어나는 것들을 분류하고, 그 이유를 파악해서 이용의 오류를 방지하기 위한 조치를 취하고 있는지를 평가. 	
	<ul style="list-style-type: none"> 자료에 대한 국제적 비교가 가능하다 	<ul style="list-style-type: none"> 필요성; 생산된 통계의 국제적 비교가 가능해야 한다. 평가내용; 통계에 대한 전반적인 사항들이(조사개요, 결과분석내용, 통계표, 조사표 등) 영문으로 작성되어 국제적 이용이 가능하고, 국제적 비교가 가능하며 또한 국제적 비교 자료가 수록되어 있는지를 평가. 	비교,8-5
5.3 부차 모집단 비교	<ul style="list-style-type: none"> 부차모집단(도메인 포함)을 비교하고 있다. 	<ul style="list-style-type: none"> 필요성; 전체 모집단을 구성하고 있는 부차모집단(영역별, 집단별 등)들의 비교가 가능해야 한다. 평가내용; 부차모집단을 정의하고, 부차모집단에 관한 통계 생산과정과 (추정)방법들을 구체적으로 언급하여 부차모집단을 비교하고 있는지를 평가. 	

○ 일관성

부차원	평가지표(항목)	필요성 및 평가내용	비고
6.1 통계 작성의 일관성	<ul style="list-style-type: none"> 통계조사시스템을 일관성 있게 유지하고 있다. 	<ul style="list-style-type: none"> 필요성; 통계조사시스템에 일관성이 있어야 통계조사의 오류를 줄이고 통계품질을 일정하게 유지할 수 있다. 평가내용; 체계적인 통계조사시스템이 마련되어 있고 이를 일정하게 유지하는 조치들이 있는지를 평가. 	
	<ul style="list-style-type: none"> 표준적인 목표 특성치들을 일관성 있게 사용한다. 	<ul style="list-style-type: none"> 필요성; 통계자료의 명확한 의미와 통계 간 비교 그리고 시계열적 분석을 위해서는 사용된 목표 특성치들이 표준적이고 일관성이 있어야 한다. 평가내용; 표준적인 목표 특성들(단위, 모집단, 도메인, 변수, 분류, 기준시간)을 일관성 있게 사용하고 있는지를 평가. 	
6.2 통계들 간의 일관성	<ul style="list-style-type: none"> 공표하고 있는 통계를 관련된 유사 통계와 비교분석한다. 	<ul style="list-style-type: none"> 필요성; 동일 기관 또는 다른 기관에서 생산된 유사 통계들 간의 비교가 필요하다. 평가내용; 서로 다른 자료들로부터 얻어지는 유사 추정치들 	비교,6-4

		을 비교하고, 차이점과 이유를 자세히 설명하고 있는지를 평가.	
	◦ 서로 다른 통계(또는 추정치)들을 결합하는 방법 등을 사용하고 있다.	◦ 필요성; 기존의 통계(또는 추정치)들로부터 새로운 통계를 만드는 과정과 방법이 상세히 설명되어야 한다. ◦ 평가내용; 서로 다른 통계들로부터 새로운 통계를 만드는 방법들이 명시되고 새로운 통계의 특성치(추정식, 추정치의 표준오차, 변이계수 등)들이 계산되었는지를 평가.	
	◦ 잠정통계에 대한 수정계획이 마련되어 있다.	◦ 필요성; 잠정적으로 사용되고 있는 잠정통계의 수정계획이 마련되어 잠정통계 이용의 오류를 방지할 수 있다. ◦ 평가내용; 잠정통계에 대한 수정 계획표가 있고 수정 내용, 수정 방법 및 수정 이유들이 상세히 설명되어 있는지를 평가.	
	◦ 잠정통계와 확정통계간의 차이를 설명하고 있다.	◦ 필요성; 잠정통계 이용의 오류를 방지하기 위해서 잠정통계와 최종 확정통계와의 차이를 설명해야 한다. ◦ 평가내용; 잠정통계와 확정통계 간에 차이 나는 항목, 차이 정도, 이유, 차이로 인한 영향 등을 상세히 설명하고 이를 잠정통계 이용자들이 숙지할 수 있도록 조치하였는지를 평가.	

○ 효율성

부차원	평가지표(항목)	필요성 및 평가내용	비고
7.1 기획과 생산 과정	◦ 통계작성을 위한 효율적인 예산산정이 이루어지고 있다.	◦ 필요성; 비용 변화에 따른 작성과정 및 생산된 통계의 효율성을 비교하여 예산을 산정한다. ◦ 평가내용; 비용의 변화에 따른 작성과정의 변화, 생산된 통계의 개선 정도 등 작성통계의 효율성 변화를 측정해서 예산산정에 반영하고 있는지를 평가.	효율,1-4
	◦ 조사방법, 자료처리, 그리고 추정방법 등의 변경(또는 개선)에 따른 효율성을 비교분석한다.	◦ 필요성; 변경된 방법의 타당성을 입증하기 위해서는 효율성분석이 요구된다. ◦ 평가내용; 조사방법, 자료처리방법, 그리고 추정방법 등 조사관련 기법들의 변화에 따른 효율성을 비교분석해서 반영하고 있는지를 평가.	
7.2 결과물 이용	◦ 투입된 비용에 따른 통계자료의 이용 빈도를 측정한다.	◦ 필요성; 투입된 비용 대비 통계자료의 이용 빈도를 측정해서 비용에 따른 이용의 효율성을 측정한다. ◦ 평가내용; 투입된 비용에 따른 이용자들의 이용률 변화와 이의 효과를 측정해서 반영하는 지를 평가.	
	◦ 투입된 비용에 따른 통계자료의 판매비용을 측정한다.	◦ 필요성; 투입된 비용 대비 통계자료의 판매비용을 측정해서 비용에 따른 판매수익을 예측한다. ◦ 평가내용; 투입된 비용에 따른 통계자료의 판매비용의 변화를 측정하고 판매수익을 높이는 전략을 마련해서 실시하고 있는지를 평가.	

5.3 절차별 품질평가지표

현재의 통계품질평가는 통계작성 절차별로 나누어 실시하고 있다. 이는 통계품질평가의 최종 목적이 통계의 품질을 향상시키는데 있으므로, 실질적인 품질향상을 위한 조치는 차원보다는 절차별로 관리하는 것이 좀더 편리하고 효율성이 있기 때문이다. 따라서 5.2절의 차원별 통계품질평가지표를 통계작성 절차로 분류하였다. 물론 최종평가는 차원별로 하는 것이 바람직하다. <표 5-3>에서는 81개의 새로운 품질평가지표를 현재 통계청에서 분류한 8개의 절차별로 분류하였다.

<표 5-3> 절차별 품질평가지표

절차	평가지표	차원	구분
1. 조사기획 (15개)	◦ 이용자들에 대한 체계적인 모니터링 프로그램이 있다.	적절성	필수
	◦ 이용자들의 요구에 대한 사업계획과 우선순위를 정한다.	적절성	
	◦ 이용자(만족도) 조사결과를 반영한다.	적절성	
	◦ 이용자 요구에 맞는 통계 개념을 사용한다.	적절성	필수
	◦ 개념간의 차이를 기술하고 있다.	적절성	
	◦ 결과물에 대한 이용자들의 의견을 반영하고 있다.	적절성	필수
	◦ 표본설계에 따른 진행과정을 점검하고 문제점등을 파악하여 조치하는 시스템을 운영하고 있다.	정확성	필수
	◦ 자료수집이 주기적으로 이루어지고 있다.	시의성	필수
	◦ 개편작업이 정기적으로 이루어지고 있다.	시의성	필수
	◦ 사용하고 있는 개념, 정의, 체계 등은 국내(국제)적 기준 또는 지침에 따르고 있다.	비교성	필수
	◦ 국내(국제)적 기준 또는 지침에 벗어나는 이유를 파악한다.	비교성	필수
	◦ 통계조사시스템을 일관성 있게 유지하고 있다.	일관성	필수
	◦ 표준적인 목표 특성치들을 일관성 있게 사용한다.	일관성	필수
	◦ 잠정통계에 대한 수정계획이 마련되어 있다.	일관성	
	◦ 통계작성을 위한 효율적인 예산산정이 이루어지고 있다.	효율성	필수
2. 모집단 및 표본 (4개)	◦ 조사를 위한 표본설계를 한다.	정확성	필수
	◦ 목표(연구)모집단과 조사모집단을 정의하고 비교하고 있다.	정확성	필수
	◦ 표본추출틀에 대한 내용을 설명하고 변동사항을 반영해 준다.	정확성	필수
	◦ 다양한 자료수집 방법들을 비교·분석 한다.	정확성	
3. 조사표 설계 (2개)	◦ 설문지의 적합성을 평가하는 시스템을 갖추고 있다.	적절성	필수
	◦ 사전조사를 실시한다.	정확성	필수
4. 조사직원 관리 (2개)	◦ 조사원에 대한 교육을 실시한다.	정확성	필수
	◦ 조사 직원의 조사관련 전문지식 숙지 여부를 파악하고, 이를 평가하는 제도적 장치가 있다.	정확성	필수
5. 조사실시 (4개)	◦ 응답자들의 응답부담을 측정하고 줄이는 조치를 취하고 있다.	정확성	필수
	◦ 응답자들의 비밀보호를 위해 적극 노력하고 있다.	정확성	필수

	◦ 체계적인 현장조사 관리시스템이 구축되어있다.	정확성	필수
	◦ 응답거절, 결측 및 부적격 응답에 대한 교체방법을 제시하고 있다.	정확성	필수
6. 자료처리 및 집계 (27개)	◦ 통계자료에 대한 오차(표본, 비표본)를 파악하고 있다.	정확성	필수
	◦ 주요 추정치들에 대한 표본분산, 분산추정치, 그리고 변이계수 등을 계산한다.	정확성	필수
	◦ 설계효과를 계산한다.	정확성	
	◦ 비표본오차의 주요 원인들을 파악하여 조치해 주고 있다.	정확성	필수
	◦ 주요 항목들에 대한 대체율을 계산한다.	정확성	
	◦ 포함오차(과소, 과대, 중복 등)를 계산한다.	정확성	
	◦ 포함오차의 영향과 문제 해결 방안을 제시하고 있다.	정확성	
	◦ 무응답오차의 원인을 파악하여 조치한다.	정확성	필수
	◦ 무응답률(단위, 항목)을 파악하고 무응답에 따른 가중치를 조정해 주고 있다.	정확성	
	◦ 응답집단과 무응답집단의 특성을 비교·분석한다.	정확성	
	◦ 무응답 및 대체효과를 평가한다.	정확성	
	◦ 측정오차의 원인을 파악하여 조치한다.	정확성	필수
	◦ 편집오류율을 측정한다.	정확성	
	◦ 대리응답률을 측정한다.	정확성	
	◦ 조사원간의 차의 특성을 측정한다.	정확성	
	◦ 처리오차의 원인을 파악하여 조치한다.	정확성	필수
	◦ 자료처리 시스템을 운영한다.	정확성	필수
	◦ 입력오류율을 계산한다.	정확성	
	◦ 입력자간의 차의 특성을 파악한다.	정확성	
	◦ 사용한 모형과 추출절차를 명시하고 있다.	정확성	필수
	◦ 이상치에 대한 조치를 취하고 있다.	정확성	
	◦ 모형에 의한 오차(편향, 분산)를 계산한다.	정확성	
	◦ 추정, 가중치부여, 추정방법 등에 대한 적용과정을 비교 한다.	정확성	
	◦ 시계열 변동의 영향을 평가한다.	비교성	
	◦ 공표하고 있는 통계를 관련된 유사 통계와 비교분석한다.	일관성	필수
	◦ 서로 다른 통계(또는 추정치)들을 결합하는 방법 등을 사용하고 있다.	일관성	
	◦ 조사방법, 자료처리, 그리고 추정방법 등의 변경(또는 개선)에 따른 효율성을 비교분석한다.	효율성	필수
7. 자료공표 (9개)	◦ 공표하기 전에 현장조사부터 집계분석 단계까지 적절한 내용검토 절차가 있다.	정확성	
	◦ 간행물 발간 이전에, 간행물에 대한 사전점검체계가 있다.	정확성	필수
	◦ 통계자료의 공표를 주기적으로 한다.	시의성	필수
	◦ 통계작성시의 기준시점과 공표시점과의 차이는 얼마나 되는가?(임시결과물과 최종결과물 모두에 대해)	시의성	필수
	◦ 주문형 통계의 경우 요청시점과 전달시점과의 차이는 얼마나 되는가?	시의성	
	◦ 새로운(또는 특별)조사의 착수시점과 공표시점간의 차이는 얼마나 되는가?	시의성	

	◦ 조사결과 자료 공표시기로부터 DB수록 시기까지의 소요 시간	시의성	
	◦ 통계자료의 공표계획 시점과 공표시점과의 차이는 얼마나 되는가?	시의성	필수
	◦ 온라인상으로 통계자료에 대한 정보를 제공한다.	접근성	필수
8. 자료이용 (18개)	◦ 결과물의 주요 이용자들을 분류하고 있다.	적절성	
	◦ 결과물에 대한 이용자 만족도 조사를 한다.	적절성	
	◦ 통계자료와 관련 내용들을 여러 매체(책자, 디스켓, CD-ROM 등)로 제공한다.	접근성	필수
	◦ 원시자료를 포함한 마이크로자료에 대한 자세한 사용방법을 기술한다.	접근성	필수
	◦ 미공표(non-published)자료에 대한 자세한 사용방법을 기술	접근성	
	◦ 결과물을 검색엔진을 통해 검색할 수 있다.	접근성	필수
	◦ 조사결과 자료에 대한 DB가 구축되어 신속하고 편리한 자료 이용이 가능하다.	접근성	필수
	◦ 관련 용어, 개념, 주석 등을 간행물에 충실하게 수록하여 이용자의 이해를 돕는다.	접근성	필수
	◦ 기술적 정보를 포함한 다양한 정보를 이용할 수 있다.	접근성	
	◦ 통계를 이해하고 사용하는데 도움을 주는 메타데이터를 제공한다.	접근성	필수
	◦ 통계생산에 관련된 사항들을 공개한다.	접근성	필수
	◦ 통계와 관계되는 모든 사항에 대한 질의 응답체계가 구축되어 있다.	접근성	필수
	◦ 시계열 자료의 불연속 횟수는?	비교성	필수
	◦ 자료에 대한 국제적 비교가 가능하다	비교성	
	◦ 부차모집단(도메인 포함)을 비교하고 있다.	비교성	
◦ 잠정통계와 확정통계간의 차이를 설명하고 있다.	일관성	필수	
◦ 투입된 비용에 따른 통계자료의 이용 빈도를 측정한다.	효율성	필수	
◦ 투입된 비용에 따른 통계자료의 판매비용을 측정한다.	효율성		

5.4 통계별 품질평가지표

통계청에서 생산하고 있는 승인통계는 통계작성 방법에 따라 조사통계 42종, 보고통계 1종, 그리고 가공통계가 9종으로 조사통계가 주를 이루고 있다(참고, <표 1-1>). 본 절에서는 가공통계와 조사통계에서 전수조사를 별도로 분류해서 통계품질평가지표를 작성하였다. 5.1절의 통계품질평가지표는 조사통계에서 표본조사의 경우이고, 이들을 기준으로 해서 가공통계와 전수조사에 의한 통계의 경우로 나누었다. 대부분의 통계품질평가지표가 중복되어서 지표만을 분류하였고, 전수조사통계와 가공통계의 경우 추가사항이나 변동이 있는 경우는 지표의 필요성과 평가내용을 첨부하였다.

5.4.1 전수조사통계

<표 5-4>에서는 전수조사통계에 대한 품질평가지표를 차원별로 분류하였고 <표 5-5>에서는 추가되거나 변경된 필요성과 평가내용을 살펴보았다.

<표 5-4> 전수조사통계 품질평가지표(62개 항목)

차원	부차원	평가지표(항목)	비고	구분
1. 적절성 (9)	1.1 이용자 요구	◦ 이용자들에 대한 체계적인 모니터링 프로그램이 있다.		필수
		◦ 이용자들의 요구에 대한 사업계획과 우선순위를 정한다.		
		◦ 이용자(만족도) 조사결과를 반영한다.	관련,1-5	
	1.2 통계개념 정의	◦ 이용자 요구에 맞는 통계개념을 사용한다.	관련,1-1	필수
		◦ 개념간의 차이를 기술하고 있다.		
	1.3 설문지작성	◦ 설문지의 적합성을 평가하는 시스템을 갖추고 있다.	정확,3-1 정확,3-2 정확,3-4	필수
	1.4 결과물	◦ 결과물의 주요 이용자들을 분류하고 있다.		
◦ 결과물에 대한 이용자 만족도 조사를 한다.				
◦ 결과물에 대한 이용자들의 의견을 반영하고 있다.			필수	
2. 정확성 (22)	2.1 일반사항	◦ 응답자들의 응답부담을 측정하고 줄이는 조치를 취하고 있다.		필수
		◦ 응답자들의 비밀보호를 위해 적극 노력하고 있다.	정확,8-8	필수
		◦ 공표하기 전에 현장조사부터 집계분석 단계까지 적절한 내용검토 절차가 있다.	정확,6-2	
		◦ 간행물 발간 이전에, 간행물에 대한 사전검체체계가 있다.	정확,7-1	필수
	2.2 비표본오차	◦ 비표본오차의 주요 원인들을 파악하여 조치해 주고 있다.		필수
		◦ 체계적인 현장조사 관리시스템이 구축되어있다.	정확,5-1 정확,5-2 정확,5-3 정확,5-4	필수
		◦ 주요 항목들에 대한 대체율을 계산한다.		필수
	2.2.1 포함오차	◦ 목표(연구)모집단과 조사모집단을 정의하고 비교하고 있다.		필수
		◦ 조사목록에 대한 포함오차(과소, 과대, 중복 등)를 계산한다.		필수
		◦ 조사목록의 포함오차 영향과 문제 해결 방안을 제시하고 있다.		필수
	2.2.2 무응답오차	◦ 무응답오차의 원인을 파악하여 조치한다.	정확,6-1	필수
		◦ 응답집단과 무응답집단의 특성을 비교·분석한다.		
		◦ 무응답 및 대체효과를 평가한다.		
	2.2.3 측정오차	◦ 측정오차의 원인을 파악하여 조치한다.		필수
◦ 사전조사를 실시한다.		정확,3-3	필수	

		◦ 조사원에 대한 교육을 실시한다.	정확,4-1 정확,4-2 정확,3-5	필수
		◦ 임시조사원의 선발, 교육 및 관리가 체계적이다.		필수
	2.2.4 처리오차	◦ 처리오차의 원인을 파악하여 조치한다.		필수
		◦ 자료처리 시스템을 운영한다.		필수
		◦ 편집오류율을 측정한다.		
		◦ 입력오류율을 계산한다.		
		◦ 입력자간의 차의 특성을 파악한다.		
3. 시의성 (5)	3.1 주기성	◦ 자료수집이 주기적으로 이루어지고 있다.		필수
		◦ 통계자료의 공표를 주기적으로 한다.		필수
	3.2 생산시간	◦ 통계작성시의 기준시점과 공표시점과의 차이는 얼마나 되는가?(잠정결과물과 최종결과물 모두에 대해) ◦ 조사결과 자료 공표시기로부터 DB수록 시기까지의 소요 시간	접근,8-3	필수
	3.3 정시성	◦ 통계자료의 공표계획 시점과 공표시점과의 차이는 얼마나 되는가?	시의,7-5 시의,7-6	필수
4. 접근성 (11)	4.1 보급	◦ 온라인상으로 통계자료에 대한 정보를 제공한다.	접근,7-4	필수
		◦ 통계자료와 관련 내용들을 여러 매체(책자, 디스켓, CD-ROM 등)로 제공한다.	접근,8-1	필수
	4.2 정보이용 및 검색지원	◦ 원시자료를 포함한 마이크로자료에 대한 자세한 사용방법을 기술한다.	접근,8-7	필수
		◦ 미공표(non-published)자료에 대한 자세한 사용방법을 기술		
		◦ 결과물을 검색엔진을 통해 검색할 수 있다.		필수
		◦ 조사결과 자료에 대한 DB가 구축되어 신속하고 편리한 자료 이용이 가능하다.	접근,8-2	필수
		◦ 관련 용어, 개념, 주석 등을 간행물에 충실하게 수록하여 이용자들의 이해를 돕는다.	접근,7-2	필수
		◦ 기술적 정보를 포함한 다양한 정보를 이용할 수 있다.		
	4.3 투명성	◦ 통계를 이해하고 사용하는데 도움을 주는 메타데이터를 제공한다.		필수
		◦ 통계생산에 관련된 사항들을 공개한다.	접근,1-9 접근,2-6 접근,2-7 효율,8-4	필수
		◦ 통계와 관계되는 모든 사항에 대한 질의 응답체계가 구축되어 있다.	접근,8-9	필수
5. 비교성 (6)	5.1 시계열 비교	◦ 시계열 자료의 불연속 횟수는?		필수
		◦ 시계열 변동의 영향을 평가한다.	비교,8-6	
	5.2 국내외 비교	◦ 사용하고 있는 개념, 정의, 체계 등은 국내(국제)적 기준 또는 지침을 따르고 있다.	정확,1-3	필수
		◦ 국내(국제)적 기준 또는 지침에 벗어나는 이유를 파악한다. ◦ 자료에 대한 국제적 비교가 가능하다	비교,8-5	필수

	5.3 부차모집단 비교	◦ 부차모집단(도메인 포함)을 비교하고 있다.		
6. 일관성 (5)	6.1 통계작성의 일관성	◦ 통계조사시스템을 일관성 있게 유지하고 있다.		필수
		◦ 표준적인 목표 특성치들을 일관성 있게 사용한다.		필수
	6.2 통계들 간의 일관성	◦ 공표하고 있는 통계를 관련된 유사통계와 비교·분석한다.	비교,6-4	필수
		◦ 잠정통계에 대한 수정계획이 마련되어 있다.		
		◦ 잠정통계와 확정통계간의 차이를 설명하고 있다.		필수
7. 효율성 (4)	7.1 기획과 생산과정	◦ 통계작성을 위한 효율적인 예산산정이 이루어지고 있다.	효율,1-4	필수
		◦ 조사방법, 자료처리, 그리고 추정방법 등의 변경(또는 개선)에 따른 효율성을 비교분석한다.		필수
	7.2 결과물 이용	◦ 투입된 비용에 따른 통계자료의 이용 빈도를 측정한다.		필수
		◦ 투입된 비용에 따른 통계자료의 판매비용을 측정한다.		

<표 5-5> 전수조사통계 품질평가지표의 필요성과 평가내용

○ 정확성

부차원	평가지표(항목)	필요성 및 평가내용
2.2 비 표 본 오 차	2.2.1 포함 오차	◦ 조사목록에 대한 포함오차(과소, 과대, 중복 등)를 계산한다. ◦ 필요성; 표본추출틀이 이용되는 표본조사와 달리 전수조사에서는 전체 조사단위들에 대한 조사목록이 사용된다. 완전한 조사목록은 현실적으로 불가능하므로 여러 가지 포함오차들이 발생한다. 이들 포함오차들은 파악되고 관리되어야 한다. ◦ 평가내용; 조사목록의 과대, 과소, 중복, 부적격과 오분류 정도가 어떻게 측정되고 있는지를 평가.
		◦ 조사목록의 포함오차 영향과 문제 해결 방안을 제시하고 있다. ◦ 필요성; 포함오차의 정도에 따라 집계 결과에 오류가 생기므로 이의 영향을 측정해서 해결방안을 마련해야 한다. ◦ 평가내용; 포함오차(과소, 과대, 중복 등)에 따라 집계 결과에서 발생하는 문제들을 파악하고 이들 포함오차를 줄이는 방안들을 마련하고 있는지를 평가.
	2.2.3 측정 오차	◦ 임시조사원의 선발, 교육 및 관리가 체계적이다. ◦ 필요성; 많은 임시조사원들을 필요로 하는 전수조사에서는 임시조사원의 선발, 교육 및 관리가 조사의 품질에 절대적으로 큰 영향을 미치므로 이들 과정들이 체계적으로 유지되어야 한다. ◦ 평가내용; 유능한 임시조사원의 선정, 선발된 조사원들에 대한 효율적이고 내실 있는 교육 그리고 조사현장에서의 효과적인 조사원 관리 등이 체계적으로 실시되고 있는지를 평가.

5.4.2 가공통계

<표 5-6>과 <표 5-7>에서는 각각 가공통계에 대한 품질평가지표와 품질평가지표에 대한 추가, 또는 변경된 필요성과 평가내용을 살펴보았다.

<표 5-6> 가공통계 품질평가지표(41개 항목)

차원	부차원	평가지표(항목)	비고	구분	
1. 적절성 (8)	1.1 이용자 요구	◦ 이용자들에 대한 체계적인 모니터링 프로그램이 있다.		필수	
		◦ 이용자들의 요구에 대한 사업계획과 우선순위를 정한다.			
		◦ 이용자(만족도) 조사결과를 반영한다.	관련,1-5		
	1.2 통계개념 정의	◦ 이용자 요구에 맞는 통계 개념을 사용한다.	관련,1-1	필수	
		◦ 개념간의 차이를 기술하고 있다.			
	1.3 결과물	◦ 결과물의 주요 이용자들을 분류하고 있다.			
◦ 결과물에 대한 이용자 만족도 조사를 한다.					
◦ 결과물에 대한 이용자의 의견을 반영하고 있다.			필수		
2. 정확성 (5)	2.1 일반사항	◦ 가공통계의 원자료인 기존통계에 대한 정확성을 파악하기 위한 주기적인 검토 체계가 있다.		필수	
		◦ 간행물 발간 이전에, 간행물에 대한 사전점검체계가 있다.	정확,7-1	필수	
	2.2 비표본오차	2.2.1 처리오차	◦ 처리오차의 원인을 파악하여 조치한다.		필수
		2.2.2 모형오차	◦ 가공통계 생산을 위해 사용되는 모형(작성 방법)과 이와 관련한 가정들(이론들)이 적절한지 주기적으로 검토한다.		필수
			◦ 가공통계 중 추계(예측)통계와 실제 조사가 실시된 후에 공표되는 통계 간의 차이를 설명하고 있다.		필수
3. 시의성 (4)	3.1 주기성	◦ 통계자료의 공표를 주기적으로 한다.		필수	
	3.2 생산시간	◦ 통계작성시의 기준시점과 공표시점과의 차이는 얼마나 되는가?		필수	
		◦ 결과 자료 공표시기로부터 DB수록 시기까지의 소요 시간	접근,8-3		
	3.3 정시성	◦ 통계자료의 공표계획 시점과 공표시점과의 차이는 얼마나 되는가?	시의,7-5 시의,7-6	필수	
4. 접근성 (11)	4.1 보급	◦ 온라인상으로 통계자료에 대한 정보를 제공한다.	접근,7-4	필수	
		◦ 통계자료와 관련 내용들을 여러 매체(책자, 디스켓, CD-ROM 등)로 제공한다.	접근,8-1	필수	
	4.2 정보이용 및 검색지원	◦ 원시자료를 포함한 마이크로자료에 대한 자세한 사용방법을 기술한다.	접근,8-7	필수	
		◦ 미공표(non-published)자료에 대한 자세한 사용방법을 기술			
	◦ 결과물을 검색엔진을 통해 검색할 수 있다.		필수		

		◦ 결과 자료에 대한 DB가 구축되어 신속하고 편리한 자료 이용이 가능하다.	접근,8-2	필수
		◦ 관련 용어, 개념, 주석 등을 간행물에 충실하게 수록하여 이용자들의 이해를 돕는다.	접근,7-2	필수
		◦ 기술적 정보를 포함한 다양한 정보를 이용할 수 있다.		
	4.3 투명성	◦ 통계를 이해하고 사용하는데 도움을 주는 메타데이터를 제공한다.		필수
		◦ 통계생산에 관련된 사항들을 공개한다.		필수
		◦ 통계와 관계되는 모든 사항에 대한 질의 응답체계가 구축되어 있다.	접근,8-9	필수
5. 비교성 (6)	5.1 시계열 비교	◦ 시계열 자료의 불연속 횟수는?	비교,8-6	필수
		◦ 시계열 변동의 영향을 평가한다.		
	5.2 국내외 비교	◦ 사용하고 있는 개념, 정의, 체제 등은 국내(국제)적 기준 또는 지침을 따르고 있다.	정확,1-3	필수
		◦ 국내(국제)적 기준 또는 지침에 벗어나는 이유를 파악한다.		필수
	5.3 부차모집단 비교	◦ 자료에 대한 국제적 비교가 가능하다	비교,8-5	
6. 일관성 (3)	6.1 통계작성의 일관성	◦ 통계조사시스템을 일관성 있게 유지하고 있다.		필수
		◦ 표준적인 목표 특성치들을 일관성 있게 사용한다.		필수
	6.2 통계들 간의 일관성	◦ 공표하고 있는 통계를 관련된 유사통계와 비교·분석한다.	비교,6-4	필수
7. 효율성 (4)	7.1 기획과 생산과정	◦ 통계작성을 위한 효율적인 예산산정이 이루어지고 있다.	효율,1-4	필수
		◦ 자료처리와 추정방법 등의 변경(또는 개선)에 따른 효율성을 비교·분석한다.		필수
	7.2 결과물 이용	◦ 투입된 비용에 따른 통계자료의 이용 빈도를 측정한다.		필수
		◦ 투입된 비용에 따른 통계자료의 판매비용을 측정한다.		

<표 5-7> 가공통계 품질평가지표의 필요성과 평가내용

○ 정확성

부차원	평가지표(항목)	필요성 및 평가내용
2.1 일반사항	◦ 가공통계의 원자료인 기존통계에 대한 정확성을 파악하기 위한 주기적인 검토 체계가 있다.	◦ 필요성; 가공통계는 각 통계기관에서 작성된 기존 통계 자료를 재분류하거나 각종 통계적 기법을 사용하여 재생산되므로 이용 전에 기존 통계의 정확성 검토는 반드시 필요하다. ◦ 평가내용; 각 통계기관에서 생산된 기존 통계들을 조사통계, 가

		공통계, 행정기록 등으로 분류하여 표본오차와 비표본 오차 등을 포함한 해당 통계자료의 정확성을 주기적으로 검토하는지 평가.
2.2 비 표 본 오 차	2.2.1 처리 오차	<ul style="list-style-type: none"> 필요성; 자료처리 과정에서 발생하는 오차를 원인별로 파악해서 오차를 줄이도록 한다. 평가내용; 기존 통계 자료에 대한 재분류 및 변환 등의 과정에서 발생할 수 있는 오차를 측정하고 이 오차를 줄이기 위한 조치를 취하고 있는지를 평가
	2.2.2 모형 오차	<ul style="list-style-type: none"> 필요성; 가공통계는 기존 자료에 대한 모형들과 여러 가지 가정(assumptions)을 기초로 생산되므로 이들을 주기적으로 검토하고 개선시켜야 한다. 평가내용; 통계 작성 시 사용되는 모형과 가정들의 적절성을 지속적으로 파악하는지를 평가
		<ul style="list-style-type: none"> 필요성; 추계를 목적으로 하는 가공통계는 일정 시간이 지난 후 실제 조사되는 통계와 비교가 가능하므로 두 통계 간에 차이에 대한 설명이 필요하다. 평가내용; 추계통계와 일정 시점 후 실제로 조사된 통계 간의 차이가 발생한 원인과 추계통계의 보완조치가 마련되어 있는지를 평가.

○ 접근성

부차원	평가지표(항목)	필요성 및 평가내용
4.3 투명성	<ul style="list-style-type: none"> 통계생산에 관련된 사항들을 공개한다. 	<ul style="list-style-type: none"> 필요성; 가공통계는 주로 기존 통계 자료의 복잡한 재분류 또는 통계적 방법론을 통해 생산되므로 이용자들이 이들 통계들을 적극적으로 활용할 수 있도록 이 내용들이 상세히 공개되어야 한다. 평가내용; 가공통계의 작성 방법 및 절차 등이 이용자들이 충분히 이해할 수 있도록 공개되고 있는지 평가.

제 6장 통계품질지수

각 통계의 품질 수준을 나타내는 종합적인 품질지수에 대한 연구는 거의 없는 실정이다. 특히 과정에 대한 품질을 측정하는 지표는 과정별로 TQM에 의해 거론되고 있지만, 결과물에 대한 품질수준의 측도에 대한 연구는 찾아보기 힘들다. 여기 소개된 논문은 네덜란드 통계청에서 발간된 것으로 결과물에 대한 품질을 측정하는 하나의 방법을 제시하고 있다. 측정의 어려움과 제안된 측도의 복잡성 그리고 그에 따른 적용의 어려움 등을 감안하더라도 충분히 연구 검토의 필요성이 있기에 세밀한 재검토를 통해 국내 통계에서의 적용 가능성을 검토해 본다. 또한 적용하기에 수월한 통계품질지수를 제안한다.

6.1 네덜란드 품질지수

본 절에서는 네덜란드 통계청에서 연구 검토한 결과물에 대한 통계품질지수를 살펴본다 (Booleman과 Brakel, 1999). 이를 통해서 우리나라에 적용 가능한 형태로의 변화를 살피고 또한 최근 널리 사용되고 있는 6시그마 기법을 도입하여 통계품질을 측정하는 방법을 제안한다.

6.1.1 요약

- (1) 제안된 품질지수(QI; Quality Index)는 결과물 중심의(Output-oriented) 품질지수임(절차 중심의(process-oriented) 품질지수가 아님)
- (2) 'the sub-dimension level of the accuracy dimension'은 다음과 같은 공식으로 계산한다.

- RMRQD(the root of the mean relative quadratic deviation from the LEVEL)

$$RQD_s = \sum_{k=1}^{H_k} \frac{(V_{ks} - D_k)^2}{D_k^2} \times 100^2$$

$$\Rightarrow RMRQD = \sqrt{\frac{\sum_{s=1}^S RQD_s}{\sum_{s=1}^S H_s}}$$

- (3) RQD_s 와 $RMRQD$ 의미

① D_k 는 부차원(sub-dimension) level 평가의 중요변수 h (key variables)에 대한 최종 확정 발표된 값(definitive figure)을 의미한다.

- ② V_{ks} 는 중요변수에 대한 최종값 이외의 다른 통계 s 에서 중요변수 k 의 통계를 의미한다. 여기서는 잠정적인 결과(provisional results) 혹은 다른 유사기관에서 발표된 다른 통계가 된다.
- ③ RQD_s 는 D_k 와 V_{ks} 의 편차의 제곱합이다. 특히 이 값은 D_k 를 기준으로 한 상대제곱편차(relative quadratic deviation)의 합이다.
- ④ $RMRQD$ 는 RQD_s 의 평균제곱근을 의미하며, 단위는 %이다.
- (4) 'the sub-dimension trend of the accuracy dimension'은 다음과 같은 공식으로 계산한다.

- RMQD(the root of the mean quadratic deviation from the TREND)

$$QD_s = \sum_{k=1}^{H_s} \left(\frac{V_{ks}(t) - V_{ks}(t-1)}{V_{ks}(t-1)} - \frac{D_k(t) - D_k(t-1)}{D_k(t-1)} \right)^2 \times 100^2$$

$$\Rightarrow RMQD = \sqrt{\frac{\sum_{s=1}^S QD_s}{\sum_{s=1}^S H_s}}$$

(5) QD_s 의미

QD_s 는 고려된 각 통계별로 시점 t 와 시점 $t-1$ 에서의 trend의 제곱차이(quadratic deviation)를 의미한다.

(6) 품질지수의 계산

① 품질지수는 앞의 공식에 의한 계산 값(정확성의 level과 trend)과 sector standard 값을 고려하여 다양한 계산 공식(Booleman과 Brakel,1999, pp.97-98)에 의해 부차원(sub-sub dimension)에서의 값이 구해지고, 이것들을 이용하여 가중치와 더불어 가중평균을 구하여 부차원과 차원에서의 지수로 사용한다. 최종 품질지수 역시 차원의 가중평균치를 다시 가중평균한 값으로 계산한다.

② 이러한 다양한 계산방법은 품질지수의 값을 '100'을 기준값으로 사용하고 있다. 즉, '100'을 기준으로 작은 값은 품질수준이 기대(sector standard)보다 '나빠진 상태'를, 큰 값은 '좋아진 상태'를 의미한다.

③ 예를 들어, 'the sub-dimension level of the accuracy dimension'의 값(realization) $RMRQD$ 이 sector standard보다 커질수록 기대보다 "나빠진 상태"를 의미하며(사익성 역시 마찬가지로의 경우) '100이하'의 값을 갖게 되고, 'the sub-dimension trend of the accuracy dimension'의 값(realization) $RMQD$ 이 sector standard보다 커질

수록 기대보다 “좋아진 상태”를 의미하여(일관성과 정보의 수준도 마찬가지로 경우) ‘100 이상’의 값을 갖게 되도록 공식이 만들어져 있다. 그 결과 제안된 공식은 다음과 같다.

$$\text{level과 시의성} \quad ==> \quad \frac{\text{sector standard}}{\text{realisation}} \times 100$$

$$\text{trend와 일관성} \quad ==> \quad \left(\frac{\text{sector standard}}{\text{realisation}} \right)^{-1}$$

그리고 정보의 수준 차원에서의 ‘SN publishes one figure’의 값은 ‘yes=1/no=0’의 값을 갖기 때문에 비(ratio) 형태가 아닌 차이(difference) 형태를 사용하여 지표를 계산한다. 즉, $100 + 100 \times (\text{realisation} - \text{sector standard})$

④ 정확성의 지수 계산 방법

중요변수의 추정값과 실제값의 차이 혹은 다른 기관에서의 평가 결과와의 차이에 대한 가중평균값으로 평가

⑤ 일관성의 지수 계산 방법

설문지 결과로 평가 점수와 sector standard와의 차이로 평가.(10점 척도)

⑥ 시의성의 지수 계산 방법

계획된 일정과 결과의 차이로 평가

⑦ 정보의 수준에 대한 지수 계산 방법

‘five most important users’에 대한 설문조사 결과와 sector standard와의 차이로 평가(10점 척도)

⑧ 최종 품질지수 계산 방법

네 개의 차원에 대한 가중평균값으로 계산한다.

6.1.2 사례연구

(structural business statistics in the construction industry for 1997)

(1) 요약

- ① 1998년 12월에 첫 번째 잠정적인 결과(provisional results)를 발표
- ② 1999년 4월에 확정결과(definitive and more detailed results)를 발표
- ③ PS 97은 SN의 GBR(SN's General Business Register)에 따라 정보를 제공하는 역할.
- ④ PS 97에는 GBR에 따라 ‘all larger companies’를 나타내는 ‘class 7-9’와, ‘remaining companies’를 나타내는 ‘class 1-6’이 있다. 또한 종업원이 없는 회사들을 ‘class 0’로 분류하였다.

- ⑤ 정확성에 대한 중요변수로는 GP(gross profit), PE(Person Employed), LC(Labour Costs), CV(Customer Value), VA(Value Added)로 결정.
- ⑥ 정확성에서 trend는 1996년과 1997년의 통계를 기준으로 한다.
- ⑦ MEB는 건설부(CD)에서 월간조사를 통해 예측한 통계값.
- ⑧ SN에서의 다른 통계(LA ; Labour Account)를 통하여 PE와 LC의 통계가 발표됨.
- ⑨ Labour Force Survey(EBB)에 PE의 통계가 생산됨.
- ⑩ NA(National Account)에서 위의 다섯 개의 output이 생산됨.

(2) 계산방법

A. 정확성

<표 6-1> 정확성에 대한 지수

Dimension	①weighting (%)	②sector standard(%)	③realization (%)	index
Accuracy	30			⑩84
level	30			⑧46
vs. pro.	30	1	④1.93	⑥52
vs. IF	50	1	3.22	31
vs. other	20	3	4.07	74
trend	70			⑨100
vs. pro.	30	1	⑤1.30	⑦100
vs. IF	50	1	0.61	100
vs. other	20	3	2.04	100

- ① -> 가중치로 각 차원의 중요도에 따라 결정 가능. 통계마다 가중치는 다를 수 있다
- ② -> 기준값으로 각 통계담당 부서에서 차원의 기준값(표준값 혹은 목표값)
- ③ -> realization으로 현재 통계로부터의 값. 예를 들면, ④의 값은 RMRQD의 값을 의미 하고, ⑤의 값은 RMQD의 값을 나타낸다.
- ④ -> 이 값은 먼저 주요변수에 대해 건설산업통계 자료인 PS에 수록된 자료 중 잠정통 계와 월간 조사통계 MEB로부터 추계된 자료를 원자료로 사용하여 RQD_s 를 구하

고 이들의 평균의 제곱근인 $RMRQD = \sqrt{\frac{33.51}{9}} = 1.93$ 으로부터 구한다. 그 아래 값 3.22는 1997년의 PS의 확정통계와 NA와 LA(Integration Framework)로 부터의 통계를 이용하여 RQD_s 를 구한 후에 같은 방식으로 얻어진 값이다. 그 밑

의 값 4.07는 EBB의 값으로부터 같은 방식으로 얻어졌다.

⑤ -> 이 값은 PS의 확정통계와 PS의 잠정통계 혹은 MEB로부터의 통계를 이용하여 QD

를 구한 후 얻어진 RMQD 값이다. 즉, $RMQD = \sqrt{\frac{10.20}{6}} = 1.30$ 이다. 그 밑의 값인 0.61은 1996년부터 97년까지의 IF(integration framework)로부터 얻어진 통계를 이용하여 구한 값이고, 그 밑의 값 2.04는 EBB 통계를 이용하여 같은 공식으로 구한 것이다.

⑥ -> 잠정적인 값과의 level 비교를 의미하는 지표로서 다음 식으로부터 구해진다.

$$\frac{\text{sector standard}}{\text{realisation}} \times 100 \implies \frac{1}{1.93} \times 100 = 51.8 \approx 52$$

그 밑의 31과 74도 같은 식으로 구한다.

⑦ -> 잠정적인 값과의 trend 비교를 의미하는 지표로서 다음 식으로부터 구해진다.

$$\left(\frac{\text{sector standard}}{\text{realisation}}\right)^{-1} \implies \frac{1.30}{1} = 1.30 < 1.5 \implies '100'$$

그 밑의 값 두 개의 '100' 역시 같은 방법으로 구한다.

⑧ -> 이 값은 그 값 아래의 세 값(52, 31 그리고 74)의 값에 대한 가중평균을 의미한다.

즉,

$$\frac{52 \times 30 + 31 \times 50 + 74 \times 20}{100} = \frac{4590}{100} \approx 46$$

⑨ -> 이 값 역시 그 값 아래 세 개의 값에 대한 가중평균이다.

⑩ -> 두 개의 부차원(level과 trend)에 대한 가중평균이다. 즉,

$$\frac{46 \times 30 + 100 \times 70}{100} = 83.8 \approx 84$$

B. 일관성

<표 6-2> 일관성에 대한 지수

Dimension	weighting (%)	sector standard	realization	index
Coherence	20			③104
vs. pro.	30	9	①8.86	② 95
vs. IF	50	8	8.57	107
vs. other	20	7	7.86	112

- ① -> 먼저 확정통계와 잠정통계에 대한 일관성 관련 7개의 설문에 대한 가중치와 점수 (10점 척도)를 이용하여 최종 점수를 가중평균값으로 구하고, 또한 확정통계와 MEB 통계와에 대한 일관성 관련 7개 설문의 가중평균의 값을 구한다. 그리고 이 두 개의 값의 평균이 위의 값 8.86이 된다. 즉,

$$\frac{10 + 7.71}{2} = 8.855 \approx 8.86$$

아래의 두 값 8.57과 7.86은 위와 같은 방식으로 구한 '확정통계와 NA와 LA 값'과의 가중평균값과 '확정통계와 EBB 값'과의 가중평균을 나타낸다.

- ② -> 지표를 의미하며, 다음과 같은 공식을 이용하여 구한다.

$$\left(\frac{\text{sector standard}}{\text{realisation}} \right)^{-1} \times 100 = \frac{8.55}{9} \times 100 = 95$$

아래의 두 값 역시 같은 방식으로 구한다.

- ③ -> 세 개의 부차원에 대한 가중평균의 값.

$$\frac{95 \times 30 + 107 \times 50 + 112 \times 20}{100} = \frac{10440}{100} = 104.4 \approx 104$$

C. 시의성

<표 6-3> 시의성에 대한 지수

Dimension	weighting (%)	sector standard(day)	realization (day)	index
Timeliness	25			⑥108
production term	60			④ 108
provisional	50	12	①11	②109
definitive	50	17	16	106
planning vs. production term	40			⑤108
provisional	50	12	11	③109
definitive	50	17	16	106

- ① -> 기록에 의한 결과

- ②와③ -> 다음과 같은 공식을 사용하여 구한다.

$$\frac{\text{sector standard}}{\text{realisation}} \times 100 = \frac{12}{11} \times 100 = 109.09 \approx 109$$

- ④와 ⑤ 그리고 ⑥ -> 앞의 예에서와 같이 가중평균으로 구한다. 즉,

$$\frac{109 \times 50 + 106 \times 50}{100} = \frac{10750}{100} = 107.5 \approx 108$$

D. 정보수준

<표 6-4> 정보수준에 대한 지수

Dimension	weighting (%)	sector standard	realization	index
Level of Information	25			⑤87
Statline user satisfaction	8	⑥	⑦	
Satisfaction of 5 most important users	90	9	①8	③89
SN publishes one figure	2	yes=1	②no=0	④0

①과 ② -> 설문조사와 계획에 의한 값

③ -> 다음과 같은 공식을 사용하여 구한다.

$$\left(\frac{\text{sector standard}}{\text{realisation}} \right)^{-1} \times 100 = \frac{8}{9} \times 100 = 88.89 \approx 89$$

④ -> 값은 다음 식에 의해 구해진다.

$$100 + 100 \times (\text{realisation} - \text{sector standard}) = 100 + 100 \times (0 - 1) = 0$$

⑤ -> 가중평균의 값으로서 다음과 같이 구한다.

$$\frac{89 \times 90 + 0 \times 2}{92} = \frac{8010}{92} = 87.06 \approx 87$$

⑥과 ⑦의 값은 없음. 따라서 위의 가중평균에서 분모가 '100'이 아니라 '92'가 됨에 유의.

E. 최종 품질지수

<표 6-5> 최종 통계품질지수

Dimension	weighting (%)	sector standard	realization	index
Quality Index	100			①95
Accuracy	30			84
Coherence	20			105
Timeliness	25			108
Level of Information	25			87

① -> 네 개의 차원에 대한 가중평균의 값으로 다음과 같이 구한다.

$$\frac{84 \times 30 + 105 \times 20 + 108 \times 25 + 87 \times 25}{100} = \frac{9495}{100} \approx 95$$

(3) 결과의 해석

전체적인 품질은 품질지수가 '100'보다 작은 값 '95'로 나타났다. 결국 품질의 수준은 기대수준에 미치지 못하였음을 알 수 있고, 이 중에서 '정확성'과 '정보수준' 차원에서 기대 이하의 통계품질이 나타났다. 그 외의 차원에서는 비교적 기대한 수준의 품질을 달성한 것으로 품질지수 값이 나타났다.

6.2 적용 가능성 및 제안

6.2.1 국내 통계 적용 가능성

- (1) 앞에서 사용한 품질지수는 적절한 수정을 통해 국내 적용이 가능하다고 판단된다. 예컨대, '정확성'을 제외한 다른 세 개의 차원은 모두 설문평가 후의 결과이거나 '시의성'에 관련된 것이므로 그대로 적용 가능하다. 단지 '정확성'에서는 중요 결과물의 Level과 Trend로 평가를 하였는데, 이것을 설문의 형태로 바꾸어 '일관성'과 '정보수준'에서와 같이 평가를 할 수 있을 것이다. 이 때는 설문문항을 적절히 만드는 것이 중요하다.
- (2) 일반적으로 '정확성' 등에 사용되는 응답률, 대체율 그리고 포함률 등에 대한 정보도 사용 가능하다. 이 때는 전년도와의 비교 혹은 올해의 목표와의 차이 등으로 손실함수를 측정할 수 있다. 이 때는 다른 것들과의 weight 등의 문제를 적절히 해결하면 될 것이다.

6.2.2 새로운 품질지수의 제안

(1) 품질지수의 복잡성

앞 절을 통해 살펴본 네덜란드 품질지수는 계산이 복잡하다. 이는 품질지수의 기준값을 '100'으로 하기 위해 검토 항목의 자료 성격에 따라 지수 계산을 달리 하기 때문이다. 즉, 다음의 세 가지 계산법을 적용하였다.

$$\frac{\text{sector standard}}{\text{realisation}} \text{ 혹은 } \left(\frac{\text{sector standard}}{\text{realisation}} \right)^{-1}$$

혹은 $100 + 100 \times (\text{realisation} - \text{sector standard})$

(2) 손실함수(Loss Function)의 이용 가능성

① 기본 아이디어

- sector standard에 미치지 못하는 품질의 통계인 경우 그 '손실(loss)'을 직선 비례하는 것으로 판단하였다. 즉, 표준치(목표치)에 미치지 못하는 경우 표준치와 실제값과의 차이에 단순선형 비례하여 손실이 발생하는 것으로 생각하였다.
- 그러나 품질관리측면에서는 목표치에서 벗어날수록 일반적으로 손실은 quadratic하게 점점 더 커지는 것으로 생각한다. 따라서 일반적으로 다음과 같은 손실함수를 적용한다.
- 이와 같은 개념을 이용하여 품질을 측정할 수 있는 측도를 만들 수 있다.

$$L = k(y - t)^2$$

여기서 t 는 목표치이고, k 는 단위손실액을 나타낸다.

② 품질지수

- 반드시 지수 값의 기준을 '100'으로 만들 필요가 없다면 다음과 같은 지수 계산법도 가능하다.

$$QIQD = \text{sgn} \times (y - t)^2$$

여기서 y, t 는 각각 실측값(실제값, realization)과 목표값(target value 혹은 sector standard)을 의미하며, sgn 은 realization의 상태가 sector standard 보다 나빠지는 상태일 경우는 “-” 이고, 그 반대의 경우에는 “+”가 된다.

- 제안된 품질지수의 장점은 단 하나의 통일된 식으로 품질을 표현하는 것이 가능해지고, 품질의 정도를 제품의 형태로 표현함으로써, 목표치와 차이에 의해 품질이 목표치에서 벗어날수록 품질은 제품의 크기로 나빠진다는 사실을 나타낼 수 있다. 이러한 개념은 일반적으로 통계학을 이용한 품질측정 문제에서 손쉽게 받아들일 수 있는 개념이다.
- 이와 같은 개념을 사용할 경우, 특성치에는 다음의 세 가지 종류가 존재한다.

- ▷ 망목특성(望目特性, nominal-is-best) : 특정한 목표치가 주어진 경우
 - ▷ 망소특성(望小特性, smaller-is-best) : 작을수록 좋은 경우
 - ▷ 망대특성(望大特性, larger-is-best) : 클수록 좋은 경우
- 통계품질을 고려할 때, 기획자가 허용하는 최소(혹은 최대)의 목표치보다 조사가 마무리된 후의 실측치가 크면 클수록(혹은 작으면 작을수록) 좋은 경우로 나뉘어 진다. 즉, 소비자(기획자)가 허용하는 목표치에 대해 망소 혹은 망대특성을 가진 특성치로 분류된다. 따라서 이러한 특성에 따라 다음과 같이 두 종류의 품질지표를 사용할 수 있다.

▷ 망소특성(望小特性, smaller-is-best) : 작을수록 좋은 경우

$$QIQDS = \begin{cases} + (y - t)^2 & , \quad y \leq t \\ - (y - t)^2 & , \quad y > t \end{cases}$$

▷ 망대특성(望大特性, larger-is-best) : 클수록 좋은 경우

$$QIQDL = \begin{cases} + (y - t)^2 & , \quad y \geq t \\ - (y - t)^2 & , \quad y < t \end{cases}$$

여기서 $QIQDS$, $QIQDL$ 에서 sign은 측정된 통계품질의 상태를 의미하며, 구체적인 내용은 다음과 같다.

" + " <=> "품질 양호"

" - " <=> "품질 불량"

③ 예제

- 제한한 손실함수와 가중평균을 사용하면 앞의 품질지수는 다음과 같이 바뀐다.

<표 6-6> 손실함수를 이용한 품질지수

Dimension	Weighting(%)	SN-Index	QIQD
QI	100	95	-0.208
Accuracy	30	84	-0.722
Coherence	20	105	0.301
Timeliness	25	108	1.0
Level of Information	25	87	-1.205

- 위의 예제에서 $y \Leftrightarrow realization$ 이며, $t \Leftrightarrow sector\ standard$ 이다.
- 또한 '-'의 값은 실측된 통계품질이 sector standard보다 나빠진 것을 의미하며, '+'의 경우는 통계품질이 sector standard보다 좋다는 것을 의미한다. 또한 절대값의 크기는

통계품질의 좋고 나쁜 정도로 해석할 수 있다. 여기서 제안된 지수 QIQD와 SN-지수 사이에는 품질의 순위가 비교적 일치하지만, '정보수준'에서는 다른 결과가 나타났다. 즉, SN-지수에서는 가장 품질이 나쁜 차원은 '정확성'으로 나타났지만, QIQD 값으로는 '정보수준'이 가장 품질이 나쁜 것을 나타냈다. 그 이유 '정보수준'의 SN-지수를 구하는 과정에서 공표에 대한 index 값이 '0'으로 계산되어 있지만, QIQD에서는 그 값이 '0'이 아닌 값으로 포함되어 있고, 단 두 개뿐인 '정보수준'의 평가지표 중 하나의 지표('5개의 중요 이용자에 의한 만족도')가 가중치 '90%'로 과도하게 치우쳐 있기 때문으로 생각된다.

- 따라서 위의 계산결과로는 전체적인 품질은 기대에 미치지 못하는 수준이고, 그 중에서 '정보수준'과 '정확성'이 통계품질을 나쁘게 만들었고, '시의성'과 '일관성'이 비교적 통계품질이 우수했다는 것을 알 수 있다.

④ 특징

- 손실함수를 이용한 새로운 품질지수는 상대적으로 매우 간단하다. 특히 이 개념은 통계에서 품질을 다룰 때, 보편적으로 사용되는 개념이다. 품질공학에서와 신뢰성부분에서는 이러한 개념이 매우 보편화되어 있다.
- 목표치로부터의 손실이라는 개념은 공식통계의 품질에서 다루고 있는 '사용자에 대한 만족도'와 동일한 개념으로 사용될 수 있다. 따라서 절차별 품질의 평가와 차원별 평가 모두에서 적절히 사용될 수 있는 측도로 생각된다.
- 많은 부분에서 절차 혹은 차원별 주요 지표가 나와야 하며, 목표치가 결정되어야 한다는 점이 단점이 될 수 있다. 그러나 이 것은 각 통계를 담당하는 부서에서 매년 혹은 장기적인 품질개선의 목표가 설정되고 이것이 외부 혹은 내부평가를 거쳐 타당성을 인정받는 시스템을 유지한다면, 오히려 품질개선에 큰 도움이 될 수 있다.
- 굳이 또 하나의 단점을 지적한다면, 품질지표의 값의 차이가 '0'을 기준으로 매우 적을 수 있기 때문에 품질의 정도가 쉽게 인식되지 않을 수 있다.

⑤ QIQD를 '100'의 값을 기준으로 조정

QIQD를 다음과 같이 변환시킨다.

$$QIQD100 = 100 + sgn \times (y - t)^2$$

- 이 값은 실측값이 $y = t$ 가 된 경우에는 $QIQD100 = 100$ 가 되고, 품질이 나빠진 경우에는 $QIQD100 < 100$ 이 되며,

품질이 좋아진 경우에는 $QIQD100 > 100$ 이 된다.

따라서 $QIQD100$ 은 “100”을 기준값으로 작은 경우에는 품질이 나빠진 상태를 그리고 커지는 경우에는 품질이 좋아진 상태를 나타낸다.

○ 예를 들어, 위의 예제에서 $QIQD100$ 의 값은 다음과 같다.

<표 6-7> ‘100’을 기준으로 변환된 통계품질지수

Dimension	Weighting(%)	SN-Index	QIQD	QIQD100
QI	100	95	-0.208	99.792
Accuracy	30	84	-0.722	99.278
Coherence	20	105	0.301	100.301
Timeliness	25	108	1.0	101.0
Level of Information	25	87	-1.205	98.795

⑥ 가중치와 sector standard의 결정

○ 가중치

이론적으로 가중치는 주관적인 값이다. 또한 통계의 특성에 따라 차원이나 부차원에 대한 가중치는 달라질 수 있다. 따라서 생산되는 통계를 담당하는 sector(부서)와 품질관리팀 및 자문위원단(가칭, 이것이 있다면)이 협의하여 매해 가중치를 다르게 작성할 수 있다.

○ sector standard

sector standard 역시 각 부서의 능력과 생산하는 통계의 특성에 따라 설정된다. 물론 sector standard를 정할 때, 품질팀과 자문위원단의 허락을 받아야 한다. 이 때는 해당통계의 품질향상 계획서(가칭)가 반드시 제출되고 평가를 받은 후에 결정하는 것이 바람직하다.

⑦ 결론

제안된 품질지수를 적용하기 위해서는 다음과 같은 문제를 고려하여야 한다.

○ 네덜란드 통계국과의 품질 차원과 부차원에 대한 차이

우리나라의 현실에 맞는 차원과 부차원을 정하여야 한다. 이 때 자료의 유무가 차원의 결정에 크게 고려된다.

○ 부차원에 따른 realization의 값, 예를 들어 $RMRQD$ 혹은 $RMQD$, 기타 설문조사를

통한 점수(10점 척도) 등을 구할 수 있는 시스템이 만들어져 있어야 한다.

통계 생산을 담당하고 있는 부서에서 유사통계를 만들어 내는 기관들의 자료가 수시로 수집되고 비교되어야 하고, 또한 이용자 등의 만족도 조사 등을 적절하게 실시하고 그 값을 이용할 수 있는 시스템이 갖추어져야 한다.

- 가중치와 기대치에 대한 결정이 이루어져야 한다.

모든 통계에 적용되는 차원 등에 대한 가중치가 일률적으로 적용되는 것은 바람직하지 않다. 이는 전체 통계의 비교에는 용이하지만, 각 통계별 특성이 전혀 반영되지 못한다는 단점이 있다. 결국 담당부서의 품질향상 의지를 일으키지 못하는 부작용을 낳을 우려가 높다. 따라서 가중치와 품질향상의 기대치(목표치)는 담당부서에서 충분한 협의를 거쳐 결정하고, 이를 추진 받도록 하는 것이 좋다.

(3) 시그마 수준(sigma-level)의 적용 가능성

① 기본 IDEA

- 이 방법은 절차별 품질(process quality)를 관리/측정하는 경우에 많은 산업부분에서 품질개선전략으로 사용되고 있는 “6-시그마 경영전략”에서 이용되는 방법이다. 즉, 현재의 품질수준을 측정하거나 개선된 품질 수준을 측정할 때, 사용되는 방법으로 모든 과정에서 동일한 단위인 “ σ ”로 품질이 표현된다는 장점이 있다.
- 물론 현재의 평가 시스템에서 이 방법을 적용하는 데는 약간의 무리가 따른다. 즉, 각 통계별로 평가 자료가 자체 및 외부 각각 1명이 평가하기 때문에 통계별 품질수준을 측정하기에는 자료가 너무 빈약하다.
- 이 방법을 적용하기 위해서는 각 통계별로 자체평가, 외부평가 혹은 이용자 평가 등 보다 많은 자료가 있을 때, 신뢰성 있는 측도가 될 수 있을 것으로 생각한다.

② 적용사례

- 구체적인 측정방법은 “6-시그마” 관련 문헌을 참조하고, 2002년도 자체 및 외부평가 자료를 각 통계별 분석이 아닌 모든 자료를 통합하여 시그마 수준(품질수준)을 측정하여 보았다.
- 현재의 각 절차별 시그마 수준
- ▷ 분석을 위해 다음과 같은 가정을 설정하였다.

<가정-1>각 절차에서 “관계없음”을 제외한 총 응답수에서 “보통” 이하의 도수를 불량(단위당 결함)으로 판정.

<가정-2>각 절차에서 “관계없음”을 제외한 총 응답수에서 “미흡” 이하의 도수를 불량(단위당 결함)으로 판정한다.

▷ 분석결과

<표 6-8> 절차별 시그마 수준

평가 절차 \ 가정	자체평가		외부평가	
	가정-1	가정-2	가정-1	가정-2
1.조사기획	2.747	3.379	2.128	2.658
2.모집단 및 표본	2.836	2.836	2.311	2.678
3.조사표 설계	6.0	6.0	2.136	2.664
4.조사직원 관리	2.344	2.344	2.569	3.278
5.조사실시	3.191	3.191	2.546	3.558
6.자료처리 및 집계	2.522	2.905	2.336	2.753
7.자료공표	3.693	3.693	2.391	3.042
8.자료이용	2.867	3.278	2.202	2.782
최종 시그마 수준	1.434	1.723	0.458	1.457

<참고>계산방법 : 자체평가 <가정-1>과 <가정-2>의 경우

① 총 응답수 --> 98

② 단위당 결함수(DPU)

<가정-1>의 경우 11(보통 8, 미흡 3) <가정-2>의 경우 3(미흡 3)

③ 불량률의 비율

<가정-1> $p = 11/98 \approx 0.1122$ <가정-2> $p = 3/98 \approx 0.0306$

④ 해당 수율(yield)

<가정-1> $e^{-p} \approx 0.893866$ <가정-2> $e^{-p} \approx 0.969863$

⑤ 장기적인 시그마수준

<가정-1> $\Phi^{-1}(e^{-p}) + 1.5 \approx 1.124735 + 1.5 = 2.74735$

<가정-2> $\Phi^{-1}(e^{-p}) + 1.5 \approx 1.878789 + 1.5 = 3.378789$

⑥ 최종 시그마 수준 : 각 절차의 수율을 곱하여 얻어진 누적수율에 의해 계산함.

즉, 자체평가 <가정-2>의 경우 다음과 같이 계산된다.

$$\text{누적수율} : \prod_{i=1}^6 e^{-p_i} \approx 0.588231$$

$$\rightarrow \Phi^{-1}(0.588231) + 1.5 \approx 1.723$$

○ 현재의 각 차원 및 절차별 시그마 수준

▷ 분석을 위해 다음과 같은 가정을 설정하였다.

<가정-1> 각 절차에서 “관계없음”을 제외한 총 응답수에서 “보통” 이하의 도수를 불량(단위당 결함)으로 판정.

<가정-2> 각 절차에서 “관계없음”을 제외한 총 응답수에서 “미흡” 이하의 도수를 불량(단위당 결함)으로 판정한다.

▷ 분석결과 (계산 방법은 동일)

<표 6-9> 각 차원 및 절차별 시그마 수준

절차	차원	정확성		접근성		시의성		관련성		비교성		효율성	
		자체	외부	자체	외부	자체	외부	자체	외부	자체	외부	자체	외부
조사 기획	가정1	3.050	2.313	2.345	1.845	6.0	2.268	2.628	2.118	-	-	2.522	1.770
	가정2	3.369	3.018	6.0	2.469	6.0	2.522	3.421	2.559	-	-	2.906	2.522
모집단 및표본	가정1	6.0	2.431	2.345	1.962	-	-	-	-	-	-	2.345	2.073
	가정2	6.0	6.0	2.345	2.073	-	-	-	-	-	-	2.345	6.0
조사표 설계	가정1	2.14	2.136	-	-	-	-	-	-	-	-	-	-
	가정2	2.66	2.664	-	-	-	-	-	-	-	-	-	-
조사직 원관리	가정1	2.57	2.569	-	-	-	-	-	-	-	-	-	-
	가정2	3.28	3.278	-	-	-	-	-	-	-	-	-	-
조사실 시	가정1	2.55	2.546	-	-	-	-	-	-	-	-	-	-
	가정2	3.56	3.558	-	-	-	-	-	-	-	-	-	-
자료처 리및집 계	가정1	2.41	2.410	-	-	-	-	-	-	2.860	2.146	-	-
	가정2	2.71	2.809	-	-	-	-	-	-	2.860	2.410	-	-
자료공 표	가정1	2.86	2.753	6.0	2.358	6.0	2.334	-	-	-	-	-	-
	가정2	6.0	6.0	6.0	3.221	6.0	2.827	-	-	2.469	1.737	-	-
자료이 용	가정1	6.0	2.410	3.029	2.536	-	-	-	-	2.469	2.496	-	-
	가정2	6.0	6.0	3.029	2.848	-	-	-	-	-	-	-	-

③ 결론

- 이 방법은 절차별(과정별)로 품질을 관리하고자 할 때 사용될 수 있다. 또한 이 전략은 TQM 이후 이미 제조업을 비롯한 많은 산업부분에서 전 세계적으로 이미 인정을 받고 있는 방법이기 때문에 적용에 있어서 대내외적인 타당성을 인정받을 수 있을뿐더러, 이미 많은 사례로 연구발표 되어 있기에 적용이 매우 용이할 것으로 생각된다.
- 그러나 이 방법을 적용하는 데는 몇 가지 어려움도 있다. 특히 각 절차별로 객관적인 품질측도(지표)가 정해져야 하고 그 목표치가 설정되어야 한다. 이러한 측도가 존재하지 않은 경우에는 설문조사 등을 통해 만족도에 대한 폭넓은 조사가 이루어지고 그 결과를 바탕으로 시그마 수준이 측정되어야 한다.

제 7장 결론 및 제언

아직까지 통계품질의 중요성에 대한 인식이 부족하고 이를 평가하는 지표도 미흡하였지만 통계청은 지난 2002년 처음으로 48개의 통계품질평가지표를 사용해서 통계청에서 생산한 12개의 통계에 대한 품질평가를 실시하였다. 그러나 주요 국제기구나 선진국들에 비해 통계품질 향상을 위한 조직과 인력이 부족하고 통계품질지표도 충분치 않는 등 여건이 제대로 마련되어 있지 않아 정확한 평가와 올바른 통계품질 향상정책을 수립하는데 어려움이 있었다. 이에 본 연구에서는 이러한 문제들을 조금이나마 해결하여 통계품질을 제대로 평가할 수 있는 기틀을 마련하고자 하였다. 본 연구에서 중점을 두고 다룬 부분들은 다음과 같다.

첫째는, 현행 통계청에서 통계품질 향상을 도모하기 위해서 사용하고 있는 통계품질평가 시스템을 검토하였다.

둘째는, OECD, IMF 등의 국제기구와 일찍부터 통계품질 향상에 노력을 기울여온 캐나다, 미국, 영국, 스웨덴 등 주요 국가들에서 사용하고 있는 통계품질평가 방법과 지표들을 비교 분석하였다.

셋째는, 현행의 통계품질평가지표들을 수정, 보완하여 우리나라에 필요한 81개의 통계품질평가 지표를 8개 절차와 7개 차원으로 나누어 개발하고, 이들 중에서 48개의 필수지표를 분류하였다.

넷째, 통계품질의 수준을 측정하기 위한 2가지의 품질지수를 제안하였다.

또한 우리나라의 통계품질향상에 효과가 있을 것으로 보이는, 연구수행과정에서 얻은 몇 가지 추가 사항들을 부분별로 나누어서 살펴보았다.

<품질평가 시스템>

(1) 통계품질평가에 대한 기본 철학의 변화

- 결과물에 대한 절대적인 평가점수에서 탈피하고 상대적인 성취도로 평가한다. 즉, 각 통계의 품질 향상의 정도(성취도)를 파악하여 포상하고 그들의 know-how를 공유한다.
- 최종적으로 실무자 중심의 자체평가가 이루어지는 문화를 형성한다. 즉, 해당 통계의 실무담당부서에서 생산하는 통계의 품질을 향상시킬 수 있는 목표를 스스로 정하고, 개선 방법을 연구하고 적용하여 최종결과의 성취도를 측정하고 이를 평가받도록 한다.

(2) 평가제도의 변화

- 통계 담당 부서에서는 (가칭)자체평가위원회(이하 '자평위'라 함. 혹은 위원회가 아니어도 내부적으로 운영되고 자문을 구할 수 있는 그룹일수도 있다)를 두어 다음과 같은 역

할을 하도록 한다.

[구성]

‘자평위’는 통계 생산 실무담당자와 주요 이용자 그리고 자문위원으로 구성하며, 품질관리팀의 1명이 참여할 수도 있다.

[임무]

- ① 각 통계 생산 전에 이용자들이 요구하는 내용을 파악하고 이를 반영하기 위한 계획을 세운다.
- ② 올해의 품질향상목표와 계획을 수립하고 이를 통계 생산계획에 반영토록 한다.
- ③ 통계 생산과정에 대한 품질평가지표에서 지적한 사항들을 수시로 점검하며 실행토록 한다. 특히 조사표와 표본설계 부분에서 변화되는 흐름을 충분히 반영하고 있는 지를 면밀히 검토한다.
- ④ 통계생산 후에 생산된 통계가 최초 목표에 맞게 생산되었는지를 적절한 방식으로 의견을 수합하고 만족도를 측정한다.
- ⑤ ‘자평위’는 품질향상계획서(목표치와 가중치 및 예산 등이 포함됨)를 품질관리팀에 제출하여 승인을 받고, 최종 자체평가서를 품질관리팀에 제출하여 평가를 받아야 한다.

- 품질관리팀에서는 (가칭)품질평가위원회(이하 ‘품평위’라 함)를 두어 ‘자평위’의 계획서와 결과보고서 그리고 그들의 활동 사항을 검토하고 좋은 안을 제안하며 평가하는 위원회이다. 구성과 임무는 다음과 같다.

[구성]

‘품평위’는 품질관리팀 실무자, 내부에서 추천된 평가자 그리고 외부 자문위원으로 구성한다.

[임무]

- ① ‘품평위’는 당해 년도의 기준 목표치 상향 조정정도(비율, %), 허용오차 비율 그리고 차원별(또는 절차별) 기준 가중치, 허용오차 범위 등을 결정하여 해당 실무부서에 알려준다.
- ② 각 통계 담당 부서에서 운영하는 ‘자평위’가 제출한 품질향상계획서(목표치 포함)를 검토하고 추진한다. 이때 계획서의 내용(목표치와 가중치 등)이 적절치 못하다고 판단될 때는 수정하도록 요구할 수 있다.
- ③ 품질 개선을 위한 방법과 예산 등에 대해서도 적절성과 효율성을 평가한다
- ④ 통계 생산 중에 발생하는 다양한 문제들을 수시로 청취하며, 생산부서의 요청이 있을 때는 이를 해결하기 위한 방안들을 제시하여야 한다.
- ⑤ ‘자평위’의 최종 자체평가서의 적절성을 검토하여 성취도를 추진한다.

⑥ 평가 대상이 되는 모든 통계들에 대해 당해 년도의 성취도를 비교 검토하기 위한 (가칭)통계품질개선보고대회(이하 '대회'라 칭함)를 개최하고 우수한 성과를 올린 부서를 선정하고 모든 사례를 공유토록 한다.

(3) 포상제도 :

- 모든 성과는 객관적이고 공정하며, 개방된 평가를 거쳐 포상을 하도록 한다. 이를 위해서는 '대회'를 매년 개최한다.
- '대회'는 '품평위'가 주관하고, 심사위원은 외부와 내부에서 선정한다. 심사의 기준은 '성취도'와 '개선 방법의 참신성' 등에 가중치를 많이 부여한다.
- '대회'의 성격은 일종의 경진대회이며, 1년 동안 수고한 통계청 직원들의 잔칫날의 형식을 담아야 한다(통계의 날을 이용하는 것도 바람직하다).
- 훌륭한 성과를 올린 품질개선 사례를 많은 사람이 보고 이를 공유할 수 있는 자리가 되어야 한다.

<품질지수 부분>

(1) 목표치(target values 혹은 sector standard)

- '자평위'는 당해 년도에 생산되는 통계의 품질에 가장 치명적인 문제를 만들어 내는 요인들(품질을 저하시키는 가장 중요한 요인들)을 찾아 이를 개선시키기 위한 계획서를 '품평위'에 제출한다. 이 때 개선 목표치를 설정한다.
- 목표치는 달성이 예상되는 값보다 약 20-30% 상향 조정하여 잡도록 하며, 목표치 설정의 타당성을 구체적으로 설명한다.
- 최종 자체평가서에서는 이 목표치를 기준으로 성취도를 측정한다. 이 때 예산 대비 효율성을 계산하고, 이러한 개선방식의 추후 유지/관리 방안도 마련한다.
- 계획서에는 품질향상 전략을 제대로 수행하여 목표치를 달성하기 위한 예산 등의 협조 사항에 대해 구체적으로 명시한다.

(2) 가중치(weight)

- 일반적으로 차원 혹은 절차에 대한 가중치는 통계별로 큰 차이가 나지 않을 것이다(이를 '품평위'에서 결정해 준다). 예를 들어, '정확도'가 가장 큰 가중치를 부여받을 것으로 판단된다. 그러나 각 통계의 특성이 다르고 담당부서의 여건과 해결방안의 난이도 등이 모두 다르기 때문에 매년 동일한 가중치로 평가를 하는 것은 바람직하지 않다.
- 따라서 가중치는 당해 년도 실무 담당 부서의 환경 등을 고려하여 '품평위'가 제안해 준 가중치의 기준과 다르게 부여될 수 있다. 물론 이를 '품평위'에서 승인 받아야 한다.
- 가중치 부여 방식은 '100'을 기준으로 중요도와 난이도 등을 고려하여 결정한다. 이때 가중치는 차원별로 그리고 해당 차원의 부차원에서 각각 부여하며, 심각하게 복잡하고

논란의 여지가 있다고 판단될 때에는 부차원 이하의 가중치는 동일하게 잡을 수 있다.

통계품질향상을 위한 하나의 시도로 이번 연구가 이루어졌고, 이러한 연구를 수행하면서 아직도 많은 부분의 연구가 부족함을 알 수 있었다. 통계품질을 위한 가이드라인을 우선적으로 마련하고, 평가시스템을 좀더 보완하며, 통계품질 향상을 위한 교육이 절실함을 실감하였다. 또한 본 연구에서 만든 품질평가지수에 대한 지속적인 수정과 개선 작업이 요구되며, 통계품질지수개발에도 노력을 기울여야 할 것이다.

통계품질의 향상은 어떤 특정 부서만의 노력으로 이루어질 수 없고, 통계생산에 직·간접적으로 참여하는 부서와 통계이용자들, 통계품질관리팀, 그리고 학계의 전문가들이 장기간에 걸친 협력이 있어야 가능하다.

참고문헌

- 영국통계청(2000), *National Statistics Quality Reviews*.
- 이동명, 김설희(2002). “국가통계 품질측정을 위한 체계적 접근”, *한국조사연구학회 2002년도 추계학술대회 발표논문집*, 111-127.
- 통계청(1999), *캐나다 통계품질 가이드라인*.
- 통계청(2000), Collection of papers on statistics quality work.
- 통계청(2002), *2002년도 해외단기 실무연수 자료 통계품질평가방법 및 제도 연구(2002.5)*
- Australian Bureau of Statistics (1996). Information Paper: Quality of Australian Balance of Payments statistics. Chapter 1. Data Quality Concepts. Statistics Concepts Library.
- Australian Bureau of Statistics (1998). Balance of Payments and International Investment Position, Australia, Concepts, Sources and Methods. Chapter 15. Data Quality. Statistics Concepts Library.
- Booleman, M and Brakel, R.(1999), Attention to Quality within Statistics Netherlands - Quantifiable, Statistics Netherlands.
- Brackstone, G.(1999), Managing Data Quality in a Statistical Agency, *Survey Methodology*, Vol. 25, No. 2, 139-149.
- Carson, C. S.(2000), What is Data Quality? A Distillation of Experience.
- Collins, M and Sykes, W.(1999), Extending the Definition of Survey Quality, *Journal of Official Statistics*, Vol. 15, No. 1, 57-66.
- Elvers E. and Rosen B.(1999). Quality Concept for Official Statistics. *Encyclopedia of Statistical Science*, 621-629. New York: John Wiley.
- Eurostat (2000). Assessment of the Quality in Statistics.
- Lessler, J. T. and Kalsbeek, W. D.(1992), *Nonsampling error in surveys*, John Wiley and Sons.
- Lyberg, L. et al.(1997), *Survey Measurement and Process Quality*, John Wiley and Sons.
- Policy on Informing Users of Data Quality and Methodology-캐나다(doc파일)
- Statistics Canada(2002), Statistics Canada's Quality Assurance Framework-2002
- U.S. Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget(2001), *Measuring and Reporting Sources of Error in Surveys*, *Statistical Policy Working Paper* 31, July 2001.

Werner G. and Hakan Linden (2001). Quality Measurement–Eurostat Experiences, *Proceedings of Statistics Canada Symposium*.

Willem de Vries and Richard van Brakel (1998). Quality Systems and Statistical Auditing. A Pragmatic Approach to Statistical Quality Management. *Netherlands Official Statistics*.

<관련사이트>

- [<http://dsbb.imf.org/vgn/images/pdfs/swerosc.pdf>].

International Monetary Fund (2001). Sweden: Report on Observance of Standards and Codes–Data Module; Response by the Authorities; and Detailed Assessments.

- [<http://dsbb.imf.org/Applications/web/dqrs/dqrsintroduction/#reference>].

International Monetary Fund (2003). Introduction to the Data Quality Reference Site.

- [http://dsbb.imf.org/vgn/images/pdfs/dqrs_generic.pdf].

International Monetary Fund (2003). Data Quality Assessment Framework–Generic.

- [<http://www.oecd.org/doc/M00029000/M00029990.doc>].

Organization for Economic Cooperation and Development (2002). Quality Framework for OECD Statistics.

- [http://www.statistics.gov.uk/about/consultations/ons_consultations/default.asp].

National Statistics (2003). Draft Quality Measurement and Reporting Framework Consultation Document.

- [<http://dsbb.imf.org/vgn/images/pdfs/swerosc.pdf>].

International Monetary Fund (2001). Sweden: Report on Observance of Standards and Codes–Data Module; Response by the Authorities; and Detailed Assessments.

- [<http://www.census.gov/qdocs/www/>].

U.S. Census Bureau (2003). Census Bureau Section 515 Information Quality Guidelines.

- [<http://unstats.un.org/unsd/goodprac/bpabout.asp>]

The UN Statistical Commission adopted the Fundamental Principles of Official Statistics in a Special Session in April 1994.

<부록 1> 절차별 세부 분석결과

(1) 조사기획

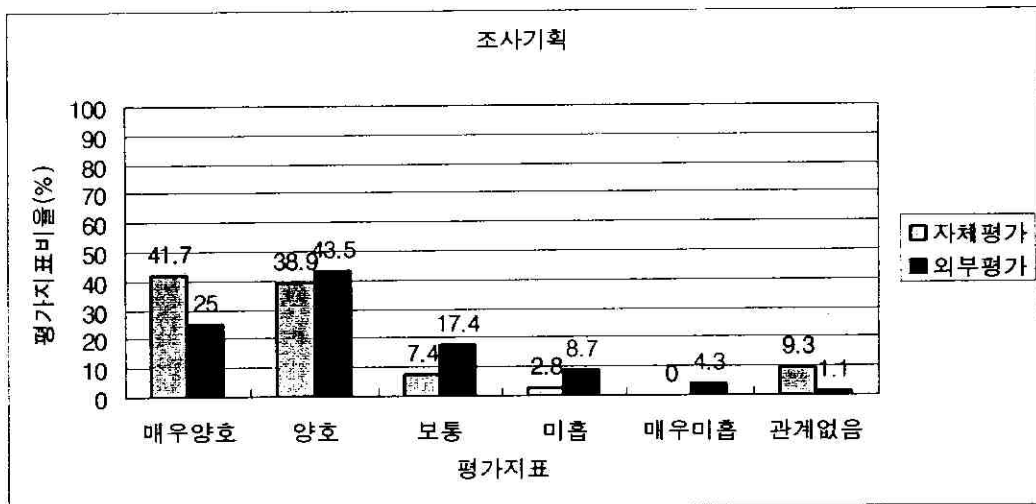
○ 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.3	0.41	0.3776	1=8-> 2-> 7-> 4-> 3-> 5=9-> 6	4.8/3.7
외부평가	3.8	0.27		1-> 2= 3= 6= 7= 8-> 5= 9-> 4	4.0/3.2

○ t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.00104	0.00119	0.00143
양측검정	0.00208	0.00238	0.00287

○ 응답현황



- <4.적정예산 확보>항목(자체평가 점수가 높음)과 <6.기획자의 직무교육>항목(외부평가 점수가 높음)은 극단적인 차이를 나타냄
- 3점 척도로 고려할 때, 외부평가에서 '양호'의 비율이 50%대로 비교적 낮은 항목이 세 개나 있음. 즉, <4.적정 예산 확보>, <5.이용자 요구사항 파악 및 반영> 그리고 <9.개

편작업의 공개성> 임. <9.개편작업의 공개성>은 자체평가에서도 58%만이 '양호' 판정.

- '관계없음' 항목 지정에도 차이가 존재. 자체평가에서 비율이 높음.
- 전반적으로 '미흡+ 매우 미흡'은 외부평가에서 높고, '매우 양호+ 양호'는 자체평가에서 높게 나타나고 있음.
- 특히 외부평가에서 10% 이상의 '미흡' 판정을 받은 항목도 <4.적정 예산 확보>, <5.이용자 요구사항 파악 및 반영>, <6.기획자의 직무교육>, <7.현장조사 업무량 파악>. <8.개편작업의 시의성> 그리고 <9.개편작업의 공개성> 등 여섯 개가 있음.

	'미흡+ 매우 미흡' 순위(내림차순)	'매우 양호+ 양호' 순위(내림차순)
자체평가	3=4=6->1=2=5=7=8=9	1=2->5->4=7->3=6->8->9
외부평가	5->9->4->8->6->7->2->1->3	1->3->8->2->7->6->5->9->4

- 자체평가 결과 <3.국제적 정의 준수>, <5.이용자 요구사항 파악 및 반영> 그리고 <6.기획자의 직무교육>에서 '양호'의 비율이 '매우 양호'의 비율보다 높게 나타남.
- 외부평가에서는 <6.기획자 직무교육>과 <8.개편작업의 시의성>에서만 '매우 양호'의 비율이 '양호'의 비율보다 높음

(2) 모집단 및 표본

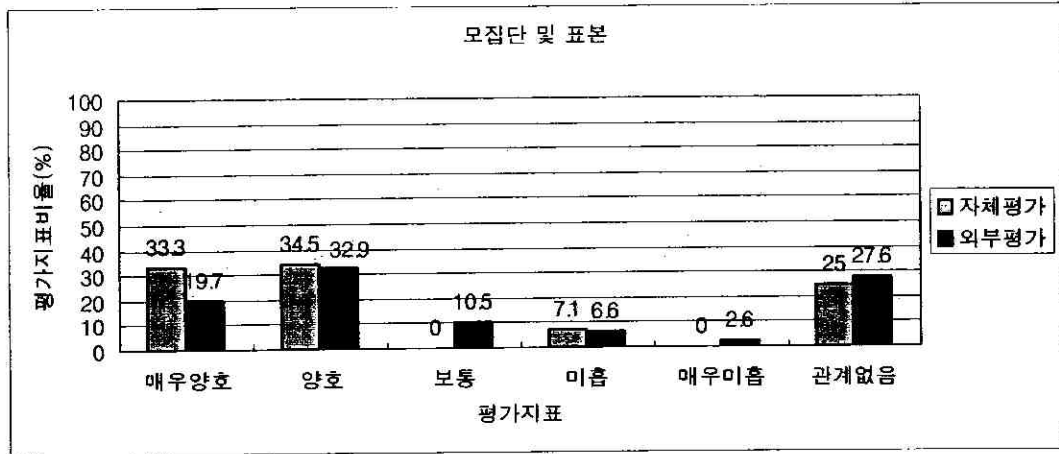
- 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.3	0.5711	0.5980	1=7->5->2->3->4->6	4.8/3.3
외부평가	3.8	0.4650		1->2->5->3->4->7->6	4.6/3.2

- t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.03005	0.08012	0.08175
양측검정	0.06010	0.16247	0.16351

○ 응답현황



- <7.표본추출 정보공개>와 <5.조사대상처 명부관리>항목은 자체평가와 외부평가의 평균 점수 차이에서 각각 1.4와 0.6으로 크게 나타나고 있으며, 자체평가 점수가 높았다.
- 3점 척도로 고려하였을 때, 외부평가에서 50% 이하의 '양호' 판정을 받은 항목이 네 개나 됨을 주목할 필요 있음. 즉, <3.표본의 일관성 유지>, <4.적정한 표본 규모>, <6.표본오차 공개> 그리고 <7.표본추출정보 공개>에서 낮게 판정 받음. 특히 <4.적정한 표본 규모>와 <6.표본오차 공개>는 자체평가에서도 58%와 42%정도로 낮은 '양호' 평가를 받은 사실에 주목.
- 그 외의 세 개의 항목 역시 외부평가에서 그리 좋은 점수를 받지 못함. '양호'의 비율이 각각 67%, 55%, 60% 임.
- <7. 표본추출 정보공개>와 <5.조사대상처 명부관리>항목은 3점 척도로서 '양호'의 비율 역시 자체평가에서는 75%이고 외부평가에서는 49%와 60%로 큰 차이를 보이고 있다.
- <4.적정한 표본규모>와 <7. 표본추출 정보공개>항목은 3점 척도로 고려했을 때, '미흡'의 비율에서 자체평가는 각각 17%와 0%인 반면 외부평가에서는 11%와 16% 정도를 나타내고 있다. <6.표본오차 공개>는 자체평가와 외부평가 모두에서 33%정도의 '미흡' 판정을 받았음.
- 자체평가 결과 <2.적정한 표본추출>, <3.표본의 일관성 유지>, <4.적정한 표본규모> 그리고 <5.표본오차의 공개> 항목에서 '양호'의 비율이 '매우 양호'의 비율보다 높음
- 외부평가에서는 <1.모집단의 명확한 정의> 항목에서만 '매우 양호'의 비율이 '양호'의 비율보다 높게 나타나고 있음.

(3) 조사표설계

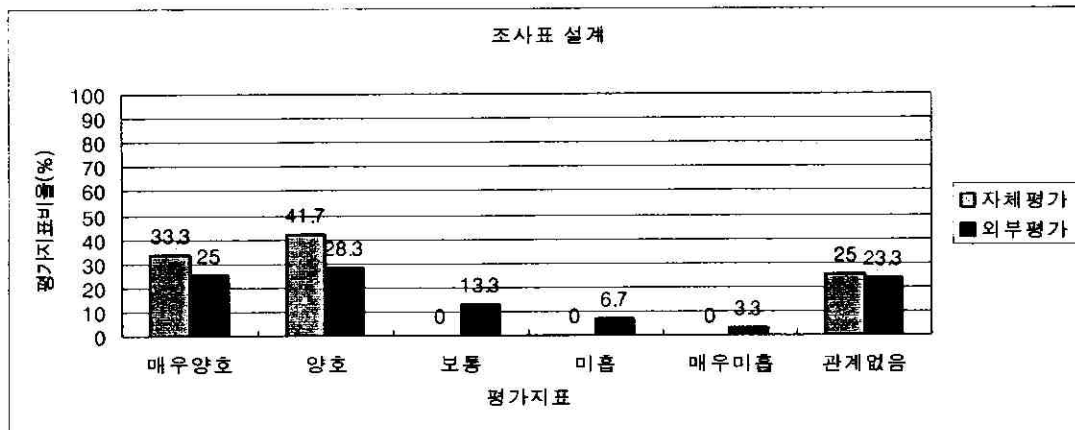
○ 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.4	0.2073	0.6780	5-> 4-> 3-> 1-> 2	4.7/4.2
외부평가	3.8	0.5263		5-> 1-> 4-> 3-> 2	4.7/3.4

○ t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.01438	0.01995	0.02794
양측검정	0.02876	0.03987	0.05587

○ 응답현황



- 평균의 차이는 <2.충분한 용어 설명>, <3.사전조사 실시> 그리고 <4.조사표 및 입력 프로그램의 친화성>에서 크게 나타남. 그 차이는 각각 0.8, 0.9 1.0 임.
- '관계없음' 항목에 비율 차이가 나타남
- 자체평가 결과 모든 항목에서 3점 척도로서의 '양호'의 비율은 75%이고 '관계없음'의 비율은 25%로 같게 나타나고 있음. 외부평가결과는 50% 이하의 '양호' 평가를 받은 항목은 <1.조사표 표현 및 흐름의 적절성>, <2.충분한 용어 설명>, <3.사전조사 실시> 그리고 <4.조사표 및 입력 프로그램의 친화성>에서 나타나고 있음.
- 심지어 자체평가에서는 모든 항목에서 0%인 '미흡'도 8.3%에서 17%까지 <5.지침서 총

실성>을 제외한 네 개의 항목에서 모두 나타나고 있음. 특히 10% 이상의 '미흡' 판정을 받은 항목은 <2.충분한 용어 설명>과 <3.사전조사 실시>로 나타났다.

- 특히 외부평가에서 '양호'의 비율은 <5.지침서의 충실성>을 제외한 네 개의 항목에서 50%에 이하의 낮은 점수를 받았고, 결국 '조사표 설계' 절차가 가장 낮은 평점을 받았음에 주목.
- 특히 <4.조사표 및 입력 프로그램의 친화성>과 <3.사전조사 실시>에서 큰 차이를 보이고 있음
- 자체평가에서 '양호'의 비율이 '매우 양호'의 비율보다 높게 평가된 항목은 <1.조사표 표현 및 흐름의 적절성>, <2.충분한 용어 설명> 그리고 <3.사전조사 실시>임.
- <5.지침서 충실성> 만이 외부평가에서 '매우 양호'의 비율이 '양호'의 비율보다 높음.

(4) 조사직원 관리

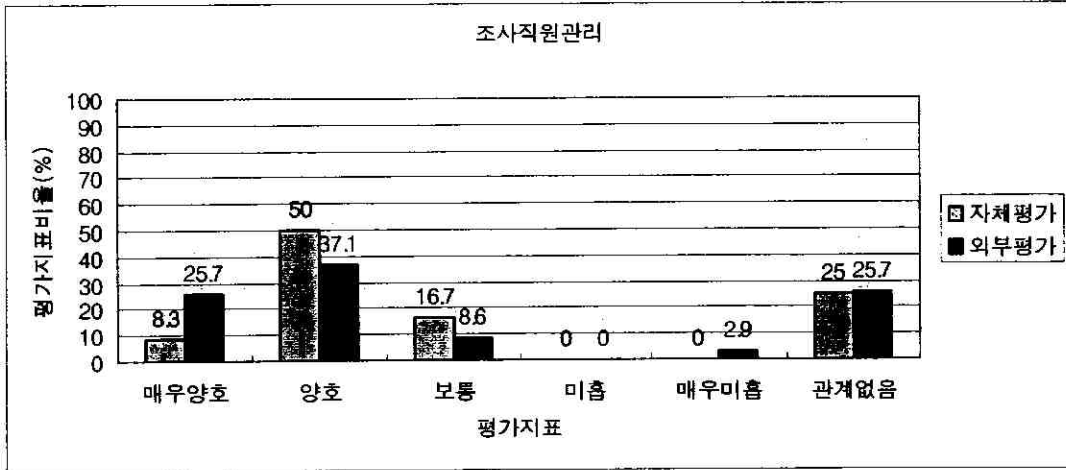
- 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	3.9	0.3	0.9608	2 ->1 ->3	4.2/3.6
외부평가	4.1	0.2082		2-> 1-> 3	4.3/3.9

- t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.03641	0.16523	0.16872
양측검정	0.07283	0.33046	0.33744

○ 응답현황



- 유일하게 외부평가가 자체평가보다 모든 항목에서 높은 점수를 받음
- '양호'의 비율은 자체평가에서는 <3.조사직원의 업무전문지식정도 파악>이 가장 낮은 반면, 외부평가에서는 가장 높게 평가되었다.
- 자체평가 결과 모든 항목에서 '양호'의 비율이 '매우 양호'의 비율보다 높음. 특히 <3.조사직원의 업무 전문지식 정도 파악> 항목은 '양호'와 '매우 양호'를 합한 비율이 41.7%로 아주 낮았음. 그럼에도 불구하고 '미흡'이나 '매우 미흡' 판정은 전혀 받지 않음.
- 외부평가가 전체적으로 자체평가보다 높은 평가를 받은 절차이다. 그러나 <3.조사직원의 업무 전문지식 정도 파악> 항목에서 '매우 미흡'의 비율이 약 8%로 나타나고 있음.
- 3점 척도로 고려할 때, 자체평가와 외부평가에서의 '양호'의 비율이 전혀 반대의 경향을 보임. 특히 <조사직원의 업무 전문지식 정도 파악> 항목은 자체평가에서는 가장 적은 비율을 나타내고 있지만, 외부평가에서는 가장 높은 비율을 나타냄.

(5) 조사실시

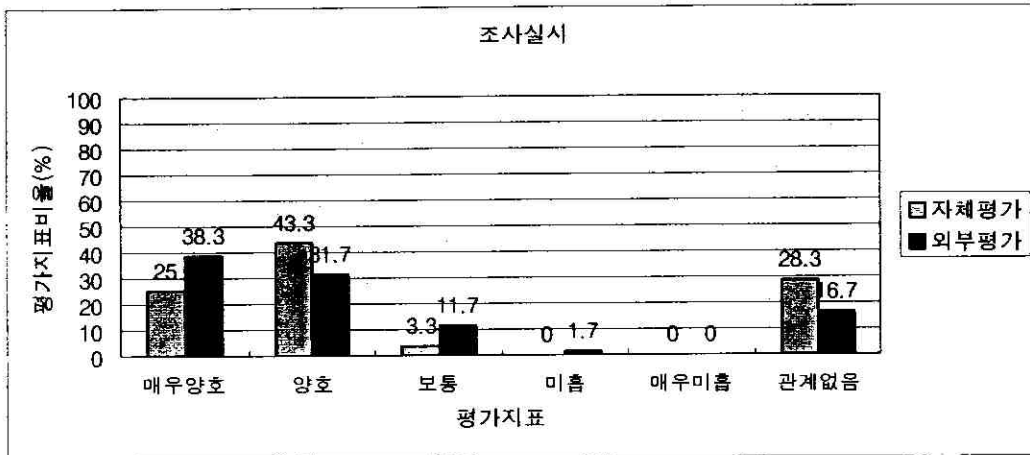
○ 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.3	0.15	0.0759	2-> 1= 3= 4-> 5	4.4/4.0
외부평가	4.3	0.13		3= 4-> 1-> 5-> 2	4.4/4.1

○ t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.4138	0.4143	0.4144
양측검정	0.8276	0.8287	0.8288

○ 응답현황



- <2.실제조사방법 파악>항목은 두 평가에서 극단적인 평가를 받음. 즉, 자체평가에서는 최고점을 외부평가에서는 최저점을 받음
- <3.조사시 문제점 해결 및 해결방법 공유>, <4.현장조사에 대한 점검 실시> 그리고 <5.현장조사의 무응답 처리지침>은 자체평가보다 외부평가에서 높은 점수를 받음
- '관계없음' 항목의 비율이 자체평가에서 많이 나타남.
- 자체평가 결과 <1.체계적인 현장관리>, <3.조사시 문제점 해결 및 해결방법 공유>, <4.현장조사에 대한 점검 실시> 그리고 <5.현장조사의 무응답 처리지침> 항목에서 '양호'의 비율이 '매우 양호'의 비율보다 높게 평가됨.
- 그러나 외부평가에서는 <3.조사시 문제점 해결 및 해결방법 공유>, <4.현장조사에 대한 점검 실시> 그리고 <5.현장조사의 무응답 처리지침> 항목에서 오히려 '매우 양호'의 비율이 '양호'의 비율보다 높게 평가됨.
- 외부평가에서 <3.조사시 문제점 해결 및 해결방법 공유> 항목에서 '미흡' 판정이 8.3%임.

(6) 자료처리 및 집계

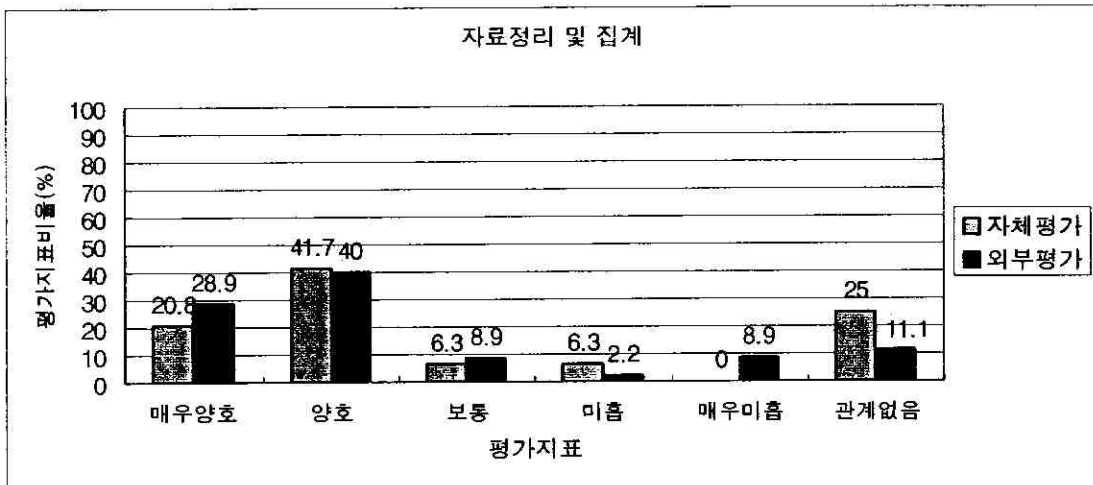
○ 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.0	0.05	0.9428	2-> 4-> 3-> 1	4.6/3.5
외부평가	3.9	0.14		2-> 1= 3-> 4	4.2/3.7

○ t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.4044	0.3751	0.3784
양측검정	0.8088	0.7502	0.7567

○ 응답현황



- <1.조사 무응답 실태 파악 및 분석>과 <3.모수추정 방법의 적절성>은 외부평가가 더 높게 나타남.
- <4.다른 통계와의 비교> 항목은 외부평가에서는 최저점을, 자체평가에서는 비교적 높은 점수를 받음.
- 외부평가는 <4.다른 통계와의 비교>항목의 '미흡'과 '매우 미흡'에 대한 비율이 20.0%로 높게 나타났음.
- 자체평가 결과 <1.조사 무응답 실태 파악 및 분석>과 <3.모수추정 방법의 적절성> 그

리고 <4.다른 통계와의 비교> 항목에서 '양호'의 비율이 '매우 양호'의 비율보다 높게 나타남. 특히 <1.조사 무응답 실태 파악 및 분석>의 '양호'와 '매우 양호'를 합한 비율이 25%로, 모든 자체평가 항목에서 가장 낮은 점수를 받았음.

- 외부평가에서 <2.입력자료의 내용 검사 시스템 적절성>에서 '매우 양호'의 비율이 '양호'의 비율보다 높게 평가되었음.
- '관계없음' 항목이 자체평가에서 더 많이 발생

(7) 자료공표

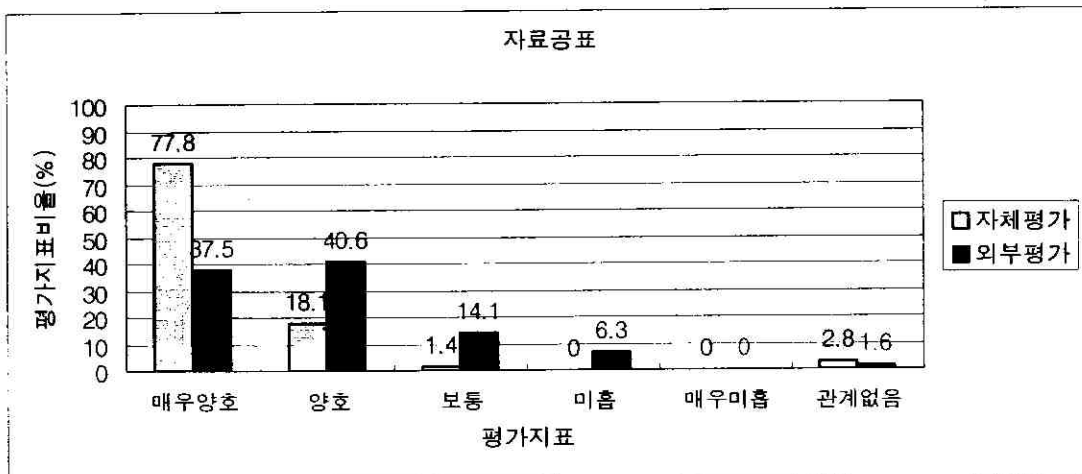
- 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.8	0.256	0.06599	5-> 3= 4= 6-> 2-> 1	5.0/4.3
외부평가	4.1	0.256		4-> 3-> 1= 6-> 2-> 5	4.5/3.8

- t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.00276	0.00057	0.00057
양측검정	0.00552	0.00113	0.00113

- 응답현황



- ‘양호’에 대한 평가가 극단적으로 바뀜. <1.간행물 사전 점검체계>와 <3.자료공표 시의성>은 외부평가에서는 높은 점수, 자체평가에서는 낮은 점수. <5.공표사전 예고>와 <6.사전예고 준수> 항목은 그 반대의 결과가 나타남.
- ‘미흡’에 대한 평가에서 자체평가는 모든 항목 ‘0’, 그러나 외부평가에서는 <5.공표사전 예고>는 무려 27.3%가 ‘미흡’
- <3.자료공표의 시의성>, <2.설명자료에 대한 충실성>, <5.공표사전 예고> 그리고 <6.사전예고 준수> 항목에서 평가 점수가 크게 차이가 남.(0.6, 0.8과 1.2점)
- 자체평가에서 <1.간행물 사전 점검 체계> 항목이 ‘양호’의 비율이 ‘매우 양호’의 비율보다 높게 나타남.
- 외부평가에서는 <4.자료접근 동시성>과 <5.공표사전 예고>항목에서 ‘매우 양호’의 비율이 ‘양호’의 비율보다 높게 나타남.
- ‘관계없음’ 항목에서 차이가 보임

(8) 자료이용

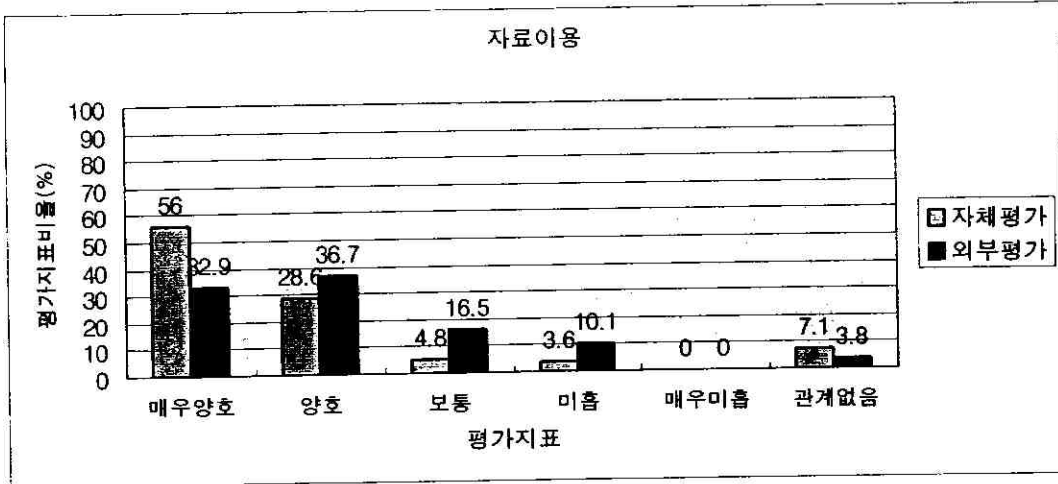
- 기초분석

	평균	표준편차	상관계수	순위(내림차순)	최대/최소
자체평가	4.5	0.47	0.9534	8-> 9-> 1-> 2-> 6-> 7-> 5	5.0/3.8
외부평가	4.0	0.63		9-> 8-> 1=> 2-> 6-> 7-> 5	4.7/3.1

- t-test(p-value)

	대응비교	등분산 비교	이분산 비교
단측검정	0.00061	0.0579	0.0590
양측검정	0.00123	0.11583	0.1180

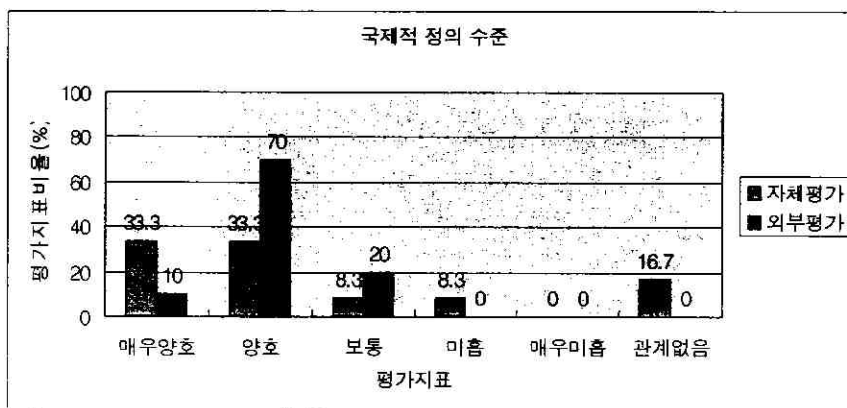
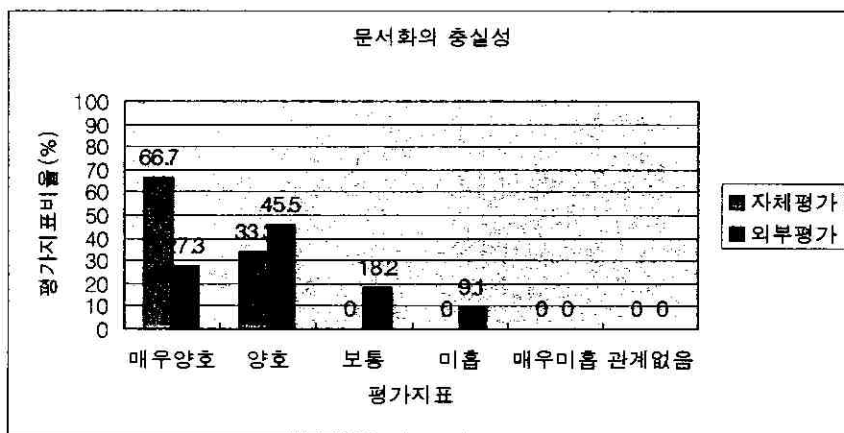
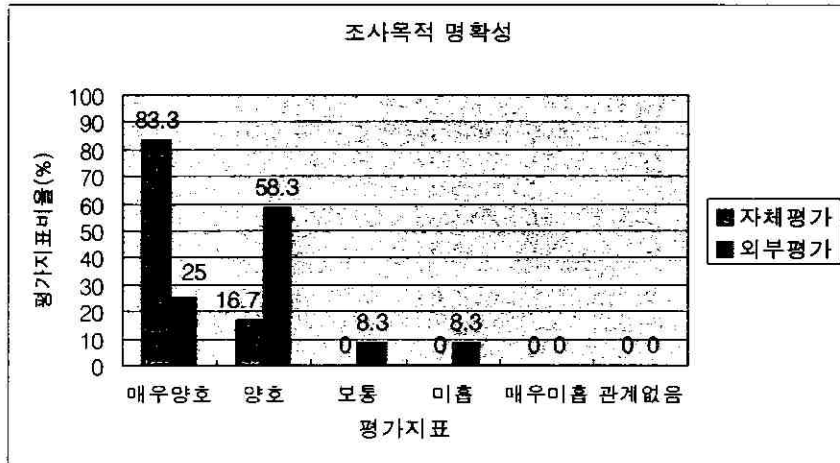
○ 응답현황

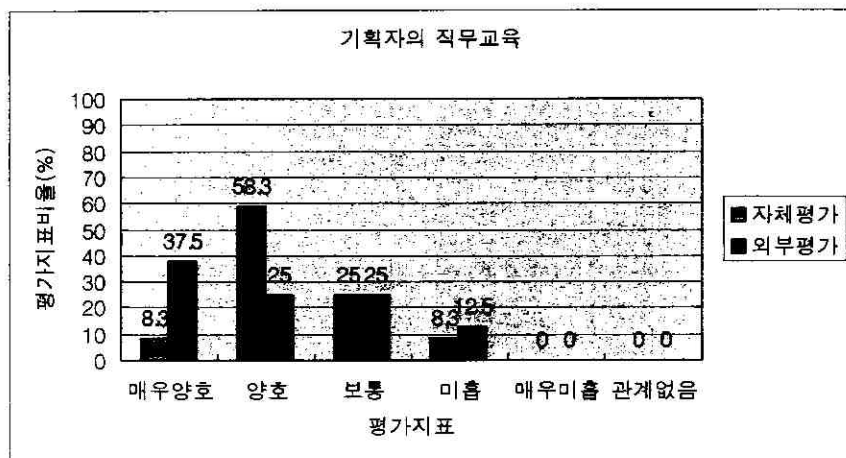
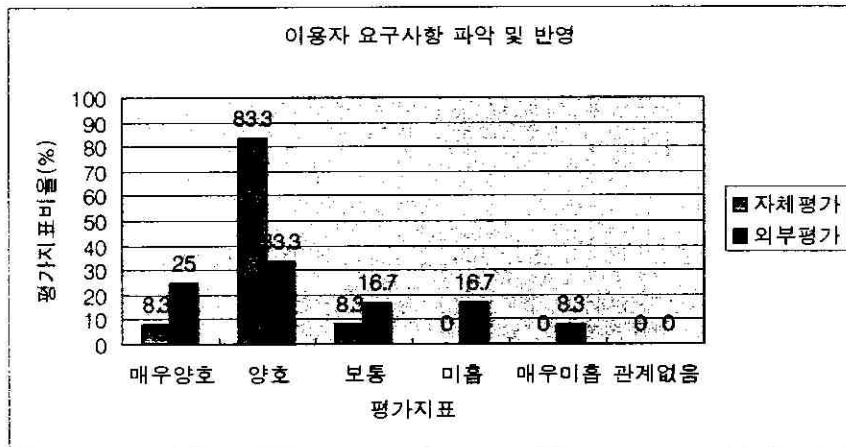
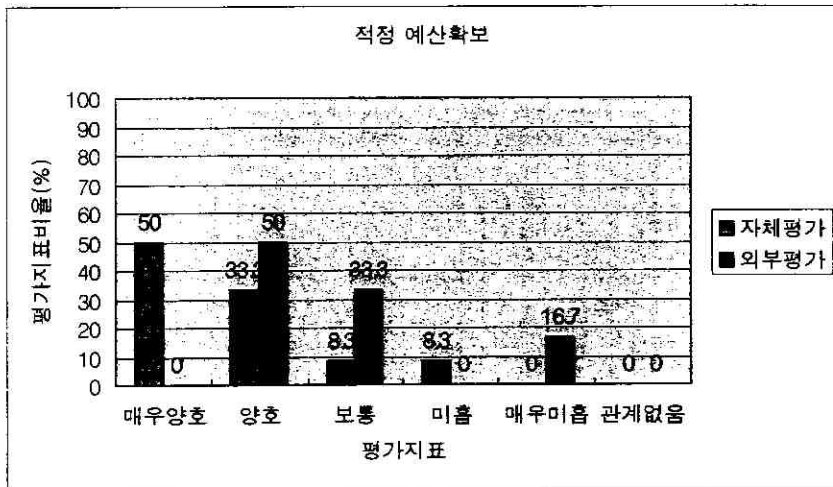


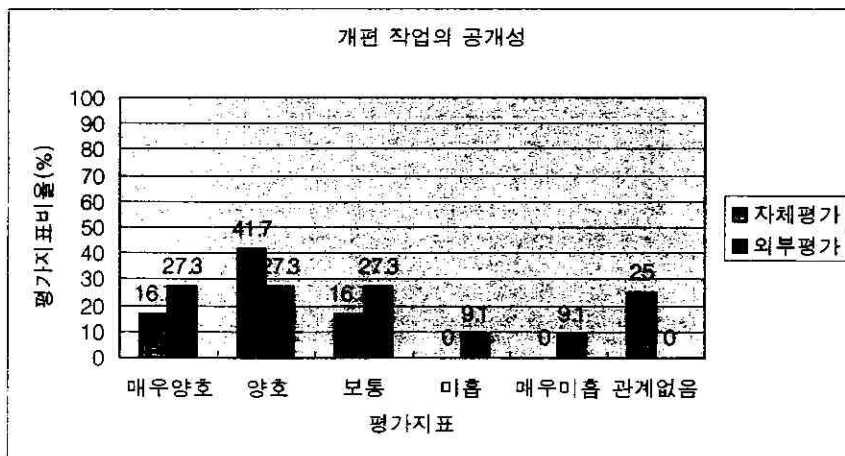
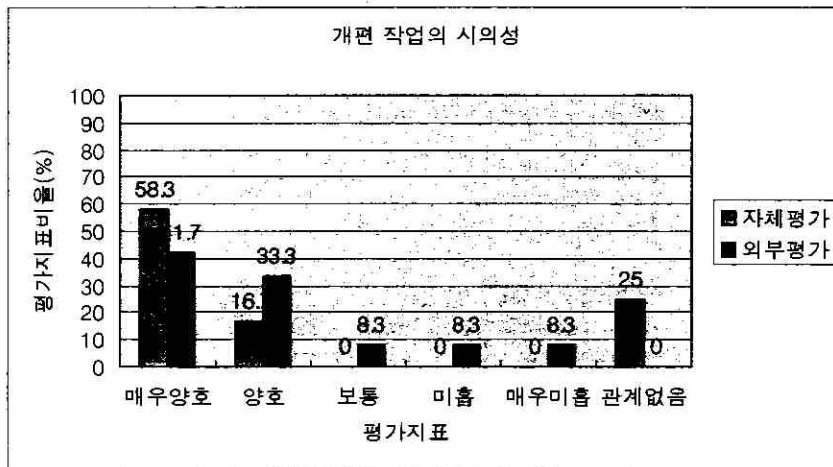
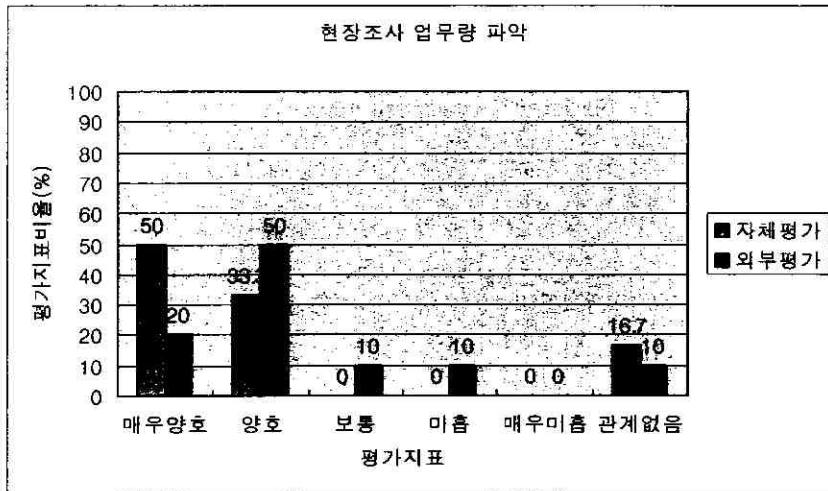
- <5.자료의 국제적 비교성>, <6.자료의 시계열 비교성>, <7.원시자료의 이용> 그리고 <8.응답자 비밀보호> 항목에서 자체평가와 외부평가의 평균점수의 차이가 크게 나타남. (각각 0.7, 0.7, 0.7 그리고 0.6)
- <7.원시자료 이용> 항목의 '미흡'에 대한 평가. 외부평가에서는 무려 40%가 미흡 판정. 자체에서는 16.7%만 '미흡' 평가.
- <5.자료의 국제적 비교성> 항목의 '미흡' 평가. 자체평가에서는 '0'. 그러나 외부평가에서는 27.3%
- 자체평가를 3점 척도로 분석할 때, <5.자료의 국제적 비교성> 항목이 50%의 '양호' 판정을 받음.
- 자체평가 결과 <5.자료의 국제적 비교성>, <6.자료의 시계열 비교성>과 <7.원시자료의 이용> 항목에서 '양호'의 비율이 '매우 양호'의 비율보다 높게 나타났음.
- 외부평가 결과 <5.자료의 국제적 비교성>과 <7.원시자료의 이용> 항목에서 40% 이하의 '양호' 및 '매우 양호' 판정을 받음. 극히 낮은 만족도임이 분명.
- 결국 <5.자료의 국제적 비교성>, <6.자료의 시계열 비교성>과 <7.원시자료의 이용> 항목에서 절차의 재검토가 필요함.

<부록 2> 각 문항 및 평가방법에 따른 비교

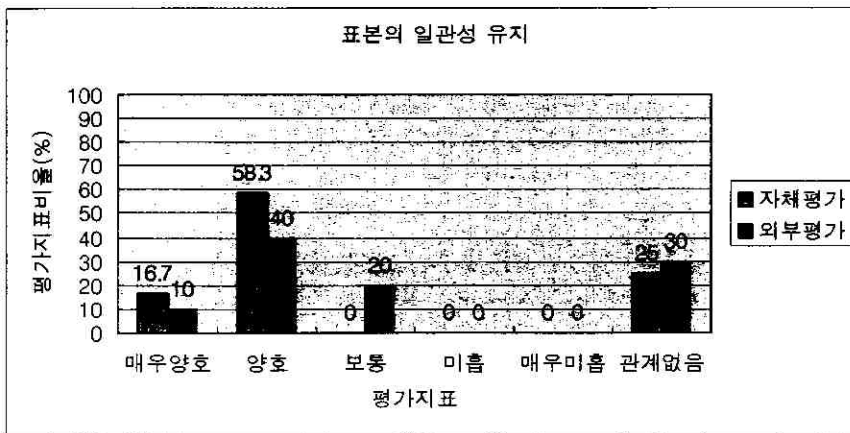
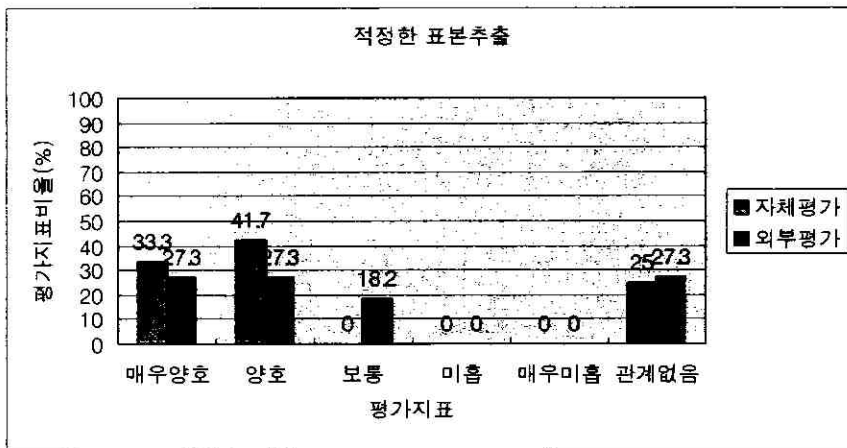
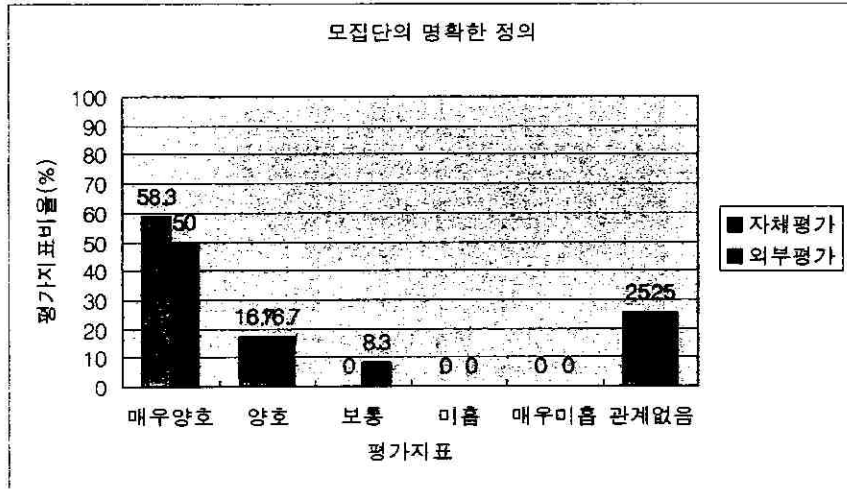
(1) 조사기획

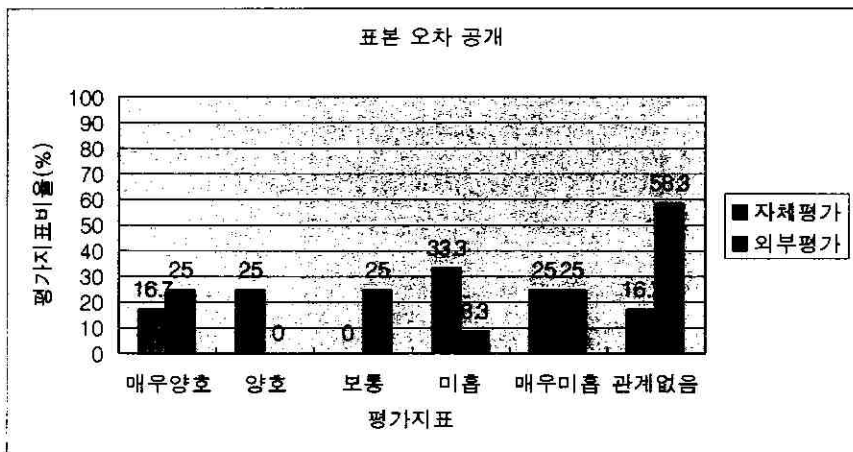
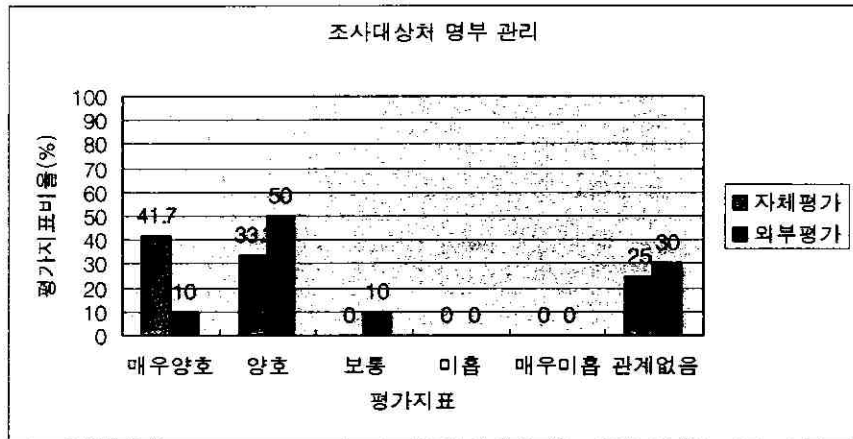
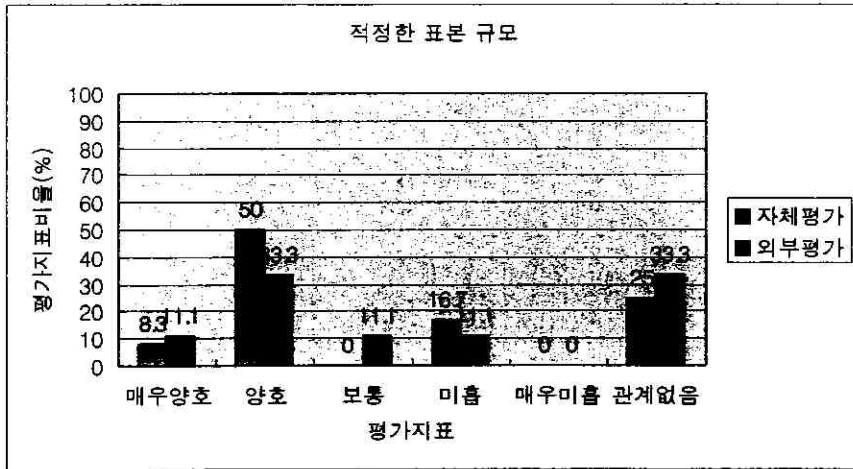


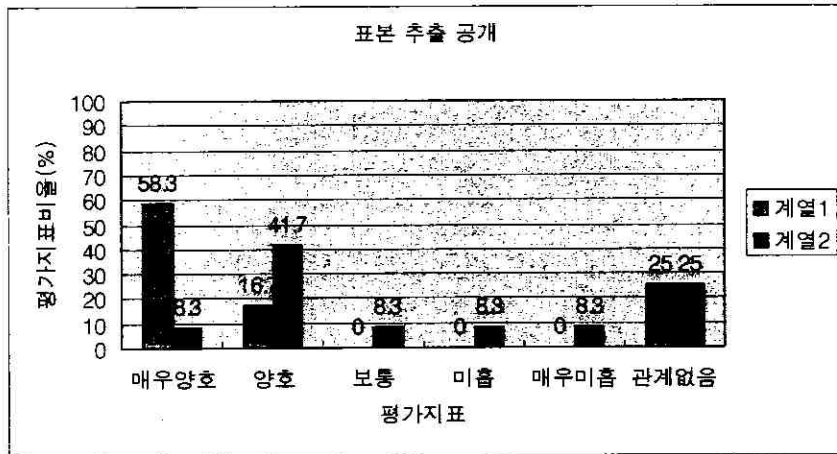




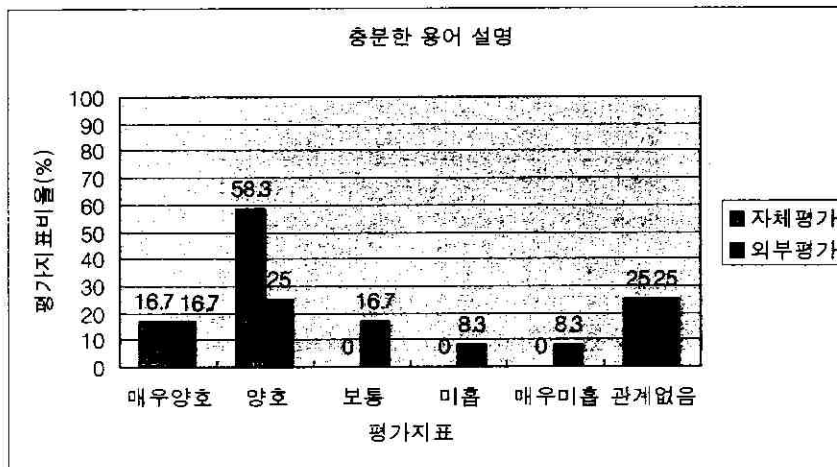
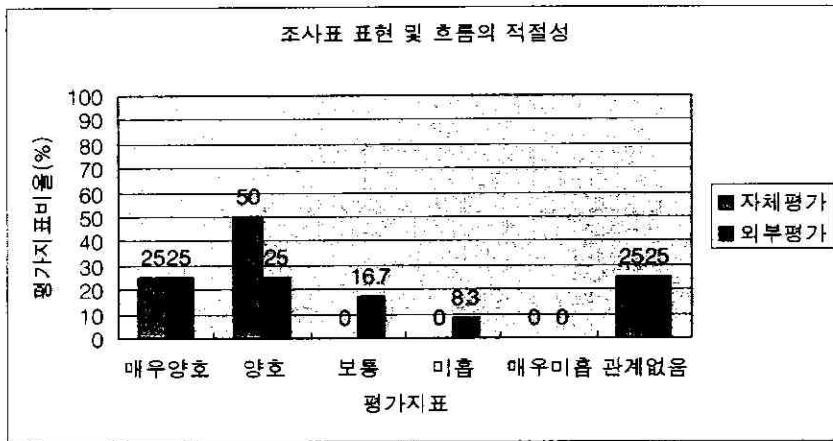
(2) 모집단 및 표본

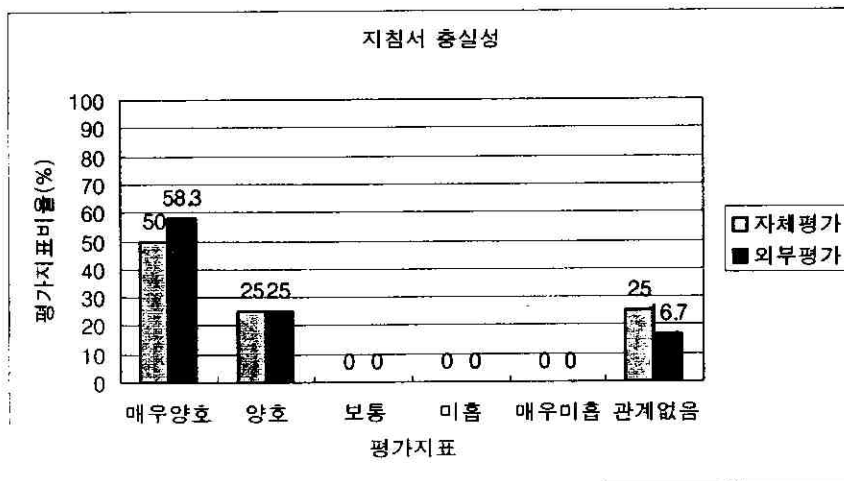
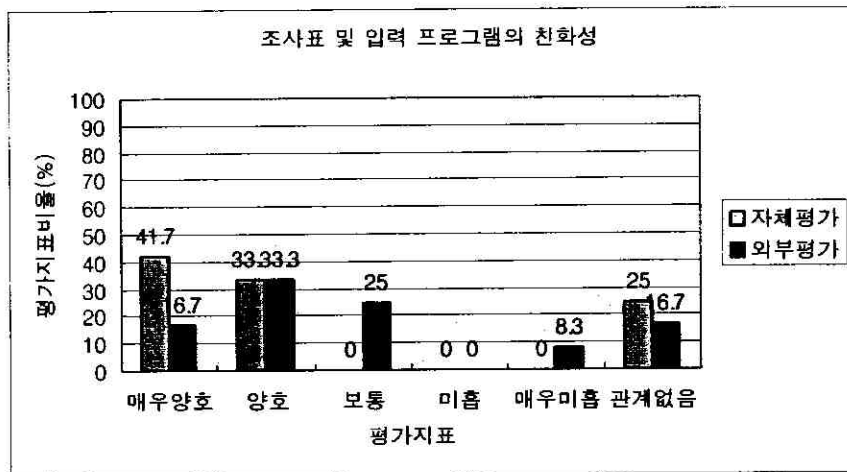
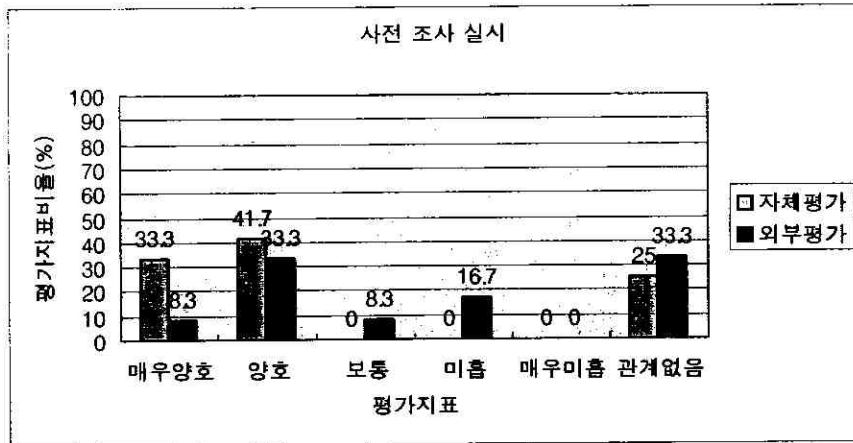




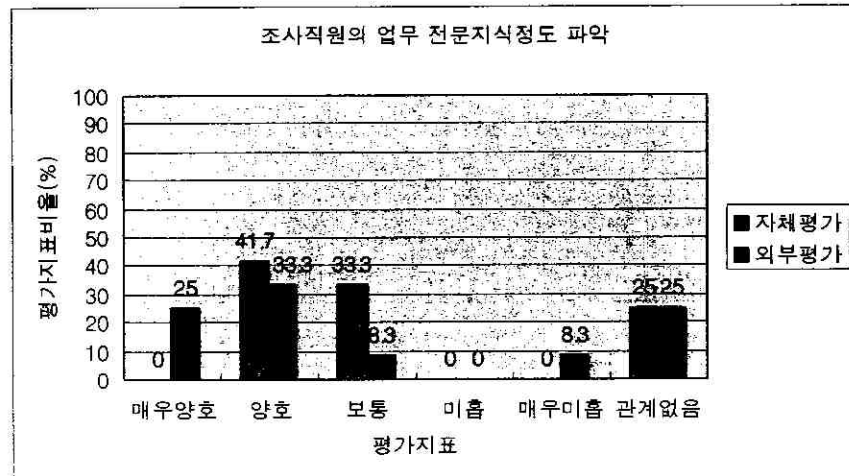
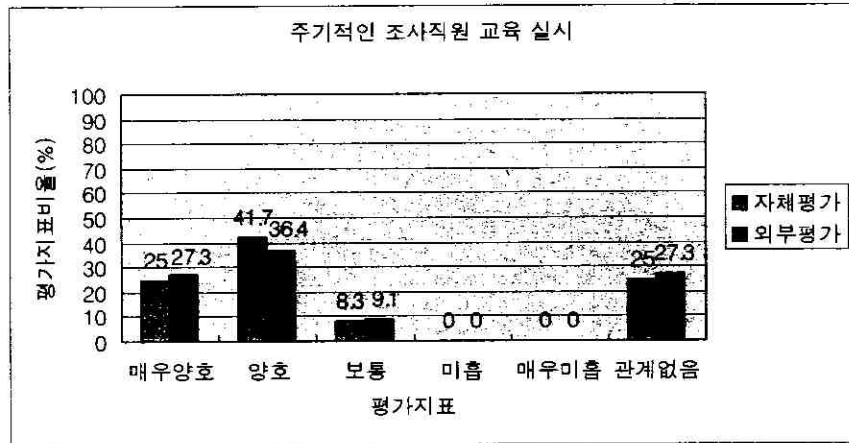
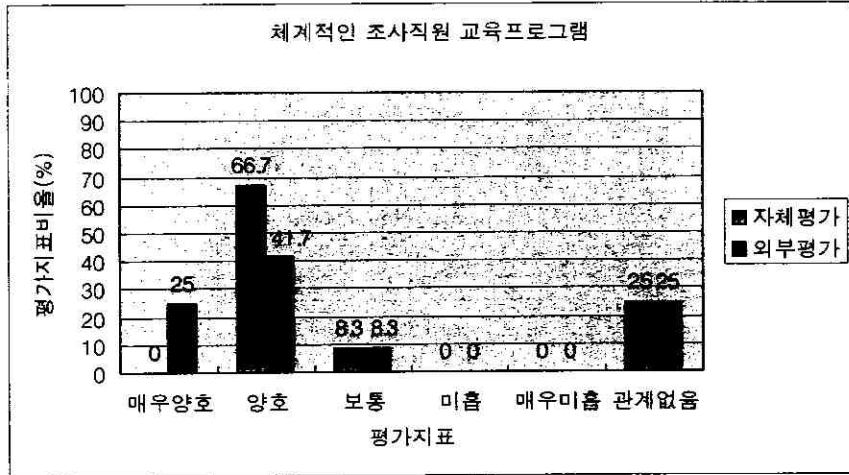


(3) 조사표설계

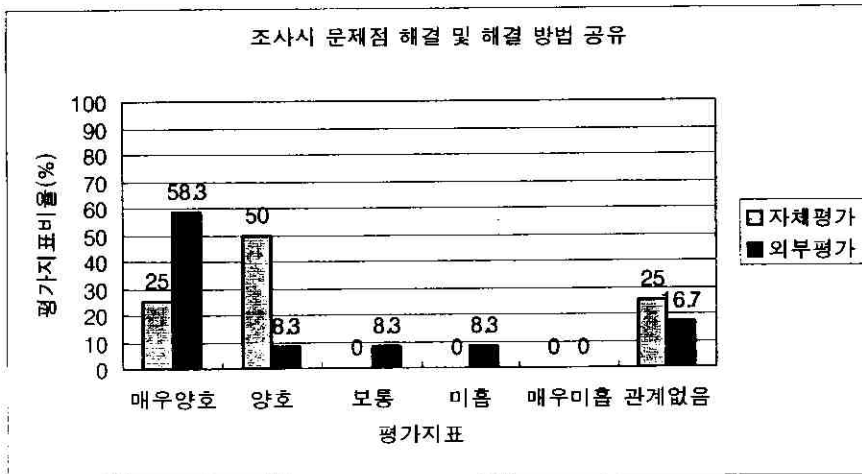
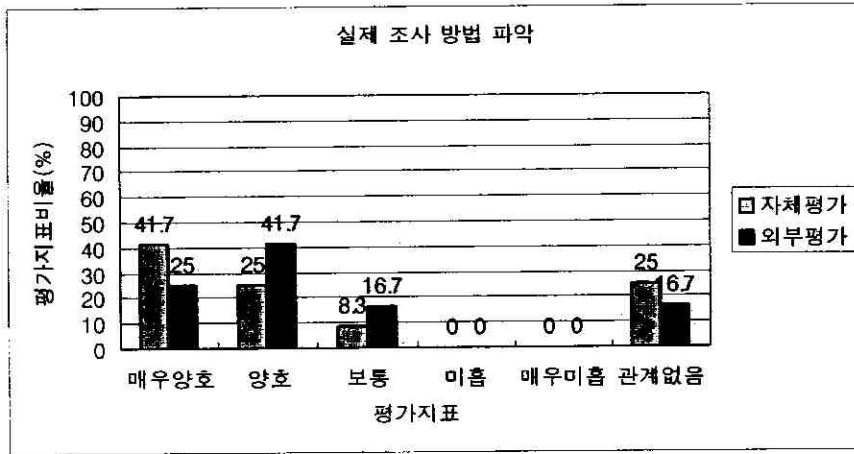
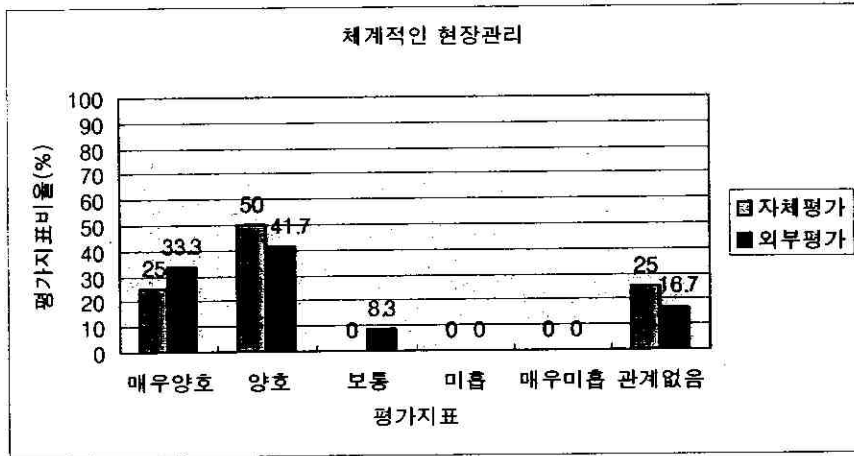


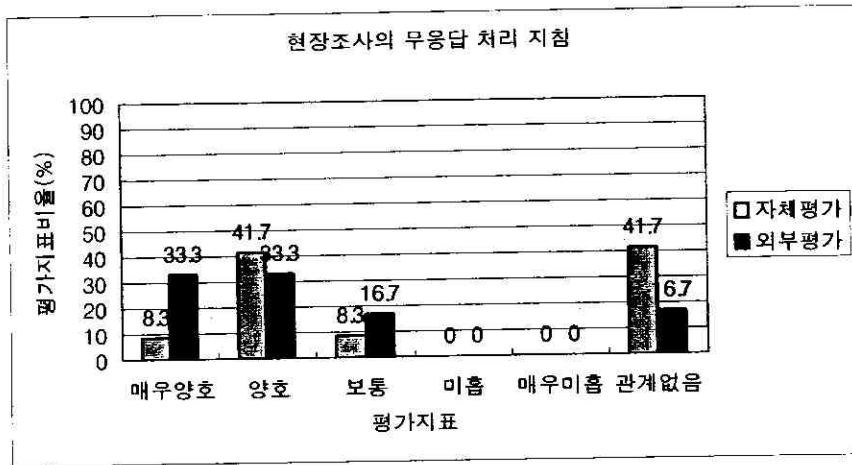
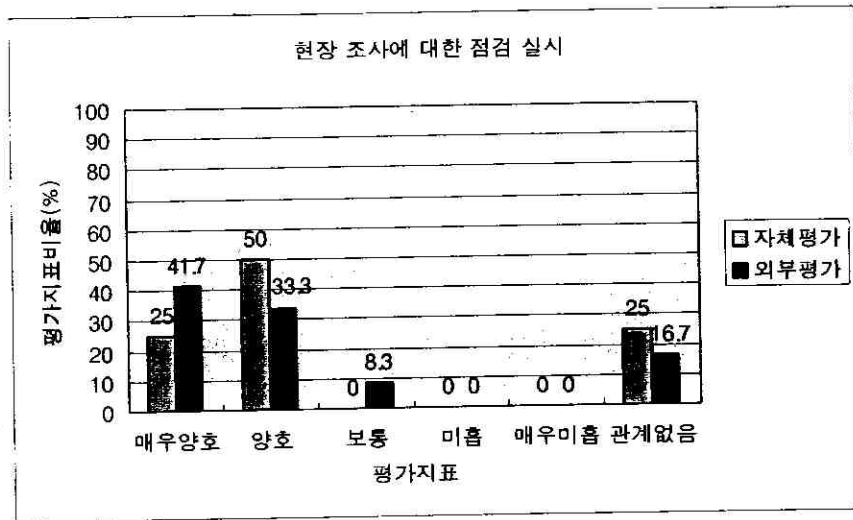


(4) 조사직원의 관리

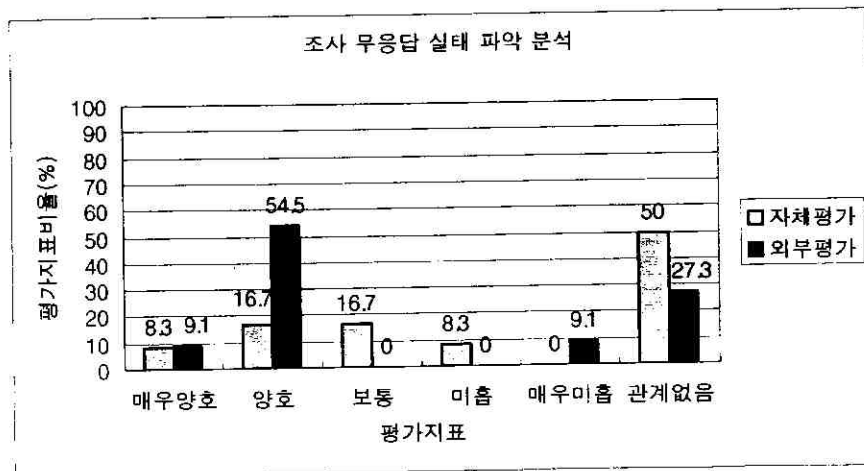


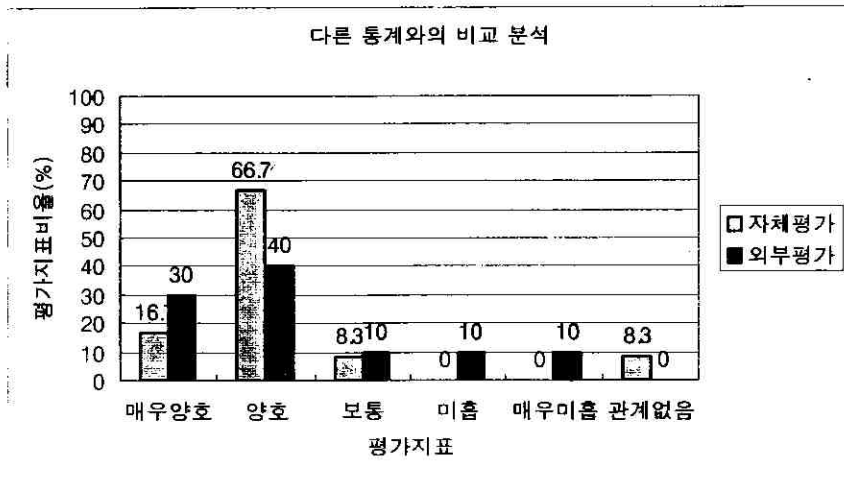
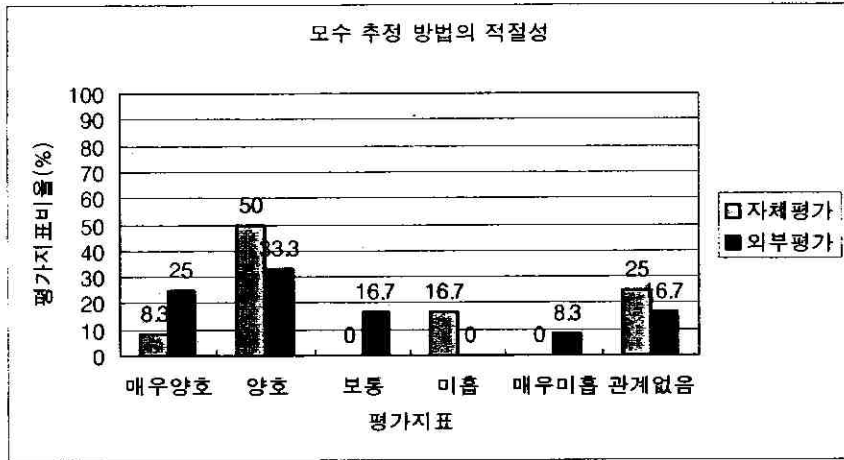
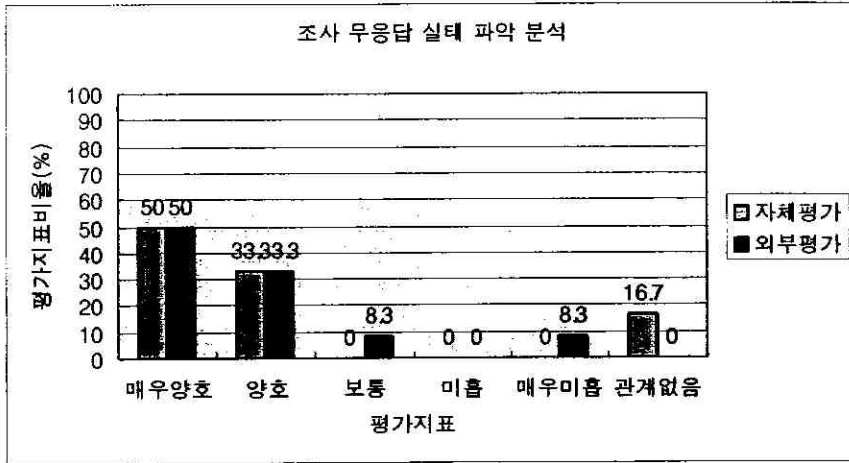
(5) 조사실시



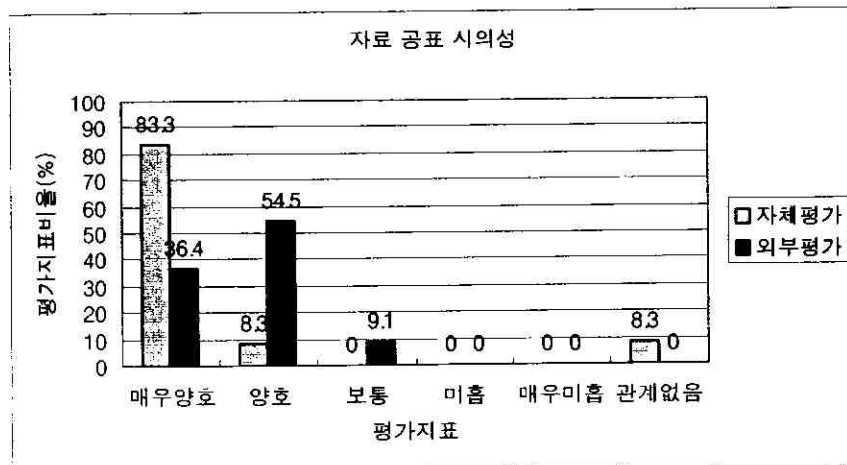
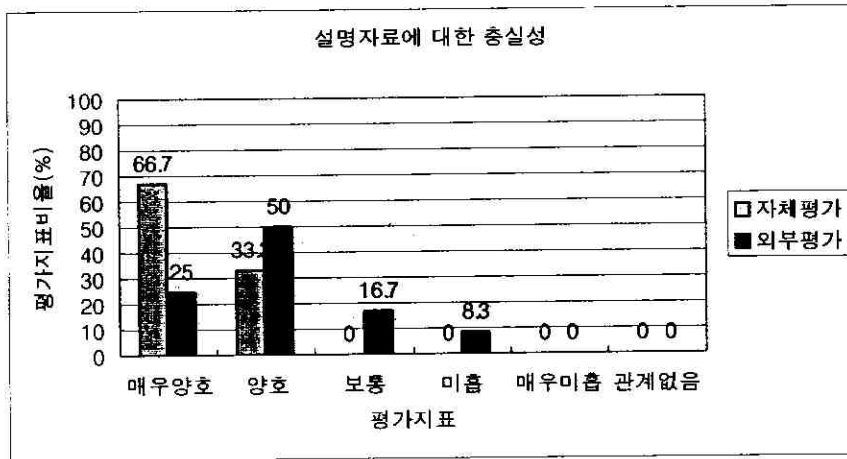
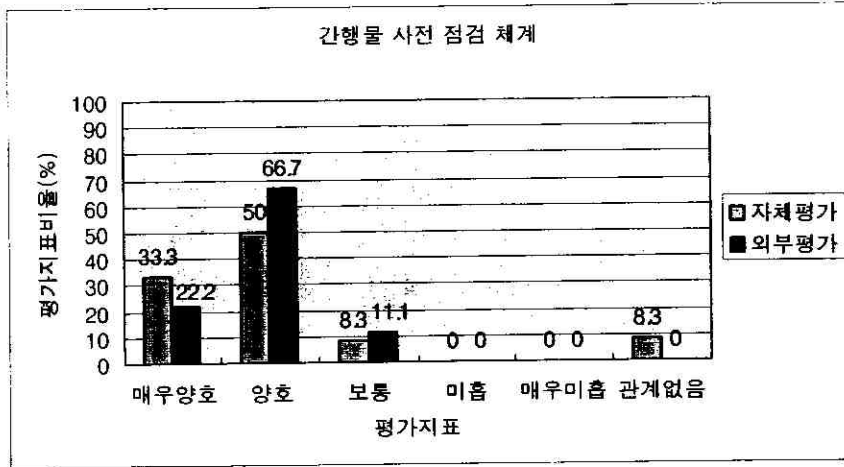


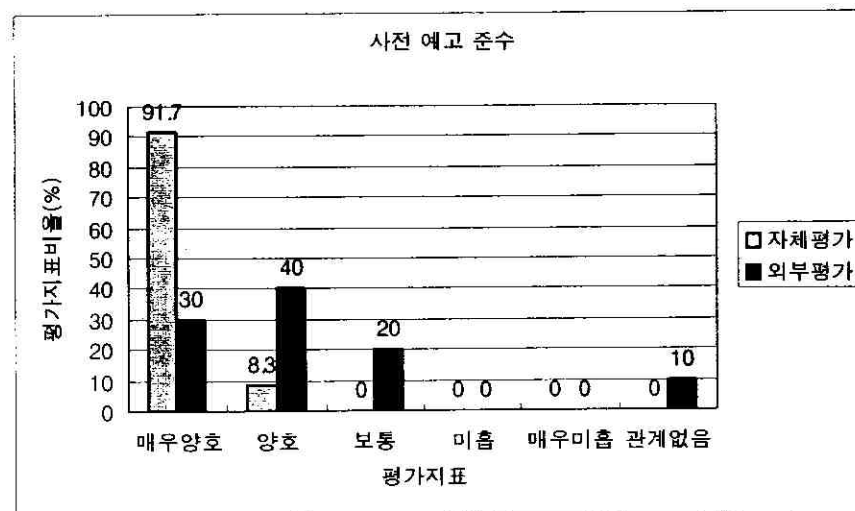
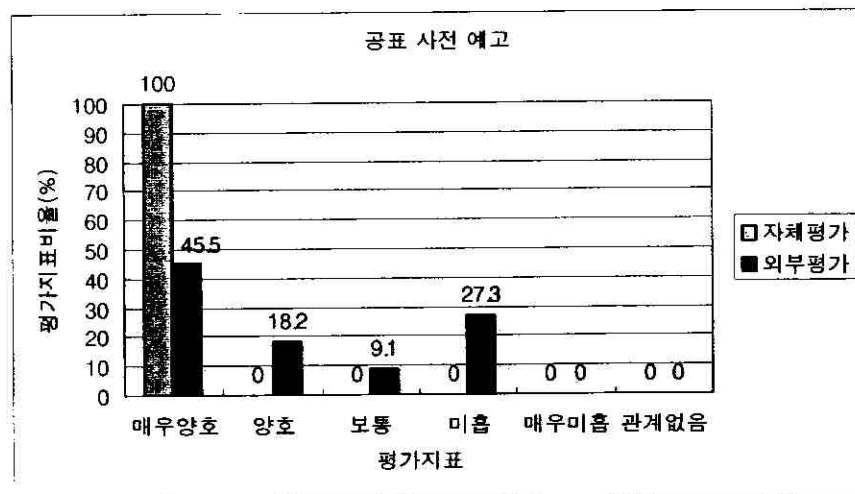
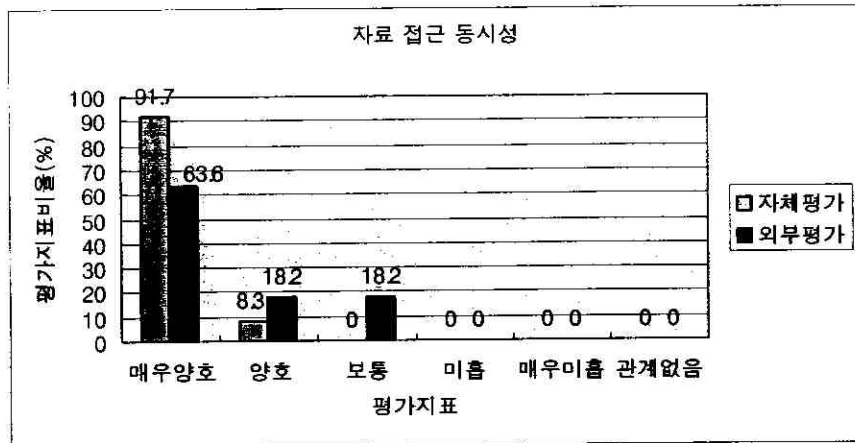
(6) 자료처리 및 집계



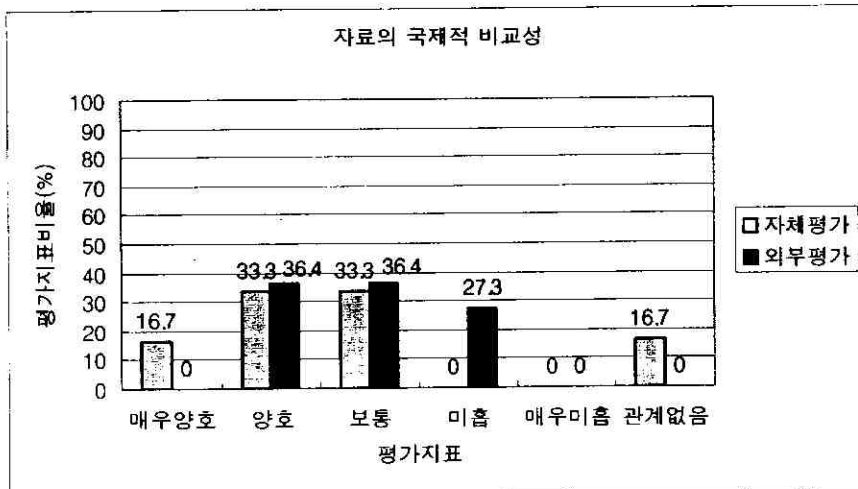
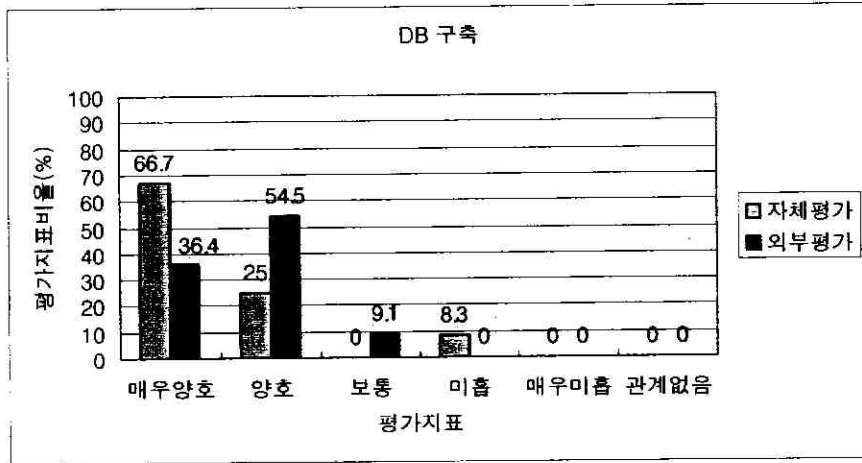
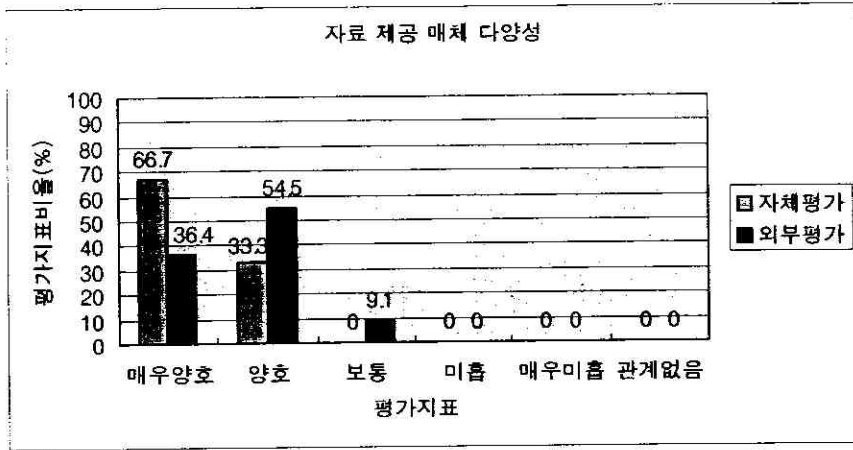


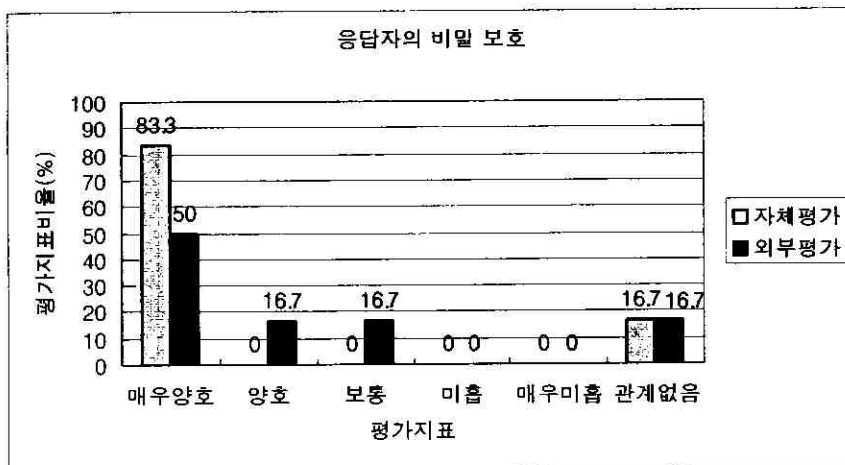
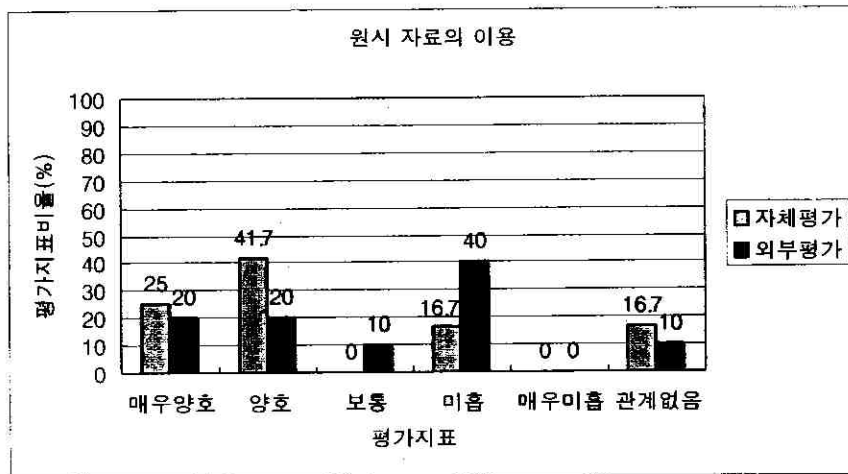
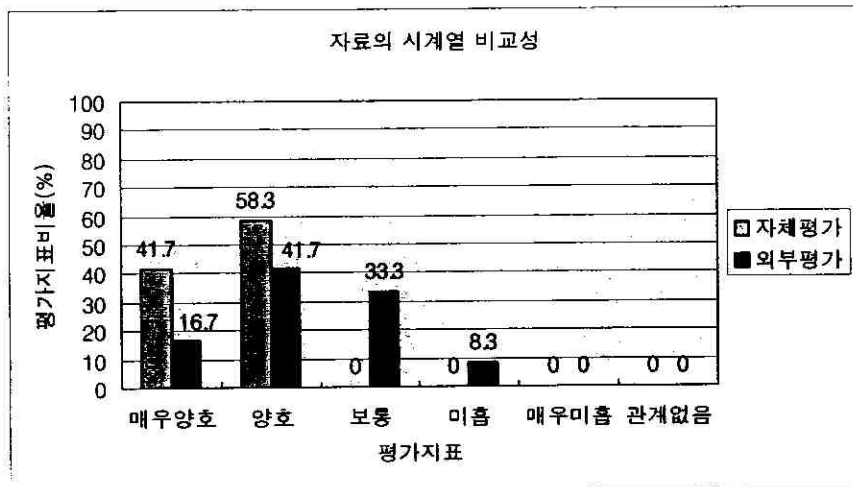
(7) 자료공표



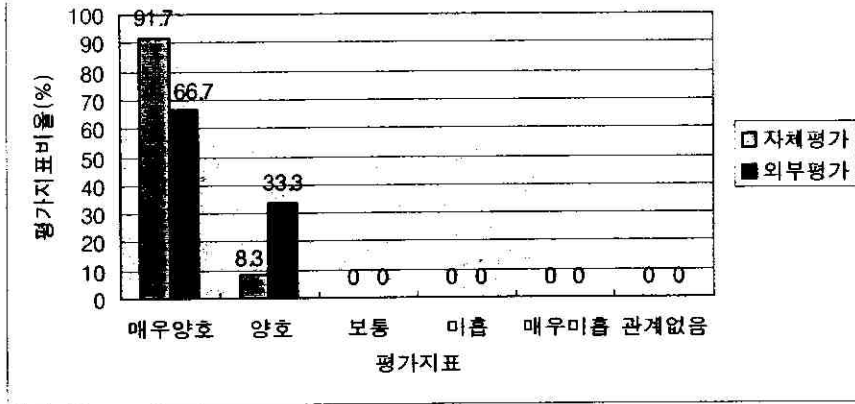


(8) 자료이용





자료의 문의처 제공



<부록 3> 평가방법에 따른 평균 점수 차이에 대한 분석

: <자체평가 점수 - 외부평가 점수>에 따른 분석

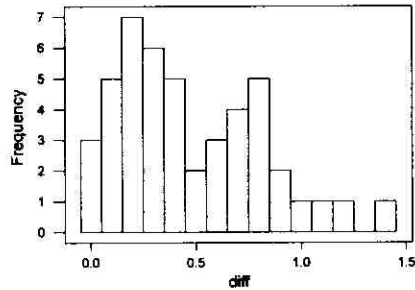
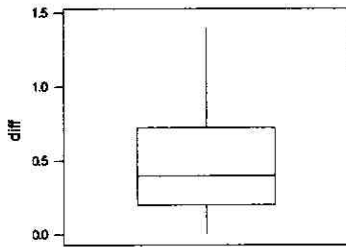
가. 문항별 차이에 의한 분석

(1) 기초통계 분석

Variable	N	Mean	Median	TrMean	StDev	SE Mean
diff	46	0.4739	0.4000	0.4571	0.3435	0.0506

Variable	Minimum	Maximum	Q1	Q3
diff	0.0000	1.4000	0.2000	0.7250

(2) 그래프에 의한 분석



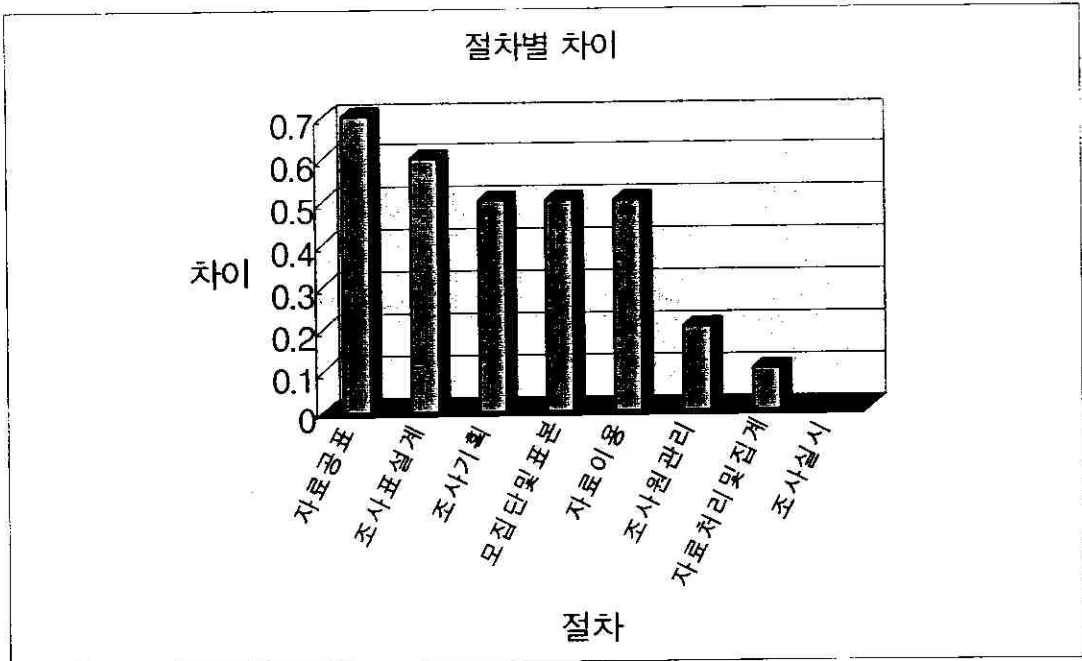
나. 절차별 차이에 의한 분석

(1) 기초통계 분석

Variable	N	Mean	Median	TrMean	StDev	SE Mean
diff-1	8	0.3875	0.5000	0.3875	0.2532	0.0895

Variable	Minimum	Maximum	Q1	Q3
diff-1	0.0000	0.7000	0.1250	0.5750

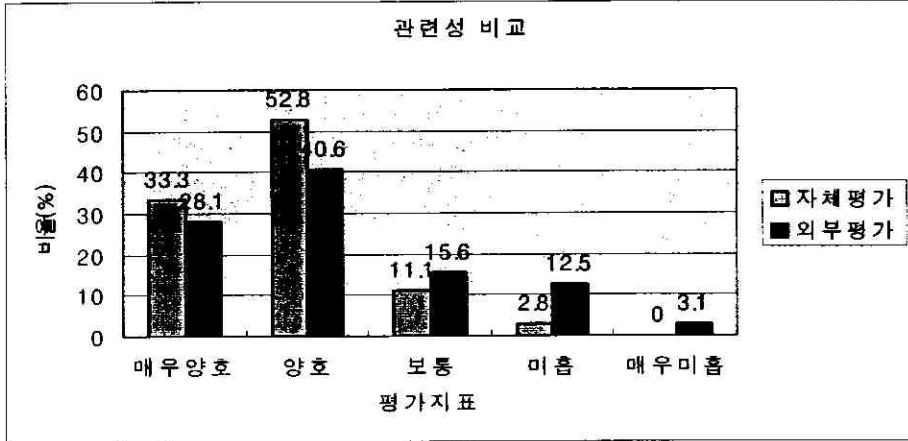
(2) 그래프에 의한 분석



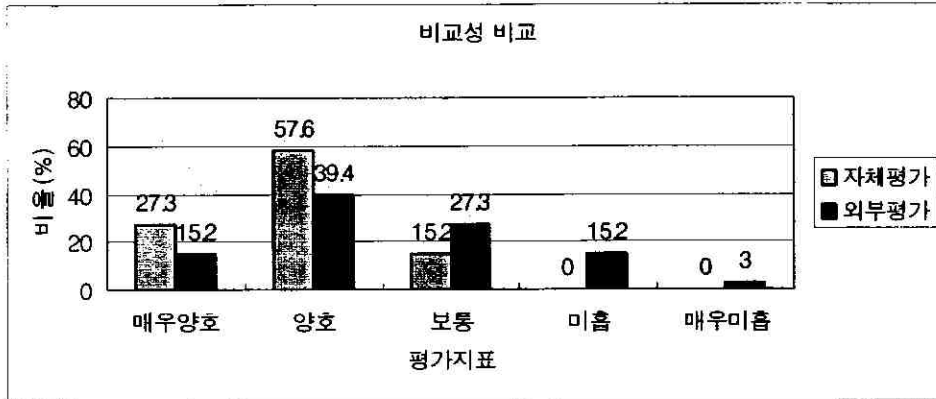
<부록 4> 차원별 분석 결과(그래프 모음)

(1) 차원 및 평가방법에 따른 응답현황

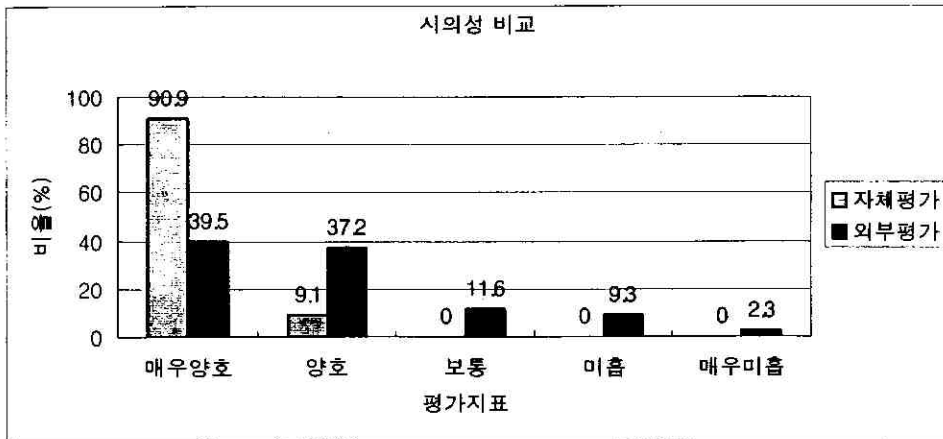
○ 관련성



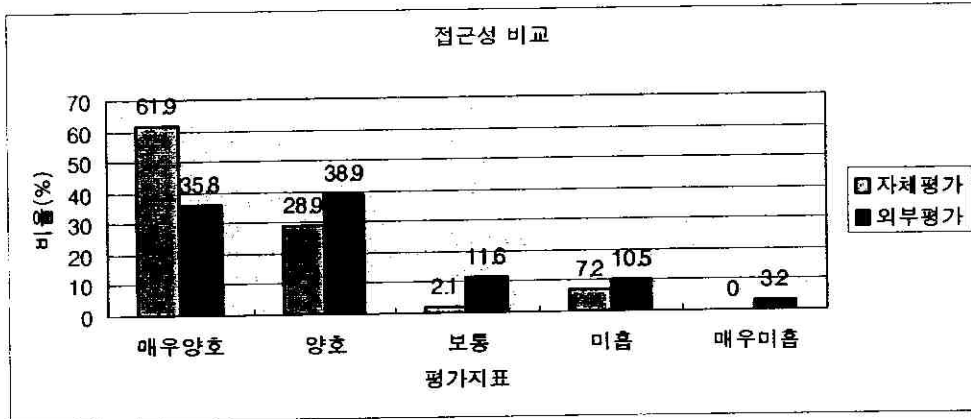
○ 비교성



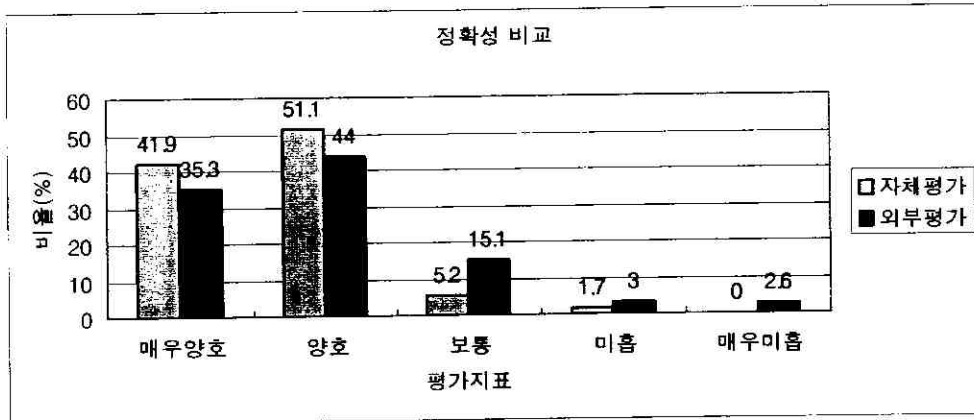
○ 시의성



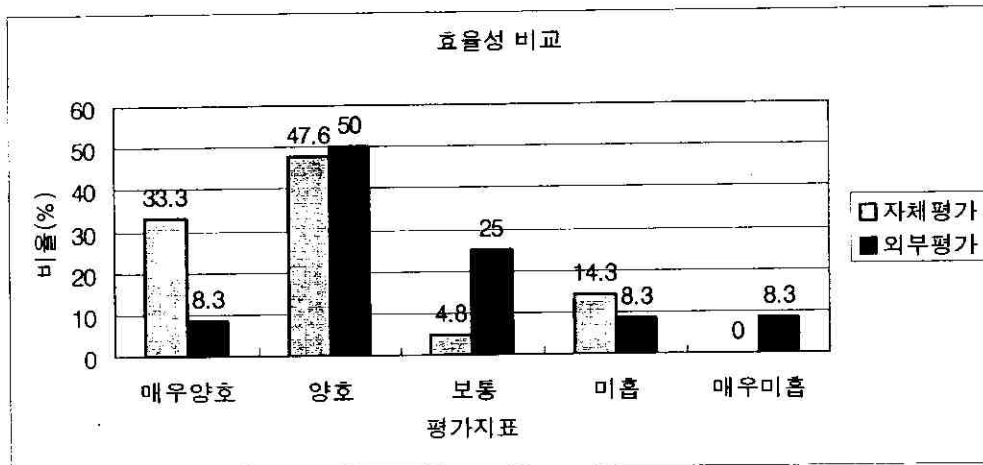
○ 접근성



○ 정확성

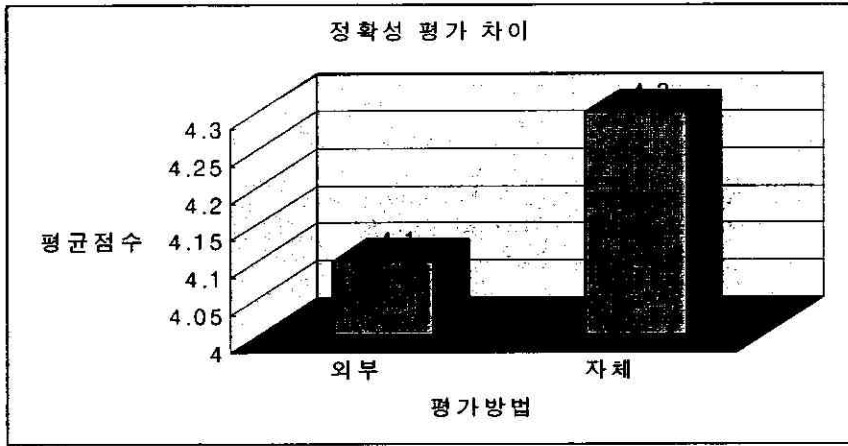


○ 효율성

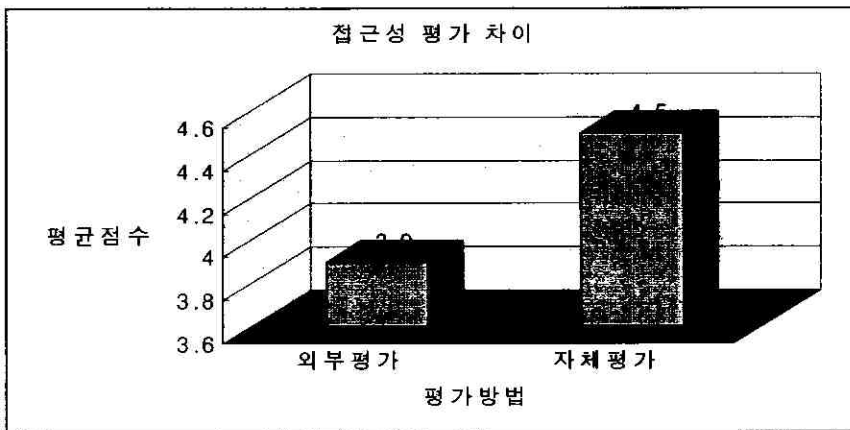


(2) 차원별 차이에 의한 분석

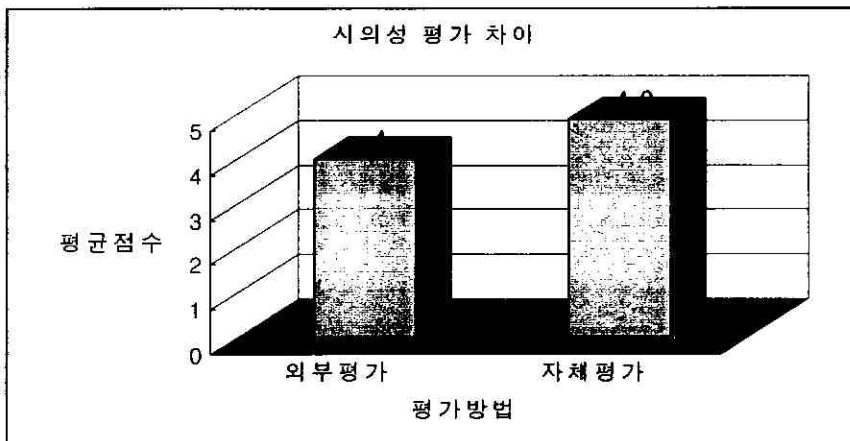
○ 정확성



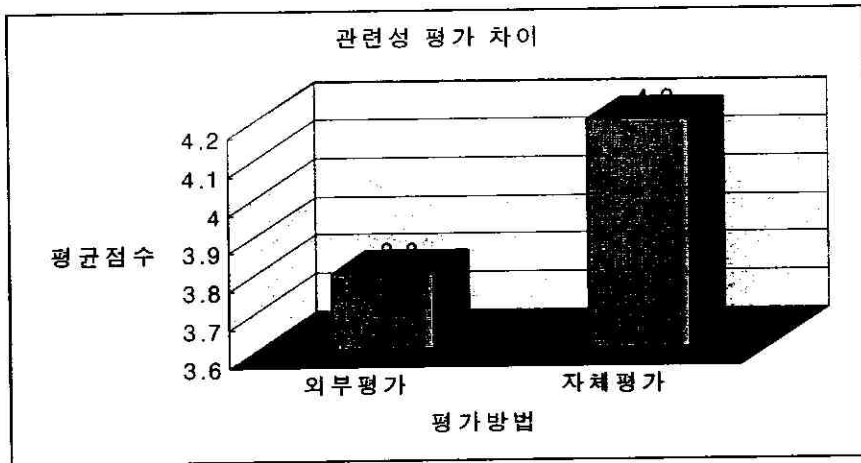
○ 접근성



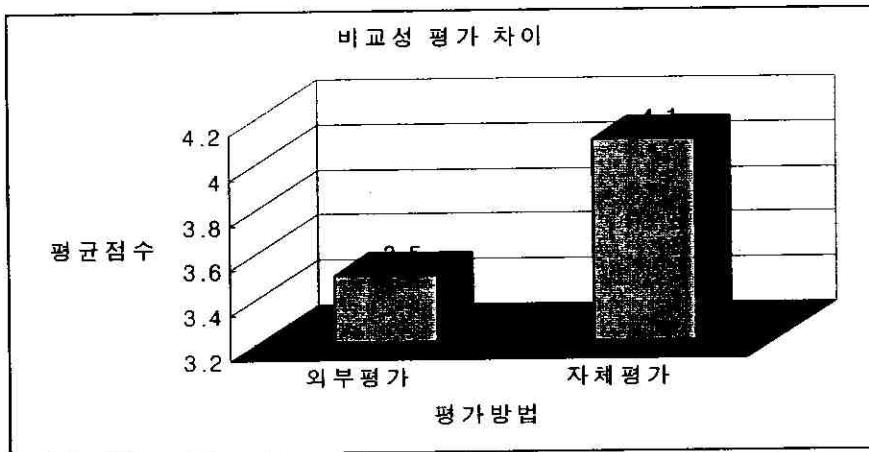
○ 시의성



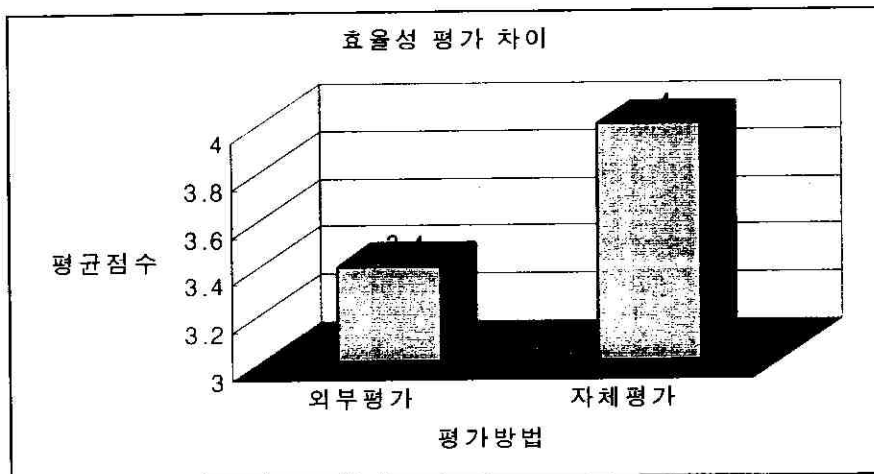
○ 관련성



○ 비교성



○ 효율성



<부록 5> 기타 : '관계없음'에 대한 도수 차이

