

연구논문

중앙정부의 사회복지시설 평가에 대한 비판적 분석 메타평가를 중심으로

임 성 옥

1998년 사회복지사업법의 개정에 따라 1999년부터 사회복지시설에 대한 평가가 중앙정부에 의해 수행되고 있으나, 평가활동의 적정성과 질, 그리고 평가결과와 활용에 대한 의문이 제기 되면서 사회복지시설 평가에 대한 메타평가의 필요성이 제기되고 있다. 이에 본 연구는 중앙정 부가 1999년부터 2001년까지 실시한 사회복지시설 평가사업을 대상으로 메타평가를 수행하 였다. 메타평가 결과, 평가투입자원 영역이 가장 낮은 평가를 받았으며, 그 다음으로 평가결과 영역이 부정적인 것으로 나타났다. 특히 평가결과 하위영역 중 평가결과보고서의 구체성과 평 가결과정보의 확산성이 보다 낮게 나타났다. 한편, 평가수행 영역은 전체 평균 이상을 약간 상 회하는 점수로 나타나 평가투입자원 및 평가결과 영역보다는 다소 긍정적이었다. 특히 자체평 가 후 현장평가를 실시하는 평가수행 절차에 대해서는 대체로 긍정적인 것으로 나타났으나, 평 가수행 영역 중 평가수행방법에서 서비스 이용자 및 종사자의 만족도 조사에 대한 신뢰성은 매 우 낮은 것으로 나타나고 있다.

이러한 결과를 바탕으로 본 연구에서는 현 평가제도의 개선 및 평가활용성을 제고하기 위한 방안을 종합적이고 구체적으로 제시하였는데, 이는 사회복지시설 평가제도의 발전방안 마련에 유용한 기초자료로 활용될 수 있을 것이다.

국문 키워드 : 메타평가, 평가활용, 평가체계의 질

임성옥은
경북대학교에서
사회복지학 박사학
위를 받고, 현재 춘
해대학 사회복지과
교수로 재직중이다.

solim@choonhae.ac.
kr

1. 서론

한국 사회복지계에서 평가에 대한 관심은 1990년대 이후 급속도로 증 대되어 왔는데, 이는 사회복지서비스 공급의 지속적인 양적 팽창과 함께 그에 따른 책임성이 제기되면서부터이다. 평가에 대한 사회복지

계의 본격적 관심은 서울시가 1996년부터 서울지역의 사회복지관 평가를 실시하면서 시작되었으며, 1998년 사회복지사업법의 개정으로 사회복지시설 평가가 법제화되면서 본격화되고 있다.

이에 따라 사회복지시설 및 복지 프로그램에 적합한 평가모형 개발 등에 관한 연구(황성철·강혜규, 1994; 이변송·김진욱, 1996; 조운희, 1998; 최일섭 외 4인, 1998; 최재성, 1998; 김영모 외 4인, 2000; 변재환 외 10인, 2000)가 집중적으로 이루어졌고, 정부에서는 1999년부터 한국보건사회연구원에 의뢰해 전국 사회복지시설을 평가하기 시작하였다. 1999년부터 2001년까지 1차적으로 평가받은 시설은 총 1,060개 사회복지시설로, 설립한 지 3년 이상이 된 시설은 거의 대부분 1회 이상 평가를 받았고, 2002년도부터는 2차 평가가 시작되었다.

법적 근거에 기반한 본격적 평가는, 평가를 실시하고 있는 정부의 입장에서는 공공자금에 대한 책임성 확보를 위한 노력이며, 서비스 이용자 입장에서는 적정한 양질의 서비스를 제대로 받고 있는지 확인하는 방편이기도 하다(Royse et al., 2001; 김통원, 2001). 이처럼 평가를 통해서 그 기관의 구성원들이 보다 더 양질의 서비스를 제공하기 위해 필요한 정보들을 얻는 것은 매우 중요한 사안이다. 그러한 정보를 바탕으로 조직의 구성원들은 기관 내부에서 자발적으로 '지속적인 질(*quality*) 개선'을 이루어 나가는 것이 바로 평가결과를 최적으로 활용하는 것이 된다.

그러나 평가활동의 양적 증가와 평가결과 활용의 중요성에도 불구하고, 평가가 제대로 이루어진 것인지, 평가목적은 제대로 달성되었는지, 또 실제로 피평가시설이 평가결과를 제대로 활용하고 있는지에 대해서는 명확한 대답을 구하기가 어렵다. 왜냐하면 그동안 사회복지시설 평가가 평가활동의 결과에 중점을 둔 사후 보고에만 치중되어, 평가 이후 결과가 어떻게 활용되고 있는지에 대해서는 평가를 실시하는 정부나 평가받는 시설 모두 큰 관심을 두지 않았기 때문이다.

뿐만 아니라 현재의 평가제도는 평가과정상의 문제, 평가지표의 타당성 문제 등 많은 문제점들을 노출해 왔으며(정선욱·김진숙, 2000;

최재성, 2000; 김통원, 2001), 특히 서울시 사회복지관 평가의 경우 현장에서 실무자들이 집단적으로 평가를 거부하는 운동까지 전개되었다. 이처럼 사회복지시설 평가가 책임성 제고 및 서비스의 질 개선이라는 목적을 달성하기 위해 많은 노력과 자원을 투입하여 실시되었으나, 그 소기의 목적을 달성하지 못하고 있다면 평가활동 자체에 문제가 있다는 것을 의미한다. 따라서 평가제도를 처음 도입할 때 의도했던 목적과 기대효과가 어느 정도 달성되었으며, 과연 현재의 방식대로 계속 진행해야 하는지에 대한 검증이 필요하다.

이제는 단지 사회복지시설을 평가하고 있다는 사실에 만족할 것이 아니라, 평가가 제대로 수행되고 있는지, 도출된 결과는 의미 있는 것인지, 사회복지시설의 나아갈 방향을 제대로 제시해 주고 있는지에 대하여 관심을 가져야 할 시점이다. 이에 사회복지시설 평가사업에 대한 메타평가(*meta-evaluation*)의 필요성이 제기되고 있다. 메타평가란 현재 행해지고 있거나 이미 행해진 평가에 대한 재평가, 즉 상위 평가를 의미하는 것으로(Scriven, 1981; 김명수, 2000), 평가자가 선택하는 방법과 도구, 평가자의 개인적 배경, 전문적 훈련과 경험 등이 상이함으로 인해 발생하는 평가의 오류(*evaluation bias*)를 방지하고 평가가 본래 의도했던 결과를 달성하기 위해 필요한 절차이다(Worthen et al., 1997).

현대국가에서 수행하는 다양한 형태의 정책평가가 현실의 정책과정을 개선하기에는 평가결과의 질이 너무 낮다는 비판이 제기되면서, 평가의 질을 개선하여 평가활용을 증진시키기 위한 방법의 하나로 일찍이 메타평가의 필요성이 등장하였다. 여러 학자들이 이에 관한 이론적 논의를 시작하였으며(Cook & Gruder, 1978; Stufflebeam, 1981; Larson & Berliner, 1983), 국내에서도 공공정책 분야의 학자들 중심으로(오무근, 1986; 문영세, 1996; 문영세, 1997; 이찬구, 1997; 김명수, 2000) 일부 연구가 이루어져 왔다. 그러나 주로 공공정책 분야나 중앙정부의 심사평가제도를 중심으로 연구가 이루어졌을 뿐 사회복지학 분야에서 체계적으로 연구가 이루어지지 못했다.

그동안 사회복지시설 평가활용에 관한 연구는 전무하였고, 단지 사회복지시설 평가활동의 문제점을 분석하고 이의 해결방안을 모색하려는 연구가 일부 있었으나 평가활동의 특정 국면에 치우치거나 이론 중심적 한계로 인해 제한적일 수밖에 없었으며, 사회복지시설 평가사업에 대한 메타평가 관련 연구 또한 부분적으로 수행되었을 뿐이다 (김통원, 1997b; 정선옥·김진숙, 2000, 김통원, 2001).¹⁾

따라서 본 연구는 중앙정부의 사회복지시설 평가사업에 대해 메타평가를 종합적, 실증적으로 실시함으로써 평가전반의 문제점과 개선방안을 도출하여 평가제도의 질을 향상시키고 평가활용성을 제고하기 위해 수행되었다. 연구범위는 1999년부터 2001년까지 보건복지부가 주관하는 사회복지시설 평가활동에 한정하였으며, 연구방법은 서베이와 면접법을 병행하였다.

2. 메타평가의 이론적 고찰

일반적으로 메타평가란 현재 행해지고 있거나 이미 행해진 평가에 대한 재평가, 즉 상위평가를 의미한다. 그러나 엄밀한 의미에서의 메타평가는 보다 복합적 의미를 지니고 있는데, 메타평가에 대한 가장 좁은 의미로는 평가결과의 종합(*evaluation synthesis*)이라는 개념으로 메타평가를 이해하는(Chemlisky, 1985) 최협의의 관점이 있다. 평가종합이란 하나 혹은 여러 개의 기존 평가로부터 얻어진 결과를 재분석하는 것을 말한다. 이와 유사하게, 평가 자체에 대한 평가만을 의미하는 것으로, 체계적인 조사설계를 이용하여 이미 수행된 경험적 총괄평가만으로 한정(Cook & Gruder, 1978)하는 견해가 있다.

보다 보편적으로 사용되는 개념은 앞의 2가지 견해와는 달리 전체적인 평가체계 및 절차에 대한 평가를 의미하는 것으로, Stufflebeam은 메타평가의 범위를 좀더 넓게 파악하여 경험적 평가연구는 물론 형성평가와 비경험적 평가도 포함시키고 있으며(Cook & Gruder,

1978), 이와 유사하게 Nilsson과 Hogben(1983)은 메타평가를 평가연구에 대한 평가의 개념과 평가 자체의 기능과 실제에 대한 평가라는 개념을 포괄한 광범위한 개념으로 이해하고 있다. 또한 Larson과 Berliner(1983)는 메타평가를 평가투입, 평가과정, 평가결과 등을 포함하는 평가체계 및 절차에 대한 평가로 이해하기도 한다.

본 연구에서는 메타평가를 실시하는 주요 목적을 평가활동 그 자체를 평가함으로써 평가의 질을 향상시켜 궁극적으로 평가결과의 활용을 증진시키고자 하는 데 있다고 보고, 평가투입자원에서부터 평가활동의 내용과 과정, 평가결과 및 사후적 활용까지 포함한 개념으로 정의한다.

한편, 메타평가를 수행하기 위해서는 구성요소가 구체적으로 결정되어야 한다. 일반적으로 메타평가의 구성요소는 크게 '평가투입'(evaluation inputs)요소, '평가과정'(evaluation process)요소, '평가결과'(evaluation outcomes)요소의 3가지로 분류·논의되고 있다(Larson & Berliner, 1983; 김명수, 2000). 먼저, 평가투입요소는 평가에 동원된 여러 자원과 평가기법, 그리고 평가환경과 사업환경의 기본요소를 포함하는 것을 말한다. 둘째, 평가과정은 평가설계의 계획과 대비시켜 진행되는 실제의 평가활동을 말하는 것으로, 이 과정에서는 평가활동이 기술적으로 적절하게 수행되었는가라는 것이 초점이 된다. 마지막으로 평가결과란 평가가 미친 영향 전체를 의미하는 것으로, 평가결과에 대한 평가는 평가결과의 활용양태에 대한 전반적인 검토를 의미하는 것이 되며, 이 문제는 기존의 문헌들에서 평가의 활용이라는 제목으로 자주 다루어져 왔던 내용이다.

이와는 달리, 문영세(1996)는 한국심사평가제도에 관한 메타평가에서 메타평가 구성요소를 평가가치, 평가투입요소, 평가방식 및 과정, 평가활용으로 분류하고 있는데, 이는 위에서 논의한 3가지 분류체계에서 평가가치요소가 더 첨가된 것이라고 볼 수 있다. 또한 이찬구(1997)는 기존 연구들의 메타평가 구성요소가 수단적이고 방법론적인 사항들에 치우쳐 있다고 비판하면서 기존 연구를 토대로 하여, 평가

기조요소와 평가과정의 절차적 사항을 새로이 추가하고 있다. 즉, 메타평가의 구성요소로 평가기준, 평가자원, 평가수행, 평가활용요소의 네 가지를 설정하고, 평가수행요소는 다시 내용적 사항과 절차적 사항으로 분류하고 있다.

이상의 연구들을 종합해 보면, Larson과 Berliner의 논의는 메타평가의 구성요소에 대한 최초의 체계적 연구라는 점에서 의의가 있으나, 평가투입 및 평가과정요소가 지나치게 수단적이고 방법론적인 사항에 치우쳐 있다는 비판을 받고 있으며, 평가결과요소에서도 평가활용을 1차적 환류로만 한정짓고 있고, 메타평가의 구체적 세부평가 항목 역시 각 세부항목이 일부 중복적이거나 구성요소별 분류가 바람직하지 않은 점이 있다는 비판을 받는데, 김명수의 논의 또한 여기서 크게 자유로울 수 없다는 한계점을 가지고 있다(이찬구, 1997).

이런 측면에서 볼 때, 문영세(1996)와 이찬구(1997)의 메타평가 모형은 평가체계 전체를 보다 광범위하게 다루고 있다는 측면은 있으나, 이들이 제시하고 있는 메타평가 구성요소는 중앙심사평가제도와 정보통신개발사업의 메타평가에 관한 것으로, 사회복지시설 평가사업을 대상으로 메타평가를 수행할 때에는 그 세부 구성요소가 달라질 수밖에 없다. 사회복지시설 평가사업의 고유하고 특이한 구성요소가 존재하기 때문이다.

따라서 사회복지시설 평가사업에 대한 메타평가 구성요소는 메타평가의 일반적 구성요소를 우선적으로 구비한 다음에, 추가적으로 사회복지시설 평가의 특징을 반영할 수 있는 세부 구성요소들이 논의되어야 한다. 본 연구에서는 위 연구들에서 공통적으로 나타나는 평가자원 영역, 평가수행 영역, 평가결과 및 활용 영역으로 일반적 구성요소를 구성한 다음, 사회복지시설 평가의 특성과 내용을 중심으로 세부항목들을 반영하고자 한다.

3. 사회복지시설 평가의 실시현황 및 특성

1990년대 이후 사회복지계에서 평가는 중대한 관심사가 되고 있는데, 이러한 평가제도의 도입배경에는 1997년도 경제위기 이후 한국사회를 관통한 구조조정의 흐름이 있었다. 외환위기 이후 우리 사회전반에 시장원리에 따른 효과성과 효율성 증진에 대한 목소리가 높아졌으며, 공공부문 또한 이러한 흐름에서 예외일 수는 없었다.

이러한 환경 속에서 사회복지계는 1990년대 들어서면서 복지관을 비롯하여 재가봉사센터 등 이용시설의 지속적인 양적 확대가 있었고,²⁾ 이에 사회복지사업에 할당된 자원이 효과적, 효율적으로 사용되고 있는가에 대한 사회의 관심이 한층 더 증가하면서 책임성 검증의 압력이 제기되었다(최재성, 1998).

또한 사회복지서비스에 대한 정책의 방향이 기존의 수용 중심의 시설보호에서 지역사회 중심의 재가서비스로 변화하게 됨으로써, 복지대상자의 욕구충족과 문제해결뿐만 아니라, 기관운영에 대한 효과성과 효율성을 확보하고 정부나 지역주민에 대한 책임성을 가지고자 하는 노력이 요구되었다(변재관 외 6인, 2002). 그 외에도 일부 사회복지

표 1
사회복지시설 평가
추진경과

년도	추진경과
1998	사회복지사업법시행규칙 개정으로 3년마다 1회 이상 사회복지시설 평가 의무화
1999	“평가지표개발 및 평가실시” 연구용역-한국보건사회연구원 사회복지시설 평가위원회 구성(1999. 6.) (총괄위원회: 6종 생활시설분과, 2종 이용시설 분과)
1999	정신요양시설(59개), 장애인복지관(36개): 평가(총 95개소)
2000	아동영양시설(28개), 노인요양시설(60개), 여성생활시설(61개), 정신지체장애인시설(52개), 부랑인시설(33개), 사회복지관(285개): 평가(총 519개소)
2001	아동시설(243개), 노인양로시설(79개), 장애인시설(79개): 평가(총 446개소)

지시설에서 국고보조금 횡령, 친인척에 의한 자의적 운영, 시설 이용 자들에 대한 인권침해 문제 등 비리문제가 제기되면서 복지시설에 대한 투명성 확보의 필요성이 제기되었고, 평가제도는 이를 위한 중요한 수단으로 부각될 수 있었다.

또한 간과될 수 없는 부분은 사회복지 내부에서도 평가의 필요성과 이에 대한 이해가 자생적으로 성숙되어 왔다는 점이다. 사회복지현장에 있는 중간관리자급의 책임성에 대한 자각은 사실상 단기간에 평가제도의 도입을 가능하게 하는 중요한 요인으로 작용하였다. 특히 1998년 사회복지사업법 개정으로 명시된 제 43조의 1항과 이에 따른 시행규칙 제 27조 1항의 “보건복지부 및 시도지사는 법 제 43조의 규정에 의하여 3년마다 1회 이상 시설에 대한 평가를 실시하여야 한다”는 규정이 마련되면서 보건복지부는 1999년도부터 평가를 실시해 오고 있다.

사회복지시설의 평가는 1999년 장애인복지관 36개소와 정신요양시설 59개소에 대해 우선적으로 실시되었으며, 2000년은 아동영아시설을 포함한 6개 유형의 시설을 대상으로 전면적 평가가 수행되었다. 2001년도 사회복지시설 평가는 아동시설 243개소, 노인 79개소, 장애인 124개소를 대상으로 이루어졌다. 따라서 1차적으로 평가받은 시설은 총 1,060개 사회복지시설로, 등록된 사회복지시설의 거의 80%가 평가를 받은 셈이다. 1999년부터 매해 실시되고 있는 평가 실시현황을 구체적으로 살펴보면 다음 <표 2>와 같다.

한편, 1999년도부터 2001년까지 정부가 실시한 사회복지시설 평가의 특성을 간략하게 살펴보면 다음과 같다. 첫째, 이 기간 동안 평가를 받은 시설은 모두 11개 유형의 시설(1,060개)로, 11개 유형의 사회복지시설에 대한 평가모형은 전반적으로 기본구조의 측면에서 유사성이 높다. 사회복지시설 평가영역 및 평가지표 등은 시설유형에 따라 다소 차이가 있으나 대부분이 시설 및 환경 영역, 조직운영 및 인력관리 영역, 서비스의 질 영역, 지역사회관계 영역 등은 기본적으로 포함하고 있다. 이에 따라 현재 사회복지시설 평가사업의 평가지표

표 2
사회복지시설 평가 현황

구분 종별		총시 설수	평가완료시설			
			소계	1999년	2000년	2001년
시설·복지관 총계		1,328	1,060 (79.8%)	95 (7.2%)	519 (39.1%)	446 (33.6%)
아 동	소계	273	271 (99.3%)	-	28 (10.3%)	243 (89.0%)
	영아	26	26 (100.0%)	-	26 (100.0%)	-
	육아	210	209 (99.5%)	-	-	209 (99.5%)
	직업훈련	5	5 (100.0%)	-	-	5 (100.0%)
	보호치료	6	6 (100.0%)	-	-	6 (100.0%)
	자립지원	13	12 (92.3%)	-	-	12 (92.3%)
	일시보호	10	8 (80.0%)	-	-	8 (80.0%)
	종합시설*	2	2 (100.0%)	-	2 (100.0%)	2 (100.0%)
	아동전용수련원	1	1 (100.0%)	-	-	1 (100.0%)
노 인	소계	237	139 (58.6%)	-	60 (25.3%)	79 (33.3%)
	양로	119	79 (66.4%)	-	-	79 (66.4%)
	요양	94	60 (63.8%)	-	60 (63.8%)	-
	전문요양	24	-	-	-	-
여 성	소계	109	61 (56.0%)	-	61 (56.0%)	-
	모자보호	38	37 (97.4%)	-	37 (97.4%)	-
	모자자립	3	2 (66.7%)	-	2 (66.7%)	-
	모자일시	8	5 (62.5%)	-	5 (62.5%)	-
	미혼모	8	8 (100.0%)	-	8 (100.0%)	-
	선호보호	21	9 (42.9%)	-	9 (42.9%)	-
	성폭력보호	6	-	-	-	-
	가정폭력보호	25	-	-	-	-
장 애 인	소계	195	176 (90.3%)	-	52 (26.7%)	124 (63.6%)
	지체장애	39	32 (82.1%)	-	-	32 (82.1%)
	시각장애	11	11 (100.0%)	-	-	11 (100.0%)
	청각·언어장애	14	13 (92.9%)	-	-	13 (92.9%)
	정신지체	61	52 (85.2%)	-	52 (85.2%)	-
	중증요양	70	68 (97.1%)	-	-	68 (97.1%)
	정신요양	61	59 (96.7%)	59 (96.7%)	-	-
부 랑 인	소계	45	33 (73.3%)	-	33 (73.3%)	-
	부랑인	36	33 (91.7%)	-	33 (91.7%)	-
	부랑아	9	-	-	-	-
복 지 관	소계	408	321 (78.7%)	36 (8.8%)	285 (69.9%)	-
	사회복지관	343	285 (83.0%)	-	285 (83.0%)	-
	장애인복지관	65	36 (55.4%)	36 (55.4%)	-	-

*종합시설이란 영아·육아시설을 모두 포함하는 시설로, 2000·2001년 2회 평가받음.

구성은 시설환경, 운영관리, 서비스의 질, 지역사회 관계, 이용자 만족도 조사 등으로 이뤄진 것이 일반적이다.

둘째, 현재 우리나라에서 행해지고 있는 사회복지시설 평가는 전담 기구가 따로 있는 것이 아니라 정부가 한국보건사회연구원에 의뢰하여 실시하고 있으며, 따라서 한국보건사회연구원의 상시 평가인력 외에는 거의 모두가 외부인력으로 수행된다. 그러므로 이들의 양적 충분성과 질적 우수성은 평가에서 매우 중요한 요소가 된다. 그리고 평가 수행시, 평가팀은 전문가(교수) 1인, 관련 담당공무원 1인, 피평가시설 유형의 타 시설 실무자 1인의 3인 1팀으로 구성³⁾되었다는 특성을 가지고 있다(변재관 외 6인, 2002).

셋째, 평가수행 절차 면에서 살펴보면, 모두 자체평가 후 현장 실사평가라는 단계로 실시되고 있다. 이러한 방식은 평가대상자가 평가 과정에 참여하는 참여적 평가의 형태로, 평가과정 속에서 평가대상자 스스로 자신이 속한 시설의 문제점을 객관적으로 인식하고 개선할 수 있다는 점에서 상당한 의미가 있다. 그러나 피평가시설에서 자체평가를 실시할 때, 서류조작 등 정직성의 문제, 높은 평가점수를 받기 위한 사실왜곡 등의 가능성도 배제할 수 없다는 측면이 있다.

넷째, 평가의 결과에 대해 공개의 원칙을 강조하여 구체적으로 목표대비 결과에 대해서 명확하게 연간백서로 발행하는 미국의 경우와는 달리, 우리나라는 많은 평가결과 정보들이 제한적으로 공개되고 있다(김통원, 2001). 1999년도에 평가받은 장애인복지관 및 정신요양시설의 경우, 평가결과가 공개된 것은 상위 20% 시설에 불과하며, 2000년도에 평가를 받은 종합사회복지관과 2001년도에 평가를 받은 아동복지시설의 경우도 각각 상위 20%와 상위 25%만 공개되었을 뿐이다. 이 경우에도 평가결과가 시설들의 상대적 순위를 매기기 위한 것일 뿐, 각 평가영역별 시설의 세부실태와 문제점을 바탕으로 개선 방안을 제시하고 있지는 않다. 따라서 이러한 측면을 세부적으로 반영하여 메타평가를 할 필요가 있다고 하겠다.

4. 연구의 분석 틀 및 연구방법

1) 연구의 분석 틀

메타평가에 관한 이론과 중앙정부의 사회복지시설 평가의 내용 및 특성에 관한 논의를 종합하여 본 연구의 분석 틀인 사회복지시설 평가 사업에 적용할 수 있는 메타평가 분석 틀을 도출하면, 먼저 본 연구의 메타평가 수행시점은 1차 사회복지시설 평가가 수행된 이후에 행해지는 사후적 활동이라는 특성을 가지며, 또한 평가활동의 특정 국면이 아닌 모든 국면을 평가대상으로 한다는 특징을 가지게 된다. 이는 메타평가가 평가조직 및 투입자원에서부터 평가활동의 내용과 과정, 평가결과의 사후적 활용형태까지 정밀하게 검토하고 평가함으로써 전체적인 평가의 질을 제고하려는 의도가 담겨있음을 의미한다. 사회복지시설 평가의 메타평가 구성요소를 종합하여 본 연구의 분석 틀을 제시하면 다음 <표 3>과 같다.

2) 연구방법

본 연구의 메타평가 대상은 보건복지부가 1999년부터 2001년까지 시행한 사회복지시설 평가사업으로, 메타평가는 먼저 평가관련 문헌검토와 1999년부터 2001년까지 실행된 사회복지시설 평가 결과보고서를 중심으로 평가 구성요소별 하위영역에 따른 세부 평가지표를 개발하여 사회복지시설 평가핵심관련자(stakeholder)의 자문을 받은 후 1차 시범평가, 2차 본 평가를 거쳐 실시되었다.

1차 시범 메타평가는 1999년부터 2001년까지 3년에 걸쳐 중앙정부로부터 평가를 받은 1,060개 사회복지시설을 메타평가의 주제로 선정하여, 2002년 8월~9월 동안 2개월에 걸쳐 이루어졌다. 이때 자료수집방법은 질문지를 활용한 우편조사를 활용했으며, 회신된 사례는 443부였다. 시범평가 과정에서 사회복지시설의 중간관리자 및 실무

자들을 대상으로 추후 전화면접이나 대인면접을 통하여 보다 구체적이고 다양한 의견을 수렴·보충하였다.

시험평가 및 면담결과를 바탕으로 세부 평가지표를 수정, 보완하여

표 3
사회복지시설 메타평가
구성요소

주요 영역		세부 측정항목
평가 자원 영역	평가인력	<ul style="list-style-type: none"> 평가팀 인력구성의 적절성 평가인력의 양적 규모 평가자의 전문적 평가능력 및 기술 평가팀간의 격차 평가자들의 평가대상시설 이해도 평가자들의 공정성
	평가조직	<ul style="list-style-type: none"> 현 평가조직의 구조적 적절성
	평가도구	<ul style="list-style-type: none"> 평가항목의 시설특성 반영 정도 평가기준의 현실적합성 평가기준의 명확성 평가기준의 변별성 평가기준의 신뢰성
평가 수행 영역	평가수행방법의 신뢰성	<ul style="list-style-type: none"> 서비스 이용자나 종사자 만족도 조사의 신뢰성 자체평가의 신뢰성 현장평가의 신뢰성 사업실적 허위작성 경험 정도 서류조작 경험 정도 서비스 이용자나 자원봉사자 명단작성의 허위작성 경험 정도
	평가수행절차의 합리성	<ul style="list-style-type: none"> 자체평가 후 현장평가 수행절차의 적절성 자체평가 전 사전평가교육의 효과성 피평가시설의 평가과정 참여 정도
	평가수행기간의 적정성	<ul style="list-style-type: none"> 평가주기의 적절성 평가자의 현장평가 시기적절성 평가자의 현장방문 평가시간의 충분성 자체평가시간의 충분성
	의사소통	<ul style="list-style-type: none"> 의사소통의 접촉빈도 상호협력 정도 피평가시설의 견해반영 정도
평가 결과 영역	충실성	<ul style="list-style-type: none"> 평가결과보고서의 구체성/용이성/유용성
	신뢰성	<ul style="list-style-type: none"> 평가결과의 기대일치 정도 평가결과의 신뢰성
	확산성	<ul style="list-style-type: none"> 평가결과 정보제공의 시기적절성 평가결과의 정책반영 정도 평가결과 정보의 접근성 평가결과에 대한 이의제기 기회의 충분성
평가 활용 영역	평가활용 유형	<ul style="list-style-type: none"> 도구적 활용 개념적 활용

최종 확정된 후 2003년 3월 3일부터 4월 5일에 걸쳐 2차 본 평가를 실시하였다.⁴⁾ 2차 본 평가는 1999년부터 2001년까지 3년에 걸쳐 중앙정부로부터 평가를 받은 1,060개 사회복지시설을 메타평가의 주체로 선정하여, 질문지를 활용한 우편조사와 시설 실무자를 대상으로 한 심층면접을 통한 질적 방법을 병행하였다.

양적 방법인 서베이에서 응답자는 사회복지시설을 대표하는 중간관리자 1명으로 한정하였다. 의사결정권이 있으며, 평가 당시 평가관련 업무를 주로 담당했던 중간관리자로, 생활시설의 경우는 사무국장이나 총무를, 이용시설에서는 부장이나 과장, 혹은 팀체제일 경우 팀장으로 제한하였다. 본 평가에서 회신된 사례는 385부였으며, 최종 분석에 사용된 사례는 347부이다. 최종분석에 사용된 사례가 347부인 것은 부실한 응답 외에도 응답자의 현 시설 근무경력에 따라 적절하지 못한 사례를 제외했기 때문이다. 메타평가 대상이 1999년부터 2001년까지의 평가이므로, 그 당시에 근무하지 않았던 것으로 볼 수 있는 사례 역시 제외하였다.⁵⁾

그리고 질적 방법인 심층면접은 울산광역시 및 대구·경북지역에 있는 사회복지시설(1999년~2001년도 평가대상시설) 실무자 25명을 대상으로 2002년 12월 9일부터 2003년 3월 31일까지 실시하였으며, 비구조화된 질문지법을 이용하였다. 면접조사의 내용은 면접대상자들의 개인적 사항 및 사회복지시설 평가제도와 관련된 전반적 내용 및 평가제도 개선을 위한 제언사항 등을 중심으로 구성되었으며, 최종적으로 총 25개의 사례를 분석하였다.

5. 메타평가 결과분석

본 연구에서는 Cronbach's Alpha Test를 이용하여 측정지표들의 신뢰도를 검증하였는데, 신뢰도 계수의 크기는 평가자원, 평가수행, 평가결과 및 평가활용 영역 모두 .6 이상의 α 계수를 나타내 신뢰도에는

큰 문제가 없는 것으로 나타났다. 각 영역별로 메타평가 결과를 분석해 보면 다음과 같다.

1) 평가자원 영역

메타평가 결과, 평가자원 영역은 4개 영역 중 가장 점수가 낮은 55.6점에 불과한 것으로 나타났다. 이 중 평가를 담당하는 조직의 구조적 적절성과 평가도구에 대한 평가는 각각 53.2점과 53.0점으로, 평가인력에 대한 메타평가 결과는 만점대비 58.1점으로, 모두 평균점에도 못 미치는 열악한 것으로 나타났다.

먼저, 평가조직의 구조적 적절성 차원에서 보면, 보건복지부가 한국보건사회연구원에 용역을 주는 평가수행 구조가 적절하지 못한 것으로 나타났다. 한국보건사회연구원의 사회복지시설평가단은 독립상설기구가 아니라 연구과제를 수행하기 위해 일시적으로 운영되는 조직이며, 연구과제가 종료됨과 동시에 해체되는 특성을 가지고 있는데, 응답시설의 72.2%가 이러한 시스템이 현 사회복지시설 평가제도의 발전에 부정적이라고 응답하였으며, 심층면접에서도 이러한 사실을 분명하게 확인할 수 있었다.

둘째, 평가자원의 하위영역인 평가인력에 대한 메타평가 결과, 교수, 실무자, 공무원으로 구성된 인력의 구성이나 양적 규모 면에서는 평균점수 이상으로 나타난 반면, 평가자의 질적 측면에서는 50점대로 매우 부정적인 것으로 나타났다. 평가자들의 질적 우수성이란 개별 평가자들이 평가수행 전문능력과 자질을 갖추고 있는지의 여부와 관련되는데, 개별 평가자들의 평가수행 능력에 대한 판단기준은 평가내용과 평가방법에 대한 이해(Sanders, 1994)를 갖추고 있어야 한다는 것이다.

하지만 평가내용에 대한 평가팀간의 편차에 대한 충분한 조정이 이루어지지 못한 것으로 나타났으며(51.2점), 평가자들의 평가대상 시설에 대한 이해도도 52.7점으로 매우 낮은 것으로 나타났다. 그 외

평가자의 전문적 평가능력 및 기술은 57.5점, 평가자들의 공정성은 59.7점으로 각각 나타났으며, 전반적으로 평가인력에 대한 피평가시설의 불신은 면접결과에서도 많이 나타났다.

셋째, 평가자원의 하위영역인 평가도구에 대한 메타평가 결과, 평가기준의 현실적합성에 대한 점수가 48.5점으로, 가장 낮은 것으로 나타났다. 그 다음으로 부정적 평가를 받은 항목이 평가항목의 시설 특성 반영정도로, 50.0점이다. 평가기준이 이론 중심적이며, 각 시설들의 특성들을 분명하게 반영하고 있지 못하다는 지적은 질적 측정 방법인 심층면접에서도 많이 언급된 부분이다. 평가대상 시설에 따라 평가의 강조점이 다르고 그에 따른 구체적 평가지표도 다르게 구성되어야 함에도 불구하고 이에 대한 고려가 거의 없다는 지적이 많았다. 특히 장애인복지시설의 경우 세부 평가항목이 장애종별이나 장애 정도에 따라 잘 반영되고 있지 못한 것으로 나타났다.

그 외에도 평가기준의 신뢰성에 대한 메타평가 결과는 54.7점, 평

표 4
평가투입자원에 대한
메타평가 결과

항 목		긍정적 N (%)	부정적 N (%)	평균 /만점	만점 대비%	표준 편차
평가 조직	평가조직의 구조적 적절성	96 (27.8)	249 (72.2)	2.13/4	53.2	.679
	평가인력					
	평가팀 인력구성의 적절성	211 (60.8)	136 (39.2)	2.61/4	65.2	.605
	평가인력의 양적 규모	189 (54.6)	157 (45.4)	2.52/4	63.0	.651
	평가자의 전문적 평가능력·기술	124 (36.0)	220 (63.0)	2.30/4	57.5	.597
	평가자들의 평가대상시설 이해도	74 (22.3)	269 (77.7)	2.11/4	52.7	.621
	평가자들의 공정성	165 (47.7)	181 (52.3)	2.39/4	59.7	.664
	평가팀간의 평가결과 일치도	75 (21.7)	270 (78.3)	2.05/4	51.2	.650
	계	347 (100.0)		13.95/24	58.1	2.366
평가 도구	평가항목의 시설특성 반영정도	62 (17.9)	284 (82.1)	2.00/4	50.0	.600
	평가기준의 현실적합성	37 (10.7)	309 (89.3)	1.94/4	48.5	.572
	평가기준의 명확성	121 (35.2)	223 (64.8)	2.29/4	57.2	.626
	평가기준의 변별성	109 (31.4)	238 (68.6)	2.21/4	55.2	.669
	평가기준의 신뢰성	101 (29.3)	244 (70.7)	2.19/4	54.7	.661
	계	347 (100.0)		10.61/20	53.0	1.795

가기준의 변별성이 55.2점, 평가기준의 명확성이 57.2점으로 각각 나타났다. 이러한 결과는 평가항목의 세부지표들이 각 시설의 특수한 상황과 특성을 잘 반영하고 있지 못할 뿐 아니라 지나친 서류 중심의 평가로 인해 나타나는 현상으로 지적되었다.

2) 평가수행 영역

평가수행이란 실제로 평가를 수행하는 과정상의 활동들을 의미하는 것으로, 평가활동이 전체적으로 합리적이고 적절한 절차에 따라 수행되었는가를 검토하려는 것인데, 이에 대한 메타평가 결과를 살펴보면 다음과 같다.

첫째, 평가수행 내용 및 절차에 대한 메타평가에서는 자체평가 후 현장평가를 하는 방식에 대해서는 70.5점, 자체평가 전에 실시하는 사전 평가교육에 대해서도 70.5점으로 상당수가 긍정적인 평가를 하고 있는 것으로 나타났다. 즉, 사전 평가교육 후 자체평가, 현장평가 순서로 실시하는 것에 큰 무리가 없었음을 알 수 있다.

반면, 평가수행방법의 신뢰성과 평가수행기간의 적절성은 대체로 부정적인 것으로 나타났다. 특히 평가수행방법의 하위평가항목 중 가장 부정적 평가를 받은 항목이 서비스 이용자 및 종사자 만족도 조사의 신뢰성인데, 55.5점에 불과하여 상당히 부정적임을 알 수 있다. 또한 자체평가의 신뢰성과 현장평가의 신뢰성에 대해서도 각각 68.5점과 62.0점으로 나타나 전반적으로 부정적 평가를 하고 있는 것으로 나타났다.

실제 평가수행 과정에서 서류조작 경험에 대해 피평가시설의 52.8%가 있는 것으로 응답하였으며, 사업실적 허위작성 경험 또한 48.7%나 되는 것으로 나타나 절반 가량의 시설들이 평가점수를 잘 받기 위한 부정적인 노력을 하는 것으로 나타났다. 결국 본 연구결과에 비추어 볼 때, 현 평가제도에서 평가방법의 정직성을 어떻게 담보해낼 것인가는 매우 시급하고도 중요한 문제임을 알 수 있다. 정직하

게 평가에 임한 시설일수록 낮은 평가를 받고, 반면 서류조작이나 사업실적을 거짓 보고한 시설이 높은 평가를 받는 현실이 개선되지 못하면, 결국 평가의 본래 목적은 사라지고 평가 자체가 시설 실무자들에게는 또 다른 하나의 일거리에 불과하기 때문이다.

둘째, 평가수행의 하위영역인 평가수행기간의 적절성에 대한 메타평가 결과, 현 사회복지시설 평가제도의 평가주기인 3년에 대해서는 68.7점으로 나타났으며, 그 다음으로 자체평가 시간의 충분성에 대해

표 5
평가수행 영역에 대한
메타평가 결과

구 분		긍정적 N (%)	부정적 N (%)	평균/ 만점	만점 대비%	표준 편차
평가 수행 방법	서비스 이용자/종사자 만족 도 조사의 신뢰성	113 (32.7)	233 (67.3)	2.22/4	55.5	.670
	자체평가의 신뢰성	242 (69.7)	105 (30.3)	2.74/4	68.5	.655
	현장평가의 신뢰성	176 (50.7)	171 (49.3)	2.48/4	62.0	.590
	서류조작 경험정도*	181 (52.8)	162 (47.2)	2.45/4	60.2	.759
	사업실적 허위작성 경험정도*	167 (48.7)	176 (51.3)	2.51/4	62.7	.700
	서비스 이용자나 자원봉사자 의 허위명단작성 경험 정도*	106 (31.3)	233 (68.7)	2.76/4	69.0	.673
	계	347 (100.0)		15.15/24	63.1	2.657
평가 수행 절차	자체평가후 현장평가수행절차의 적절성	272 (78.6)	74 (21.4)	2.82/4	70.5	.574
	자체평가 전 사전평가교육의 효과성	261 (75.4)	85 (24.6)	2.82/4	70.5	.592
	피평가시설의 평가과정 참여 정도	114 (32.9)	233 (67.1)	2.27/4	56.7	.615
	계	347 (100.0)		7.91/12	65.9	1.241
평가 수행 기간	평가주기의 적절성	255 (73.5)	92 (26.5)	2.75/4	68.7	.649
	평가자의 현장평가시기 적절성	135 (39.0)	211 (61.0)	2.26/4	56.5	.723
	평가자의 현장방문평가시간의 충분성	177 (51.0)	170 (49.0)	2.45/4	61.2	.658
	자체평가시간의 충분성	223 (64.3)	124 (35.7)	2.64/4	66.0	.617
계	347 (100.0)		10.09/16	63.0	1.686	
의사 소통	의사소통의 접촉빈도	113 (32.6)	234 (67.4)	2.22/4	55.5	.642
	피평가시설의 협력 정도	335 (96.8)	11 (3.2)	3.17/4	79.2	.453
	피평가시설의 견해 반영 정도	98 (28.2)	249 (71.8)	2.22/4	55.5	.581
계	347 (100.0)		7.61/12	63.4	1.144	

* 평균에서는 역점수 처리(평균점수가 높을수록 서류조작이나 허위작성 경험 정도가 낮음을 의미).

서는 66점으로 나타났다. 반면, 가장 부정적 평가를 받은 항목은 평가자의 현장방문 시기로 56.5점에 불과한데, 이는 현장방문 시기가 평가대상인 사회복지시설의 실무자들이 휴가철인 7~8월에 주로 이루어지기 때문인 것으로 여겨진다.

또한 평가자의 현장방문 평가시간이 충분하지 않다는 평가결과가 나타났는데(61.2점), 이러한 결과는 면접에서도 많이 지적된 부분이다. 이러한 평가시간에 대한 부정적 결과는 현 사회복지시설 평가제도의 평가지표가 시설환경, 운영관리, 서비스의 질, 지역사회 관계, 이용자 만족도 조사 등 상당히 광범위하게 구성되어 있음에도 불구하고, 이를 평균 1일 정도의 단시간에 수행함으로써 평가결과를 믿을 수 있는지의 여부와 직결되는 문제를 야기시킨다.

셋째, 평가수행 과정에서 평가자와 피평가시설 간의 의사소통에 대한 메타평가 결과, 평가대상인 피평가시설의 협력 정도에 대해서는 79.2점으로 대체로 긍정적 평가로 나타났다. 이러한 결과는 평가대상인 피평가시설 측에서의 협력 정도를 평가하였기 때문으로, 향후 평가자의 입장에서 피평가시설의 협력 정도 또한 메타평가할 필요가 있다고 보여진다. 한편, 평가자와의 의사소통 빈도나 피평가시설의 견해가 반영되는 정도에 대해서는 모두 55.5점으로 나타났다. 이러한 결과는 평가수행 과정에서 피평가시설과 평가자 간의 적극적 상호교류가 상당히 미흡한 현실을 반영하는 결과로, 이에 대한 개선책이 필요함을 시사한다.

3) 평가결과 영역

평가목적을 효과적으로 달성하기 위해, 평가결과 영역에서는 평가결과보고서의 내용과 형태가 구체적인지, 그 결과는 믿을만한지, 보고서 배포시기는 적절한지, 평가결과 정보에 대한 피드백은 적절하게 이루어지는지 등을 파악할 필요가 있는데, 이에 대한 메타평가 결과를 살펴보면 다음과 같다.

첫째, 평가결과 영역의 하위영역인 평가결과의 충실성에 대한 메타평가 결과, 평가결과보고서의 구체성이 58.0점으로 가장 낮은 평가를 받았다. 이는 평가결과를 활용하고 싶어도 그 결과가 너무 포괄적이고 구체적이지 못하기 때문이다. 이 문제는 다시 평가결과의 유용성 과도 이어진다. 평가결과의 유용성에 대해서는 58.2점으로 나타났는데, 평가결과보고서가 실제 사회복지사업 개선이나 서비스의 질 향상에 별로 도움이 되지 못하였다는 것을 알 수 있다. 실제 평가결과가 구체적이지 못하여 사업이나 시설운영 개선에 전혀 도움이 되지 않는다는 시설 실무자의 면접결과는 이러한 현실을 잘 보여준다고 하겠다.

둘째, 평가결과의 하위영역인 신뢰성에 대한 메타평가 결과, 평가결과가 어느 정도 기대했던 것과 일치하는지에 대해서는 62.5점으로 나타났으며, 평가결과를 어느 정도 믿을 수 있는지에 대해서는 61.2점으로 나타나 대체로 부정적임을 알 수 있다.

셋째, 평가결과의 하위영역인 확산성에 대한 메타평가 결과, 평가결과의 정책반영 정도가 가장 부정적인 것으로 나타났고(46.5점), 그 다음으로 평가결과에 대한 이의제기 기회의 충분성이 평균 52.2점으

표 6
평가결과 영역에 대한
메타평가 결과

구 분		긍정적 N (%)	부정적 N (%)	평균/ 만점	만점 대비%	표준 편차
평가결 과의 충실성	평가결과보고서의 구체성	133 (38.7)	211 (60.8)	2.32/4	58.0	.626
	평가결과보고서의 용이성	206 (60.2)	136 (39.8)	2.59/4	64.7	.565
	평가결과보고서의 유용성	132 (38.7)	209 (61.3)	2.33/4	58.2	.594
계		347 (100.0)		7.23/12	60.2	1.405
평가결 과의 신뢰성	평가결과의 신뢰성	169 (49.6)	172 (50.4)	2.45/4	61.2	.624
	평가결과의 기대일치정도	185 (53.9)	158 (46.1)	2.50/4	62.5	.611
	계	347 (100.0)		4.95/8	61.8	1.110
평가결 과의 확산성	평가결과 정보제공의 적시성	113 (33.0)	229 (67.0)	2.24/4	56.0	.629
	평가결과의 정책반영 정도	44 (12.9)	296 (87.1)	1.86/4	46.5	.633
	평가결과정보의 접근성	105 (30.6)	238 (69.4)	2.22/4	55.5	.618
	평가결과이의제기 기회충분성	86 (25.3)	254 (74.7)	2.09/4	52.2	.645
계		347 (100.0)		8.40/16	52.5	1.752

로 낮은 평가를 받은 것으로 나타났다. 특히, 면접결과에서 실무자들이 가장 많이 지적하고 있는 영역이 평가결과의 확산성 중 평가결과의 정책반영 정도에 대한 것으로 나타났다. 이것은 평가 후 평가대상 시설에서 개선할 능력이 없는 부분에 대한 정부의 지원이 대단히 부족하다는 것을 반영하는 결과이다. 이들의 주장처럼, 정부가 평가를 실시한 이후 시설개선을 위한 정책적 노력이 수반되지 않을 때 본연의 평가목적은 성취되지 않을 것이므로 이에 대한 개선이 필요하다고 하겠다.

그리고 평가결과의 확산성 중 평가결과정보의 접근성이 55.5점으로, 응답시설의 69.4%가 평가결과정보에 대한 접근의 어려움이 있는 것으로 나타났다. 평가결과 정보제공의 적시성은 56.0점으로, 응답시설의 67%가 평가결과에 대한 정보제공이 시설에서 필요한 시기에 적절하게 제공되지 않는다고 평가하였다. 심층면접에서도 평가결과를 개별통지 받지 못한 시설이 많은 것으로 나타났는데, 이러한 평가결과의 제한적 공개는 시설개선을 위한 학습의 과정으로 평가결과가 활용될 수 없는 상황임을 명백하게 보여준다.

요컨대, 평가를 위한 평가가 아니라 진정으로 시설의 운영개선과 서비스의 질을 향상시키고자 한다면, 평가결과정보가 평가대상인 사회복지시설 당사자에게 보다 구체적이고 적시에 제공되어야 할 뿐 아니라, 정부에서도 평가결과를 반영하는 후속조치가 매우 필요하다는 것을 알 수 있다.

4) 평가활용 영역

평가결과를 어떻게 활용할 것인가 하는 문제는 특히 평가대상 시설들에게는 초미의 관심사라고 할 수 있다. 평가결과의 활용 정도에 관한 기술적 통계분석결과를 보면, 먼저 도구적 활용의 평균값은 33.65(100점 만점 대비 64.7점), 반면 개념적 활용의 경우, 평균 26.52(100점 만점 대비 73.6점)로 도구적 활용보다 상대적으로 높게 나타났는데

표 7
평가활동에 대한
메타평가(n=347)

구 분		긍정적 N(%)	부정적 N(%)	평균/ 만점	만점 대비%	표준 편차	계
도 구 적 활 용	복지사업 및 프로그램 개선에 관한 의사결정	236 (68.8)	107 (31.2)	2.71/4	67.7	.592	33.65/ 52 (64.7)
	복지사업 및 프로그램 중단여부 결정	181 (52.8)	162 (47.2)	2.51/4	62.7	.653	
	프로그램 내용 수정에 대한 결정	215 (62.9)	127 (37.1)	2.63/4	65.7	.603	
	시설홍보	127 (37.0)	216 (63.0)	2.29/4	57.2	.703	
	서비스이용자들의 요구에 부응하는 계획수립	210 (61.2)	133 (38.8)	2.61/4	65.2	.630	
	서비스 이용자들의 권익향상	229 (66.0)	114 (33.2)	2.68/4	67.0	.593	
	자원봉사자 및 후원자 관리	182 (53.2)	160 (46.8)	2.53/4	65.7	.601	
	직원교육 및 훈련기회 제공	244 (71.1)	99 (28.9)	2.76/4	69.0	.583	
	시설의 설비 개선	225 (65.6)	118 (34.4)	2.67/4	66.7	.610	
	장, 단기복지사업의 기획	217 (63.6)	124 (36.4)	2.68/4	67.0	.600	
	시설운영 및 관리체계 변화	222 (64.7)	121 (35.3)	2.65/4	66.2	.577	
	예산증가 또는 감소	133 (38.9)	209 (61.1)	2.35/4	58.7	.583	
	시설인력의 양적, 질적 개선	199 (58.2)	143 (41.8)	2.59/4	64.7	.610	
개 념 적 활 용	평가결과에 대해 잔여하게 고려	293 (85.4)	50 (14.6)	3.06/4	76.5	.611	26.52/ 36 (73.6)
	전문적 지식에 비추어 평가보고서 검토	264 (77.6)	76 (22.4)	2.85/4	71.2	.568	
	평가가 시설에 미치는 의미에 대해 심사숙고	287 (83.7)	56 (16.3)	2.99/4	74.7	.592	
	평가결과의 타당도 비판	290 (84.5)	53 (15.5)	3.01/4	75.2	.609	
	평가결과이의제기방법고려	195 (57.4)	145 (42.6)	2.62/4	65.5	.692	
	직원과 함께 평가보고에 관해 논의	291 (85.1)	51 (14.9)	2.99/4	74.7	.610	
	평가결과의 실행가능성에 대한 비판	264 (77.2)	78 (22.8)	2.88/4	72.0	.605	
	평가결과적용방법에 대해 생각	302 (88.0)	41 (12.0)	3.01/4	75.2	.535	
	평가보고서의 요약 및 발췌본 숙독	299 (88.2)	40 (11.8)	3.07/4	76.7	.656	

데, 이러한 결과는 평가활용이 직접적 행동상의 변화보다는 간접적 사고의 변화에 더 많은 효과가 있었음을 의미한다. Heisenberg(2000)의 불확실성 원리(Uncertainty Principle)에 의하면, 평가 및 측정과 관련된 이야기가 시작되면서 이미 간접적 변화가 나타난다는 것이다.

하지만 이러한 간접활용이 직접적 의사결정으로까지 그대로 이어지고 있지는 못하고 있는 것으로 나타났다. 즉, 도구적 활용은 64.7점으로 상대적으로 낮은 것으로 나타나 행동상의 변화로까지 그대로 이어지지 못하는 현실을 알 수 있다. 한편, 도구적 활용의 경우, 주로 복지사업 및 프로그램의 개선에 대한 의사결정에 평가결과를 많이 활용한 것으로 나타났으며, 예산을 증가시키거나 감소시키는 것과 관련된 의사결정, 시설홍보 등에는 상대적으로 적게 활용한 것으로 나타나 이에 대한 개선이 필요함을 알 수 있다. 한편, 평가에 대한 인식변화나 교육차원의 간접적 활용유형인 개념적 활용의 경우, 주로 평가결과보고서를 진지하게 검토하거나 평가결과 적합성을 심사숙고하는 경향이 높았다.

한편, <표 8>에 나타나듯이 평가결과의 신뢰정도⁶⁾에 따른 평가활용 정도의 차이를 살펴본 결과, 간접적 활용인 개념적 활용의 경우는 통계적으로 유의한 차이를 보이지 않았다. 반면, 직접적 활용의 형태인 도구적 활용의 경우 평가결과 신뢰 정도에 따른 집단간 차이가 있는 것으로 나타났다($p < .05$). 즉, 평가결과를 신뢰하는 정도가 높은

표 8
평가결과 신뢰 정도에
따른 평가활용 차이

집 단		사례수	평균	표준편차	t	자유도	유의확률
도구 적 활용	평가결과 신뢰 낮은 집단	128	32.78	5.9189	-2.201	330	.028
	평가결과 신뢰 높은 집단	204	34.14	5.1389			
개념 적 활용	평가결과 신뢰 낮은 집단	127	26.40	3.5731	-.406	327	.685
	평가결과 신뢰 높은 집단	202	26.56	3.2289			

시설이 평가결과에 대한 신뢰 정도가 낮은 집단에 비해 도구적 평가 활용 정도가 더 높은 것으로 나타났는데, 이러한 결과는 평가결과에 대한 신뢰 정도를 높일수록 피평가시설에서 평가결과를 직접적으로 활용할 가능성이 높아짐을 시사한다.

6. 사회복지시설 평가제도의 개선방안

1) 평가자원 영역

평가자원 영역의 메타평가 결과를 바탕으로 향후 평가제도의 발전방안을 제시해보면 다음과 같다.

첫째, 무엇보다 평가전담 기구 및 평가관련 법규정의 정비가 필요하다. 사회복지시설 평가사업의 평가활동이 원활히 수행되기 위해서는 평가관련 전담조직과 이를 뒷받침하는 평가규정의 정비가 이루어져야 한다. 사회복지시설 평가사업은 전문 연구인력뿐만 아니라 시설 평가 과정에서 방대한 행정력이 요구된다. 평가인력에 대한 업무연락 및 평가일정 조정, 민원접수 및 해결뿐만 아니라 관련 중앙부서 및 시·도와의 원활하고 긴밀한 협조를 요청하는 행정체계가 필요하다.

그러나 현재 사회복지시설 평가를 담당하고 있는 한국보건사회연구원의 사회복지시설평가단은 독립상설기구가 아니라 연구과제를 수행하기 위해 일시적으로 운영되는 조직이며, 연구과제가 종료됨과 동시에 해체되는 특성을 가진다. 이러한 특성은 평가사업의 수행에 대한 사전·사후작업이 지속적으로 이루어지지 못하도록 하며, 따라서 평가수행에 대한 전문기술의 축적을 어렵게 할 뿐 아니라 평가의 전문성을 유지할 수 없고, 평가제도의 안정화를 어렵게 한다. 현재 우리나라 사회복지시설에 대한 평가제도는 정부의 비용부담이 클 뿐만 아니라 평가대상인 시설 측의 준비부담과 평가자의 부담도 매우 크다고 볼 수 있다.

따라서 평가결과를 통해 사회복지시설이 계속적으로 발전할 수 있도록 평가결과가 활용되기 위해서는 가장 먼저 상시적으로 평가업무를 수행하는 전담기구가 반드시 필요하다. 평가전담기구가 수행해야 할 역할은 정부 내 여러 계층에서 수행되고 있는 평가활동을 총괄·조정하고, 사회복지시설들로 하여금 체계적으로 자체평가를 수행할 수 있도록 지원해주며, 또한 평가결과에 따라 사회복지시설뿐만 아니라 정부가 취하는 조치사항들에 대해 정기적으로 점검하는 것이 무엇보다 필요하다.

둘째, 사회복지시설 평가에 대한 법규정이 구체화되어 있지 못하다. 사회복지사업법에서 사회복지시설은 3년마다 평가를 받아야 한다는 규정으로 언급해놓은 것이 전부다. 평가규정은 평가자와 피평가자 모두에게 평가의 예측성과 일관성을 확보해주는 제도적 장치로 작용하게 되므로(이찬구, 1997), 평가목적, 평가주체, 평가기준, 평가방법, 평가활용 등 가능한 많은 평가관련 사항을 포함하는 것이 바람직하다. 따라서 이에 대한 구체적 법규정이 요구된다.

셋째, 평가 본연의 목적을 달성하기 위해서는 객관성과 신뢰성을 갖춘 체계가 필수적인데, 이를 위해서는 전문인력의 확보가 무엇보다도 중요하다. 평가자는 평가요소 및 절차의 효율성을 결정하는 가장 중요한 요소가 된다. 따라서 우수한 평가인력을 확보하는 것이야말로 평가결과의 질을 높이는 첩경이 될 수 있다.

하지만 본 연구결과, 평가인력의 구성이나 양적 규모 면에서는 다소 긍정적인 편이었으나 질적 우수성에 대해서는 상당히 부정적인 것으로 나타났는데, 특히 평가팀간의 평가결과 일치 정도와 평가자들의 시설 이해 정도가 매우 낮은 점수를 보였다. 각 평가팀은 평가팀간의 편차가 없도록 평가지표를 해석하는 데 있어 일치성을 가져야 하며, 이를 위해서는 평가팀간의 충분한 논의와 교육이 있어야 할 것이다.

이를 위해서는 무엇보다 평가자의 평가업무에의 전담 정도를 높이는 방법이 고려되어야 한다. 한국보건사회연구원 내부 인력의 평가업무에의 전담도는 전담기구의 설치와 직접적으로 연결되는 사항이므로

더 이상의 논의는 생략한다. 외부 인력인 시설평가팀 평가위원의 평가업무 전담도 제고방안을 위해, 우선적으로는 평가자들에게 지급되는 평가관련 수당을 좀더 현실화하고, 다음에는 평가위원들에게 실질적 혜택을 부여해야 한다.

제 1차 사회복지시설 평가사업의 경우 평가팀장에 해당하는 관련 교수의 평가수당은 1개소 당 50,000원에서 70,000원이었다. 이는 시설별 약 120개 이상의 평가지표를 검토하고, 평가소요 시간은 평균 6시간에서 10시간 정도로 볼 때(변재관 외 8인, 2001), 전문인력의 참여를 유도할 수 있는 기제가 부족한 실정이다. 뿐만 아니라 시설평가 수행 과정에서 사고가 발생할지라도 이를 보상해 줄 수 있는 방안이 마련되어 있지 않다. 따라서 현장평가를 수행한 경험이 있는 평가위원의 안정적 참여를 지속적으로 보장하기 위해서는 평가위원의 처우개선은 필수적이다. 평가에 참여한 대학교수에게는 전문가로서의 사회봉사 경력을 정부차원에서 인정하고, 공무원이나 실무현장의 사회복지사에게는 인사고과에서 우대하도록 하는 방안을 고려할 필요가 있다.

요컨대, 평가인력의 질적 우수성은 무엇보다도 한국보건사회연구원의 사회복지시설 평가단 내·외부 인력 모두의 전문지식 향상과 평가경험의 축적을 통해 제고되어야 한다. 하지만 현재와 같이 전담기구의 자체직원이 아닌 관련 연구기관에서 시설평가단의 직원을 구성하는 방법은 직원들의 소속감 결여와 업무의 불연속성으로 인해 평가경험의 축적에 결정적 저해요인으로 작용하고 있어 일정부분 한계를 가질 수밖에 없다. 따라서 근본적으로는 평가전담기구의 설립을 통해 전담직원의 양적 확보를 갖추고, 2차적으로는 외부 평가인력들의 평가경험을 축적하기 위해 평가자의 임기를 3~4년 정도로 하는 것이 바람직하다.

신뢰성과 공정성을 갖춘 평가자를 선정하는 일은 쉽지 않기 때문에 선진국에서는 평가자 후보들에 대한 정보를 DB화하여 운영하고 있다(이진주 외 6인, 1990). 우리나라도 향후 사회복지시설 평가를 보다 체계화하기 위해서는 평가자 후보들에 대한 정보를 DB화하고, 이러

한 정보를 바탕으로 엄격한 기준에 의하여 평가자를 선정, 교육할 필요가 있다.

또한 사회복지계에서 평가전문가를 체계적으로 양성하는 방안도 양질의 평가인력 확보를 위한 하나의 방법이 될 것이다. 평가팀별로 1주일에서 2주일 이상 평가활동을 하는 경우가 일반적이는데, 이로 인해 평가위원으로 위촉되는 교수, 공무원, 실무자 모두 본업에 막대한 지장을 주게 되어 평가위원 확보가 어려운 현실이다(최재성, 2000). 이런 측면에서 양질의 평가팀을 안정적으로 확보하기 위해서는 사회복지사 중 평가전문가를 체계적으로 양성하여 활용할 수 있는 방안도 신속히 강구하여야 할 것이다.

넷째, 평가도구에 대한 메타평가 결과는 전반적으로 매우 부정적임을 알 수 있는데, 이러한 평가도구의 문제점 개선을 위해서는 평가를 수행하는 정부와 평가대상인 사회복지시설 간의 지속적 상호교류가 무엇보다 중요하다. 지속적인 상호 의견교환을 통해 평가도구의 질적 개선을 도모해야 하며, 또한 현재의 상대평가 방식의 평가도구는 향후 절대평가의 체계로 전환되어야 한다. 즉, 서비스의 표준을 제시하고 이 조건을 충족하는 시설에 대해서는 인정을 해주고, 인정을 받지 못하는 시설은 개선하도록 유도할 필요가 있다. 이러한 상대평가에서 절대평가로의 전환은 중장기적으로 평가인증제도의 도입과 연관되는 과제라고 하겠다.

2) 평가수행 영역

사회복지시설 평가수행은 합리적 평가절차에 따라 공정하게, 적정 기간 내에 수행되어야 한다. 이러한 측면에서 사회복지시설 평가수행과 관련된 문제점으로 드러난 사항들을 중심으로 개선방안을 제시하면 다음과 같다.

첫째, 자체평가의 자의적이고 주관적인 측면을 방지하기 위해 자체 평가의 결과를 DB화하여 평가신용도를 엄격하게 관리할 필요가 있

다. 앞에서 살펴보았듯이, 평가수행 방법에 대한 메타평가 결과, 서류의 조작과 같은 정직성의 문제가 제기되었다. 따라서 시설의 자체 평가 결과가 외부 평가자의 평가결과와 지나치게 상치되고 이러한 현상이 반복적으로 나타나게 되면, 상세하게 재평가를 실시하여 해당 시설의 자체평가 결과를 인정하지 않을 뿐 아니라 불이익을 주는 등 정직성을 유도할 제도적 뒷받침이 필요하다.

평가대상 시설과 관련이 있는 인근의 관련 복지시설 및 시민단체로부터 의견을 수렴하여 평가결과에 반영하는 접근도 하나의 방법이 될 것이다. 미국의 경우, 사회복지분야의 대표적 평가인증기관인 COA (Council on Accreditation)는 평가과정의 중요한 일부로 20개소의 주변 시설로부터 평가대상시설에 대한 견해를 수집하여 평가에 반영하고 있는데 (Council on Accreditation, 1997), 이는 시설간의 연계를 강화하고 정직성을 상호 유도하는 기제가 될 수 있다고 보여진다.

둘째, 평가결과가 평가대상인 사회복지시설에 객관적이고 신뢰성 있는 정보로 수용·활용되기 위해서는 절차적 합리성도 확보되어야 한다. 이런 측면에서 사회복지시설 평가사업의 메타평가를 통해 나타난 평가절차상의 문제점을 개선하기 위한 방안이 필요하다. 사회복지시설 평가사업은 평가시기가 주로 여름 휴가철인 7~8월에 이루어졌는데, 이에 대한 메타평가 결과 부적절한 것으로 나타났다. 뿐만 아니라 평가자들의 현장방문 평가시간도 지나치게 짧다는 의견이 지배적이었다. 따라서 현장평가의 기간을 현재보다 늘리고, 시설이 원하는 시기를 조정하여 사전에 협의를 거쳐 평가시점을 결정할 수 있도록 해야 할 것이다.

셋째, 3년인 평가주기를 재조정할 필요가 있다. 너무 잦은 평가는 평가의 형식화를 가져올 가능성 있으며, 또한 업무량의 과중으로 인해 정작 클라이언트에 대한 서비스 제공시간이 제한될 수 있다. 따라서 현 3년인 평가제도를 평가결과에 따라 3~5년으로 조정할 필요가 있겠다. 즉, 평가결과에 따라 우수한 시설의 경우, 재평가기간을 5년 후로 늦추어주는 것도 한 방법이 될 것이다.

넷째, 전문화된 평가인력이 일정 기간을 시설업무에 직접 참여하면서 구체적, 전문적 평가를 하는 일종의 참여 중심적 평가수행을 장기적으로 고려해 볼 만하다. 그러나 단기적 측면에서 볼 때, 비용문제나 운영상의 문제 등 많은 현실적 제약이 뒤따르므로 자체평가 결과가 현장평가와 상치되는 시설에 우선적으로 적용해 볼 수 있을 것이다.

3) 평가결과 및 활용 영역

메타평가에서 나타난 평가결과 및 활용상의 문제점을 바탕으로 개선 방안을 모색해보면 다음과 같다.

첫째, 평가결과가 널리 활용되기 위해서는 무엇보다도 평가결과보고서가 구체적 정보를 포함하고 있어야 한다. 보고서에 포함되어야 할 구체적이고 필수적인 정보란 평가의 목적을 달성하는 데 기여할 수 있는 정보들로, 예를 들면, 평가결과의 활용방법이라든가, 시설개선을 위한 구체적 조언들이다(정선욱·김진숙, 2000). 그러나 본 연구 결과에 의하면, 1차 사회복지시설 평가결과는 시설유형에 따라 다소 차이가 있긴 하나 대체적으로 상, 중, 하나 상위 20%만 공개되었을 뿐, 세부적이고 구체적인 정보공개가 이루어지지 못했으며, 일부 배포된 보고서 또한 너무 포괄적이거나 추상적이어서 원재료에 불과하다는 지적이 많았다.

평가는 분명히 시설의 발전과 서비스 향상에 목적을 두고 이루어져야 할 것이며, 따라서 해당시설에게 평가결과를 구체적으로 공개하여 그 시설들이 스스로 인정하고 개선되도록 유도할 필요가 있다. 평가결과가 시설들의 순위를 매기는 데 사용되기보다는 각 시설의 세부 실태와 주요 문제점들을 밝히고 개선하는 자료로 활용되기 위해서는 무엇보다 구체적인 평가결과의 공개가 필요하다. 즉, 평가결과보고서가 시설개선 및 문제해결을 담보해내기 위해서는 시설의 현실에 대한 객관적 사실과 분명하고 타당한 원인분석, 구체적이고 실현가능한 대안제시가 포함되어야 한다.

둘째, 개별 시설별 보고서 배포가 필요하다. 연구결과, 평가보고서가 배포되지 않은 시설도 있었는데, 이것은 시설개선을 위한 학습의 과정으로 평가가 전혀 활용될 수 없는 상황임을 보여주는 결과이다. 평가가 시설을 심판하기 위한 것이 아니라 시설의 개선과 서비스 질을 향상하기 위해서라면, 개별 시설별로 평가보고서를 배포하는 것이 효과적일 것이다. 또한 개별 평가보고서를 배포하기 전에 시설로부터 의견을 수렴할 수 있는 제도적 방안들이 필요하다.

셋째, 평가결과정보의 정책반영이 필요하다. 평가결과정보가 평가대상인 사회복지시설 당사자에게 보다 구체적이고 적시에 제공되어야 할 뿐 아니라, 정부에서도 평가결과를 반영하는 후속조치가 매우 필요하다. 평가 실시 이후, 시설개선을 위한 정책적 노력이 수반되지 않을 때 평가목적의 성취는 이루어지지 않을 것이므로 이에 대한 개선이 필요한데, 이를 위해서는 무엇보다도 먼저 사회복지시설에 대한 정부의 예산지원이 현실화되어야 한다.

사회복지시설에 대한 예산지원의 비현실성은 평가제도의 본래 취지를 왜곡시키며, 뿐만 아니라 정부의 적절한 지원이 이루어지지 않은 상태에서 시설 측의 변화와 개선만 요구한다면 평가제도를 규제의 연장수단으로 받아들일 가능성마저 있다(최재성, 2000). 따라서 평가대상 시설의 차이가 이미 구조화되어 있는 상태에서 계속 시설 측에만 변화를 요구하기보다 정부의 적극적인 시설개선정책 마련이 요구되며, 이를 위해서는 근본적으로 예산지원의 현실화가 이루어질 필요가 있다.

넷째, 평가결과 및 평가관련 정보를 DB화하여 관련시설이나 개인들이 쉽게 접근, 활용할 수 있도록 평가정보시스템을 구축·발전시킬 필요가 있다. 이러한 정보시스템을 통한 체계적 평가관련 정보는 평가를 받은 사회복지시설뿐만 아니라 시설을 운영하고자 하는 개인이나 법인에게도 유용한 도움을 제공할 수 있을 것이다. 이러한 방안은 평가결과 정보에 대한 접근성을 높여 사회복지시설의 개선과 발전을 위한 학습과정을 확산시킬 수 있을 것으로, 향후 평가전담기구의 설립과 함께 보다 구체적으로 논의되어야 할 것이다.

마지막으로, 평가활용 정도를 높이기 위해서는 무엇보다도 평가결과의 신뢰 정도를 높일 필요가 있다. 특히, 평가활용 유형 중 도구적 활용의 경우 평가결과 신뢰 정도에 따른 집단간 차이가 있는 것으로 나타났는데, 이러한 결과는 평가결과에 대한 피평가시설의 신뢰 정도가 높아질수록 피평가시설에서 평가결과를 직접적으로 활용할 가능성이 높아질 수 있음을 반영하는 결과이다. 결국 이러한 현안들은 평가전체의 전반적 질과 상관 있는 것으로 평가자체의 질을 향상시키기 위한 노력이 뒷받침될 때, 궁극적으로 시설개선과 서비스의 질도 향상시킬 수 있을 것이다.

7. 결론

본 연구는 사회복지시설 평가사업을 메타평가함으로써 현 평가활동의 문제점과 개선방안을 도출하여 평가제도의 질을 향상시키고 평가활용성을 제고하는 데 그 목적이 있다. 이를 위해 보건복지부가 1999년부터 2001년까지 실시한 1차 사회복지시설 평가사업을 메타평가하였다. 메타평가를 위한 분석틀의 범주로는 평가자원, 평가수행, 평가결과, 평가활용 측면으로 구분하였다. 자료수집방법은 3년 동안 중앙정부로부터 평가를 받은 1,060개 사회복지시설을 대상으로 질문지를 이용한 우편조사와 시설실무자 25명을 대상으로 한 심층면접을 통한 질적 방법을 보충하였다.

본 연구에서는 메타평가를 바탕으로 현 평가제도의 개선 및 평가활용성을 제고하기 위한 방안으로, 전담조직의 설치와 평가인력의 업무전담도의 향상 방안, 평가수행방법과 평가결과 및 활용의 구체적 방안 등을 제시하였다. 이러한 평가제도의 개선방안은 평가제도의 질을 향상시켜, 궁극적으로 평가의 주된 목적이었던 시설운영의 개선 및 서비스의 질 향상이라는 목적달성에 기여하게 되는 실천적 함의를 지닌다.

그러나 본 연구는 연구수행방법에서 평가대상인 사회복지시설을 메

타평가의 주체로 선정하였는데, 이들이 이해관계에 따라 메타평가에 임할 수 있는 측면이 있어 공정하고 객관적 평가가 저해될 가능성이 다소 있다. 따라서 향후 평가자들을 대상으로 서베이와 심층면접을 실시함으로써 사회복지시설 평가사업의 메타평가를 좀더 다각적이고 역동적인 측면에서 보완할 필요가 있다. 향후 이러한 후속연구들은 사회복지시설의 질적 개선과 보다 고품질의 복지서비스를 제공할 수 있는 평가제도 정착에 기여할 수 있을 것이다.

■ 주

- 1) 특히 정선욱과 김진숙(2000)은 1999년도에 실시된 사회복지시설 평가 중 정신요양시설에 대한 평가사업만을 선택하여 유용성을 기준으로 메타평가를 실시하였다. 국내 사회복지제에서 메타평가에 관한 연구로서는 유일한 연구이나 하나의 시설평가에 한정되었다는 한계가 있다.
- 2) 사회복지관의 경우, 1985년도 기준으로 전국에 29개소만이 운영되었으나, 1997년 말 기준으로 329개소가 운영되어 그 수가 무려 10배 이상 증가하였으며, 여기에 종사하는 사회복지사의 수도 3,000여 명에 이를 것으로 추산된다.
- 3) 1999년에서 2001년까지 3개년 동안 실시한 평가에 투입된 평가인력은 1999년 장애인복지관 30명, 정신요양시설 18명으로 총 48명, 2000년은 총 176명, 2001년은 노인시설 31명, 장애인시설 50명, 아동시설 99명으로 총 180명이었다. 3개년 평균 연구진 총 31명, 평가인력 총 404명으로 매년 약 148명이 투입된 셈이다
- 4) 1차 시범 메타평가에서 중복되는 항목이 많다는 지적에 따라 단순화시킬 필요성이 제기되었으며, 척도의 점수분포를 볼 때 중앙집중경향의 오류가 일부 있는 것으로 나타났다. 이에 본 평가에서는 5점척도에서 중간 값을 제외한 4점척도로 구성할 필요가 있는 것으로 나타났다. 4점척도는 전혀 그렇지 않다—그렇지 않은 편이다—그런 편이다—매우 그렇다로 구성되었다. 그리고 평가활용 영역에서 평가활용 측정도구는 McCormick(1997)이 평가결과 활용의 행위유형을 구체화하여 개발한 질문지를 번안, 수정하여 시범평가에서 사용하였는데, 사회복지시설 평가의 특성보다는 프로그램 평가의 내용이 보다 많은 것으로 지적되었다. 따라서 기관평가라고 할 수 있는 사회복지시설 평가의 특성에 따라 항목변경 및 측정문항 축약이 필요한 것으로 나타났다. 이러한 시범평가 및 면담결과를 바탕으로 세부 평가지표를 수정, 보완하여 최종 확정된 후 2차 본 평가를 실시하였다.
- 5) 1999년도에 평가를 받은 시설은 응답자의 현 시설 근무경력이 4년 이하, 2000년도는 3년, 2001년도는 2년 이하의 경우를 분석에서 제외하였다.
- 6) 평가결과 신뢰 정도는 평균점수를 기준으로 평균점수보다 높은 집단과 평균점수보다 낮은 두 집단으로 분류하였다.

■ 참고 문헌

- 권기성·나운환. 1995. "장애인복지기관의 조직효과성 평가지표에 관한 연구," 한국사회복지학회, 《한국사회복지학》 통권 제27호, pp. 9~33.
- 권선진. 1996. "사회복지관의 조직효과성: 평가와 과제," 연세대학교 사회복지연구소, 《연세사회복지연구》 제3권, pp. 306~329.
- 김명수. 2000. 《공공정책평가론》, 서울: 박영사.
- 김명수. 2003. "중앙행정기관 평가제도의 운영에 대한 비판적 검토," (한국정책분석평가학회 2003년 춘계학술대회 자료집), pp. 1~16.
- 김영모 외 4인. 2000. "사회복지관 평가지표," 한국복지정책연구소.
- 김통원. 1997a. "사회복지시설의 표준 평가지표 개발에 관한 연구," (한국사회복지학회 1997년 춘계학술대회 자료집), pp. 127~163.
- _____. 1997b. "사회복지시설 평가 및 운영지원에 관한 문제점 및 대안: 서울특별시를 중심으로," 한국사회복지관협회·한국사회복지프로그램 연구회, 《사회복지관 평가의 문제점과 개선방안》, pp. 5~28.
- _____. 2001. "한국과 미국의 사회복지 평가체계의 특성 비교," 한국사회복지학회, 《한국사회복지학》 통권 45호, pp. 126~155.
- 문영세. 1996. "한국중앙심사평가제도의 메타평가적 분석," (한국행정학회 1996년도 동계학술대회 논문집 I), pp. 1~37.
- _____. 1997. "한국중앙심사평가제도의 위상 재정립을 위한 메타평가: 설문조사를 중심으로," 대구·경북행정학회, 《한국행정논집》 9권 2호, pp. 231~255.
- 변재관 외 10인. 2000. 《사회복지시설 평가지표 개발 및 평가결과》, 한국보건사회연구원·보건복지부.
- _____. 외 8인. 2001. 《2000년 사회복지시설 평가(I)(II)》, 한국보건사회연구원·보건복지부.
- _____. 외 6인. 2002. 《2001년 사회복지시설 평가-양로, 장애인, 아동복지시설-》, 한국보건사회연구원·보건복지부.
- 오무근. 1986. "메타평가에 관한 연구," 명지실업전문대학 교육문제연구소, 《교육문제연구》, pp. 151~167.
- 이번송·김진욱. 1996. "서울시 사회복지관 단순평가모형," 한국사회복지학회, 《한국사회복지학》 통권 29호, pp. 105~131.
- 이찬구. 1997. "연구개발사업의 메타평가에 관한 연구-정보통신 분야를 중심으로-", (충남대학교 박사학위논문).
- 정선옥·김진숙. 2000. "정신요양시설 평가에 대한 메타평가적 접근-유용성 기준을 중심으로-", 《상황과 복지》 제8호, pp. 131~156.
- 조운희. 1998. "사회복지관 운영관리의 평가모형 개발에 관한 연구," (대구대학교 박사학위논문).
- 최일섭 외 3인. 1993. 《한국장애인종합복지관 조직효과성 평가연구》, 한국장애인재활협회.
- 최일섭 외 4인. 1998. 《사회복지관 평가모형 개발연구》, 한국사회복지관협회.
- 최재성. 1998. "사회복지관 평가모델개발의 방향과 과제: 장애인복지관 운영평가프로그램 개발사례를

- 중심으로”, 연세대학교 사회복지연구소, 《연세사회복지연구》 Vol. 5, pp. 265~290.
- _____. 2000. “사회복지평가의 진단과 정책과제”, (한국사회복지행정학회 2000년도 추계학술대회 및 Workshop 자료집), pp. 3~28.
- 황성철. 2000. “종합복지관의 평가”, 한국사회복지협의회, 《사회복지》 통권 제144호, pp. 89~101.
- 황성철·강혜규. 1994. 《사회복지관 운영평가 및 모형개발》, 한국보건사회연구원.
- Alkin, M. C. (ed.). 1990. *Debates on Evaluation*, CA: Sage.
- Chelimsky, E. & Shadish, W. R. (eds.). 1997. *Evaluation for the 21st Century: A Handbook*, CA: Sage.
- Chelimsky, E. 1985. *Program Evaluation: Patterns and Directions*, Washington D. C.: American Society for Public Administration.
- Cook, T. D. & Gruder, C. 1978. “Metaevaluation,” *Evaluation Quarterly* 2(1), pp. 5~51.
- _____. 1979. “Metaevaluation research,” *Evaluation Studies Review Annual* 4, pp. 469~515.
- Cook, T. D., Judith, L. R. & Pollard, W. E. 1981. “The Misutilization of Evaluation Research Some Pitfalls of Definition,” *Evaluation Studies Review Annual* 6, pp. 727~748.
- _____. 1995. “Assessing the Framework Programmes: A Metaevaluation,” *Evaluation* 1(2), pp. 171~188.
- Council on Accreditation. 1997. *Council on Accreditation 1997 Self-Study Manual for Behavioral Health Care Services and Community Support and Education Services*, N.Y: NY. Council on Accreditation.
- Heisenberg, W. K. 2000. “Heisenberg Uncertainty Principle,” on the World Wide Web: physics.hallm.ac.kr.
- Larson, R. & Berliner, L. 1983. “On Evaluating Evaluations,” *Policy Sciences* 16(2), pp. 147~163.
- McCormick, E. R. 1997. “Factors Influencing the Use of Evaluation Results,” *Doctoral dissertation*, The University of Minnesota.
- Nilsson, N. & Hogben, D. 1983. “Metaevaluation,” in House, E. R. (ed.). *Philosophy of Evaluation*, San Francisco: Jossey-Bass, pp. 83~97.
- Rossi, P. H., Freeman, H. E. & Lipsey, M. W. 1999. *Evaluation: A Systemic Approach*, CA: Sage.
- Scriven, M. 1969. “An Introduction to Metaevaluation,” *Educational Product Report* 2, pp. 36~38.
- Stevenson, J. F., Longabaugh, R. H. & McNeill, D. N. 1979. “Metaevaluation in the Human Services,” in Schulberg, H. C. & Jerrell, J. M. (eds.). *The Evaluator and Management*, CA: Sage, pp. 37~54.
- Stufflebeam, D. 1981. “Metaevaluation: Concept, Standard and Uses,” in Berk, R. A. (ed.). *Educational Evaluation Methodology*, The Johns Hopkins University Press.
- Worthen, B. R., Sanders, J. R. & Fitzpatrick, J. L. 1997. *Program Evaluation-Alternative Approaches and Practical Guidelines-*, Longman.

Abstract

Evaluating the Evaluation of Social Welfare Agencies of Central Government: A Meta-evaluation

Sung-Ok Lim

The purpose of this study was to investigate the existing evaluation system of social welfare agencies using a meta-evaluation method, which will eventually contribute to improve the quality and utility of the evaluation system. In order to accomplish the study purpose, this study meta-evaluated the social welfare agency evaluation project that was originally conducted by the Ministry of Health and Welfare (MHW) from 1999 to 2001. The meta-evaluation method employed in this study comprehensively evaluated the existing evaluation system by dividing the whole evaluation processes into four broad domains including 1) the evaluation resources, 2) the evaluation performance, 3) the evaluation results, and 4) the evaluation utilization. This study employed both quantitative and qualitative data collection methods.

The main findings of the study results were summarized as follows. First, the grand total score of meta-evaluation reported by agency staff concerning the current social welfare agency evaluation project was very low. Out of the highest possible score of 100, they only gave 62.7, indicating that the social welfare agencies as main subjects of the evaluation had very negative perceptions about the current evaluation system.

Second, the evaluation resources domain (55.6) had the lowest

score among the four domains. Third, the evaluation results domain also showed a low score of 57.1 out of 100. In particular, the concreteness of the final report and the dissemination of the evaluation results showed relatively low scores.

Fourth, the evaluation performance domain had relatively higher score compared to other domains indicating that subjects had more positive attitudes toward evaluation performance domain. However, the analysis also revealed that the duration of the evaluation process and the dishonesty issue in the process of each agency's self-evaluation should be modified immediately. Fifth, the conceptual utilization which is a type of indirect utilization showed higher scores than the instrumental utilization, a type of direct utilization. And, it was found that there were group differences based on the credibility level of evaluation results in the area of the instrumental utilization within the utilization patterns. These findings suggest that the credibility level of evaluation results is positively related to the degree of utilization of the evaluation results.

Based on the above results of the meta-evaluation, this study proposed the implications to improve the current evaluation system and to increase the utilization of evaluation results.

In spite of a great importance, little attention has been paid to meta-evaluation. This study contributed to the meta-evaluation research area in that a comprehensive and practical meta-evaluation methods are employed as the framework of analysis. Furthermore, the findings of this study provide crucial implications for the improvement of the current evaluation system including in the domains of resource, performance, results, and utilization of evaluation. Consequently, it is expected that this study can be the

first step to improve evaluation quality and enhancement of result utilization.

Key Words : meta-evaluation, evaluation utilization, quality of the evaluation system