

한국과 미국의 사회복지 평가체계의 특성비교

김통원*

지난 2년 동안 우리 정부는 대부분 영역의 사회복지기관들에 대해서 전면적인 평가를 하였다. 본격적인 '평가시대'는 도래하였지만 보건복지부에서 시행중인 인증제(Accreditation) 평가방식에 많은 문제점들이 일어나고 있다. 특히 운영 및 관리상의 부실이 대두되고 있으며 그 중에서도 심사인력과 관련한 객관성 및 공정성에 대한 시비는 계속 일어나리라 예상된다. 이외에도 인증제는 정부에서 수많은 기관들에 대한 평가비용을 모두 부담하기에 엄청난 비용이 든다. 반면에 미국을 위시한 OECD 국가들의 평가제는 우리와는 상당한 차이점을 보이고 있다. 이들 국가들은 간단 명료한 수행평가(Performance Measurement)를 사용하여 객관성 그리고 공정성을 확보하려고 한다. 비록 정보의 양이나 질에서 인증제보다 제한적이지만, 정부가 주체가 되어 평가하는 데 있어 우리보다 공정하고 객관성을 확보하고 있다. 따라서 본 연구는 사회복지관련 부문에서 평가를 진행해온 미국모델과 우리의 평가체계를 비교 연구하였다. 구체적으로 보건복지부가 시행하고 있는 사회복지기관 평가의 현황과 문제점, 그리고 미국의 주요 평가체계 내용들과 제한점을 제시하면서 두 나라의 평가체계들의 장단점들을 비교하였다. 끝으로 우리의 평가체계의 전략을 단기적 관점과 장기적 관점으로 나누어서 제시하였다.

* 성균관대학교 사회복지학과 교수.

1. 서론

한국 사회복지현장에 평가체도가 실제로 도입된 지 벌써 5년이라는 세월이 지났다. 1996년 서울시가 종합사회복지관을 중심으로 평가를 처음 시작하였으며, 보건복지부에서도 1999년부터 장애인복지관과 정신요양시설을 대상으로 전국적인 규모의 평가를 하였다. 다음 해인 2000년도에는 우리나라 대부분 영역의 사회복지기관들에 대해서 전면적인 평가를 하였다. 지난 2년 동안 총 1,267개 사회복지기관 중에서 절반정도인 617개 기관들이 평가 경험을 가졌다. 이러한 본격적인 '평가시대'의 도래는 평가를 위한 사회복지사업법의 법적 근거에 기반하였으며, 특히 평가의 주체인 정부의 입장에서는 공공자금에 대한 책무성(*accountability*) 확보를 위한 노력이었다. 한편 입소자 혹은 이용자 입장에서는 적절한 양질의 서비스를 제대로 받고 있는지 확인하는 방편이 되기도 한다(Royse et al., 2001).

그동안 일부 사회복지기관에 입소자 인권유린행위, 개방성 부족, 투명성 부족, 효율성 부족, 이용자 불만족 등 여러 가지 문제들이 존재하고 있었던 것은 사실이었다. 그래서 기존의 사회복지기관 관리형태인 행정지도와 회계감사형태로는 이러한 문제들을 해결할 수가 없으므로 근본적인 변화가 요청되었다. 다시 말해서 사회복지기관들의 개방성, 투명성 그리고 효율성을 보다 더 확보하기 위한 새로운 관리형태로서 사회복지기관 '평가체계'¹⁾가 대두하였다고 할 수 있다. 현존하는 사회복지관련 문제들의 해결과 책무성 확보는 평가시스템 도입에 의해서 해결할 수 있다고 기대한다.

최근 공공부문에서 신관리주의(*New Managerialism*)라는 새로운 행정관리의 형태로써 주도 성과를 중시하고 그 성과를 평가함으로써 공공부문의 경쟁력이나 책무성을 강화하려고 한다(Clarke et al., 2000). 평가는 비단 사회복지분야뿐만이 아니라 공공부문의 모든 분야에서 강조되고 있는 일종의 관리도구라고도 할 수 있다. 실제로 우리 사회에서 염원하고 있는 모든 분야에서의 개혁은 온전한 평가체계를 마련함으로써

1) 본 연구에서 평가를 '평가체계'라고 부른다. 이것은 평가를 단순히 평가지표만을 의미하는 것이 아니고 평가의 목적, 평가의 기준, 평가지표, 평가수행 절차 그리고 결과의 활용 등을 포함하는 포괄적인 성격을 가진 것으로 규정하기 때문이다. 평가체계는 구성요소들이 서로 논리적으로 일관성을 가져야 하며, 체계적으로 연계되어야 한다는 것을 의미한다.

가능한지도 모른다. Ernest와 Kenneth(1999)는 평가를 통해서 사회정의와 민주주의 까지 가져올 수 있다고 주장한다. 그만큼 평가를 우리 사회 모든 분야에서 필수적인 사회적 장치로 보고있다. 그러나, 서울시가 1996년부터 시행한 사회복지관평가는 많은 문제들을 노출하여 왔으며(김통원, 1997a), 전국적인 규모로 본격적으로 시행하고 있는 보건복지부의 평가체계에 대해서도 유사한 문제들이 보이고 있어, 평가자체의 효용성에 대해서 의문을 가지게 되는 경우도 있다. 정부에서 평가체도를 처음 도입할 때 의도했던 평가의 목적과 기대효과가 어느 정도로 성과를 거두었으며, 구체적인 피드백이 무엇이며 과연 현재 방식대로 계속해서 진행해야 하는지 검증이 필요하다. 더욱이 최근 정부에서조차도 이러한 평가와 관련한 미비점들을 평가지표를 만들 때에 참여한 학자들에게 책임을 돌리는 경우도 있다고 하니, 전면적인 정부평가 후 그 결과를 중심으로 심도 있게 연구할 필요가 있다. 하지만 현재 정부평가에 대한 관련 연구는 평가를 직접 수행한 기관인 보건사회연구원에서 나온 보고서밖에 없다.

따라서 본 연구는 1960년대부터 사회복지관련 부문에서 평가를 진행해온 미국모델과 우리의 평가체계의 특성들을 비교 연구함으로써 기본적 차이점을 비롯하여 많은 시사점들을 제공할 수 있다고 본다. 특히 1990년대부터 대두하기 시작한 수행평가(Performance Measurement)²⁾를 중심으로 미국의 평가체계와 우리의 평가체계를 비교하는 것은 최신 평가관련 지식을 국내 사회복지계에 접목할 수 있는 기회가 되며, 국가정책적 측면에서도 기존의 무리한 행정중심의 평가제 운영에 실증적 제동의 근거로 활용할 수 있을 것이다.

본 연구목적은 먼저 보건복지부의 현행 사회복지기관 평가의 현황과 그 문제점들을 살펴본다. 동시에 SEA(Service Efforts and Accomplishments) 보고서를 위시한 미국평가체계와 관련한 주요 평가내용들과 제한점을 제시한다. 마지막으로 미국평가체계와

2) 본 연구에서는 Performance Measurement를 '수행평가'라고 번역한다. 한국의 행정학분야에서는 이를 '성과평가'(outcomes evaluation)라고 부르며, 교육학에서는 본 연구처럼 수행평가라 한다. 수행평가는 성과평가처럼 결과만을 측정하는 것이 아니라 현재 수행중인 프로그램이나 조직의 활동까지 평가하는 것을 의미한다. 기획예산처에서 나온 성과주의예산제도에 관한 1999년도 자료에 의하면 performance를 성과로 번역하였으나, 2000년도 자료에는 outcomes를 성과로 번역하고 있다. 또한 미국 "Government Performance and Results Act"에서 수행(performance)과 결과(results)로 구분하며 performance를 결과적인 의미보다도 수행적인 의미(on-going)로 나타내고 있다. 따라서 performance measurement를 수행평가라고 번역하는 것이 타당하다고 본다.

우리 평가체계의 특성들을 비교한 뒤, 우리나라 사회복지 평가체계를 전망해 보고자 한다.

2. 연구내용 및 국내외 연구동향

1) 연구내용

(1) 연구범위

본 연구대상인 한국 사회복지평가체계의 범위를 보건복지부에서 시행한 제1차 평가(1999년도)와 2차 평가(2000년도)로 국한시켰다. 본 연구가 진행되는 동안에도 2000년도 보고서가 공식적으로 발행되지 않아서 일부 정보가 부족한 가운데 연구가 진행되었음도 밝혀둔다. 또한 국내외 연구동향에서 외국대상을 미국으로 국한시키고 있으며, 비교하는 미국의 평가체계 또한 대부분 '수행평가'로 국한시켰다. 대부분의 선진국들이 공공부문에서의 평가체계를 미국과 같이 수행평가로 하므로 선진국들의 평가체계 내용들이 미국과 유사하다고 보아도 무방하다(OECD, 1998). 한국과 미국의 평가체계를 비교할 때, 평가의 주체가 정부가 수행하는 '외부평가'라는 관점에 초점을 두고 연구하였다.

(2) 연구방법

본 연구의 수행과정은 다음과 같이 크게 3단계로 구분된다.

① 제1단계 : 국내외 문헌연구 및 이차적 자료분석

주로 문헌연구 중심으로 기존의 평가관련 연구문헌들을 종합하고 분석하였다. 특히 보건사회연구원(2000)에서 발간한 "사회복지시설 평가지표개발 및 평가결과" 보고서를 가지고 이차적 자료분석(*secondary data analysis*)을 함으로써 한국의 평가체계의 현황과 문제점을 파악하려고 하였다.

② 제2단계 : 평가에 참여한 핵심 정보제공자와 인터뷰

보건복지부 및 서울시 평가에 실제로 평가자 혹은 피평가자로 참여한 관련전문가들과의 면담 및 토론들을 통해서 심층적인 내용들을 수집하였다. 이들을 방문하거나 전화로 개방형 설문을 사용하여 인터뷰를 하였다. 일종의 인상주의적 접근법 (*impressionistic approach*) 을 시도하였다(Royse et al., 2001). 외부 평가자 6명 그리고 피평가 경험을 한 6명으로 총 12명을 조사하였다. 비록 이러한 접근이 주관적이고 정보가 제한적일 수 있으나, 보건사회연구원의 평가보고서 정보가 평가상의 문제점을 공개하는 데 있어 너무 제한적이므로 보완적 측면에서 채택하였다.

③ 제3단계 : 평가체계 비교연구

본 연구는 사회복지관련 미국 평가체계들을 비교 대상으로 하여 한국의 평가체계와 비교하였으며 비교·분석틀은 PUMA(Public Management, 1999)에서 사용한 OECD 국가들의 평가체계를 비교한 것을 사용하였다. 총 6가지 영역에서 비교했으나 본 연구에서는 결과 중심의 관리를 빼고 준비성을 첨가하여 6개 영역에서 비교하였다.

2) 국내외 연구동향

미국 사회복지기관은 1960년대부터 책무성을 확보하기 위한 수단으로서 가능한 한 '평가적 요소'를 포함하도록 요청함으로써 평가가 시작되었다(Brawley & Brawley, 1988). 1980년대까지 약 30년 동안 평가는 사회복지기관에서 전체 운영사업 중에 하나로 정례화 되었으며 도입된 평가방법들은 대부분이 결과지향적 혹은 목표중심의 평가방법이었다. 그들 중에 대표적인 MBO (*Management by Objectives*)는 결과지향적인 패러다임으로 목표(*objectives*)를 측정할 수 있는 성과들(*outcomes*)로 규정되었으며, 사회복지의 책무성은 이러한 목표들을 달성하는 것을 중심으로 진행되어 왔다(Gunther & Hawkins, 1996). 이러한 흐름이 1990년대에 들어와서는 보다 더 포괄적인 성격을 가진 수행평가(*Performance Measurement*)라는 개념으로 확장된다(Martin & Ketter, 1996). 수행평가란 사회복지실천프로그램의 효율성, 품질성, 효과성에 관한 정보를 얻기 위해서 정규적으로 수집 및 보고하는 것을 의미한다(Rossi & Freeman, 1993). Martin & Ketter(1996)는 사회복지 프로그램에 관한 수행평가를 체계모델

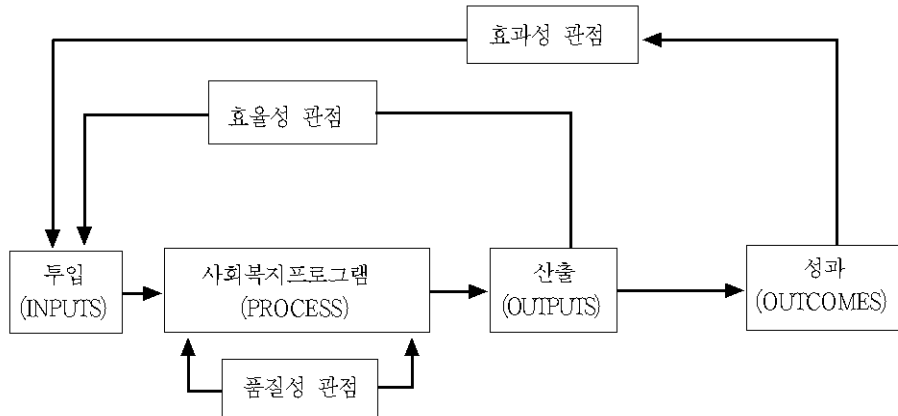
(투입, 과정, 산출, 성과, 환류)에 적용하여 설명하였다(〈표 1〉 참조). 수행평가란 다음의 3가지 측면을 통합한 것이라고 할 수가 있다.

1) 효율성 관점 : 투입과 산출의 비교를 통해서 얻어지는 생산성을 평가하려는 방법이다. 이 접근은 그동안 사회복지 실천에서 제대로 환영받지 못했었다(Pruger & Miller, 1991). 그렇지만 효율성에 대한 강조는 계속될 것이다. 2) 품질성 관점³⁾: 기업의 품질경영(Total Quality Management)과 연유한 품질성 관점은 일부 학자들(Cassam & Gupta, 1992; Dickens, 1994; Martin, 1993; Brinkerhoff & Dressler, 1990)에 의해서 주장되고 연구되어 왔다. 주로 프로그램 활동과정이나 조직내의 활동들을 평가한다. 하지만 이러한 접근은 사회복지분야에서는 아직도 낯설다. 3) 효과성 관점 : 성과를 강조하는 접근으로 사회복지관련 프로그램에서 비교적 그동안 선호되어 왔던 관점이다. 투입과 성과를 비교하는 효과성을 강조하는 접근으로 GPRA(The Government Performance and Results Act, 1993), *National Performance Review*, *Managed Care* 등에 의해서 사회복지분야에서 많이 사용하는 평가접근 방법이다. 동시에 많은 연구의 관심과 문헌들이 존재한다(Gore, 1993; Osborn & Gaebler, 1992; Cornelius, 1994; Austin, Blum, & Murtaza, 1995).

특히 미국 공공분야 중에서도 사회복지분야에서의 수행평가체계는 정부회계표준위원회(GASB, Governmental Accounting Standards Board)의 SEA 보고서로 이루어진다. Epstein(1992), Fountain & Robb(1994) 그리고 Martin & Ketter(1996) 등은 GASB의 SEA 보고서가 사회복지실천분야를 포함한 모든 공공부문에서 가장 강력한 평가시스템이 될 것이라고 주장하고 있다. 이러한 수행평가중심의 접근은 미국만이 아니라 OECD 국가들 모두가 참여를 하고 있다(OECD, 1998). 영국, 유럽 국가들, 호주, 뉴질랜드, 캐나다 등 거의 모든 선진국가들에서 사회복지분야를 포함한 모든 공공부문에서의 개혁과정은 수행평가를 통해서 이루어지고 있다고 해도 과언이 아니다.

3) 본 연구에서 품질성 관점이란 과정(process) 중심의 품질보증(Quality Assurance) 측면에서 넓게 해석한다. 비록 Martin과 Kettner(1996)가 주장하는 품질(quality)은 주로 서비스의 질로 해석되는 quality outputs로 좁게 보고 있지만, 본 연구에서는 평가의 양대 산맥중 하나인 과정중심의 접근인 품질보증과 동일하게 취급한다. 미국 공공부문에서의 수행평가는 실제로는 품질성 관점을 생략하고 효율성과 효과성 관점에 집중하고 있다. 이것은 과정(process) 중심의 품질성을 측정하는 것이 복잡하고 계량화하기 어렵기 때문이다.

〈표 1〉 체계이론에 입각한 평가모델



자료: Martin, Lawrence L. & Kettner, Peter M. 1996. "Measuring the Performance of Human Service Programs". Sage. p.8. 을 일부 수정함.

한편 국내연구동향을 살펴보면, 지금까지 사회복지시설 평가와 관련되어 국내에서 발표된 논문들은 1994년부터 발표되기 시작하여, 특히 1998년도에 본격적으로 많은 연구결과들이 나왔다. 이것은 1999년부터 사회복지시설평가가 법률적인 뒷받침에 의해서 실제로 시행되기 직전에 연구자들이 다양한 평가안들을 제시하였기 때문이라고 추정된다. 그동안 발표된 주요 선행연구들을 살펴보면 다음과 같다. 서울시 사회복지관 단순평가모형개발(이번송·김진옥·함건식, 1995)은 측정지표들에서 지나친 계량화를 시도하고 있다. 특히 일부 경제학 배경의 연구자를 중심으로 쉽게 수량화할 수 있는 실적 지표들을 중심으로 하고 있는 한계점을 보이고 있다. 아동복지수용시설 운영평가 및 개선방안(김응석 외 5인, 1995)과 사회복지관 운영평가 및 모형개발(황성철·강혜규, 1994)은 비록 시설별 특성을 반영한 평가체계를 제시하였지만 평가시행 방법이나 절차를 구체적으로 제시하지 못하고 있다. 사회복지관 평가의 문제점과 개선방안(김통원, 1997a)은 서울시에서 처음 시행한 평가에 대한 문제점들을 주로 측정상의 오류들을 중심으로 문제점들을 제시하였다. 또한 사회복지시설의 표준 평가지표 개발에 관한 연구(김통원, 1997b)는 이러한 서울시 평가의 문제점들을 사회복지학회에서 처음으로 공론화하였다. 사회복지 수용시설 평가제의 도입과 대비전략: Self-Assessment 결과를 중심으로(김통원, 1998)는 영국 사회복지사협회에서 개발한

평가틀을 가지고 수용시설 종사자들을 대상으로 조사한 것으로, 실증적인 자료를 통해서 평가제의 도입에 대한 현장의 준비도를 살펴보려고 노력하였다. 무료노인입소 시설 종사자의 소속기관평가(김통원·김혜란, 1998) 역시 실증적 자료를 통하여 준비기간이 필요함으로 평가의 점진적인 실시를 주장하고 있다. 사회복지관 평가모형개발연구(최일섭 외 4인, 1998), 서울시 사회복지시설 운영실태 조사 및 사회복지시설 평가지표 개발(최성재, 1998), 그리고 장애인복지관 평가프로그램개발(최재성, 1998)은 실제적으로 서울시 평가와 보건복지부 평가에 직·간접적으로 활용된 평가개발 연구들이다. 비록 일부 평가항목들이 아직 정교화되지 못하고 평가절차 등을 구체적으로 제시하지 못하고 있지만 평가체계의 모습을 나름대로 보여주고 있다. 한편 사회복지시설평가의 길잡이(이정호·김성이·김통원, 1999)는 평가의 주체가 정부가 아닌 민간부문의 역할을 강조한 평가틀을 제시하고 있어 다른 특색을 나타내고 있다. 이러한 모형들의 제시와는 달리 직접적으로 국가기관에 의해서 실행된 평가로는 바로 보건복지부의 위탁으로 보건사회연구원에서 개발한 평가지표와 평가결과보고서(변재관 외 10인, 2000)가 있다. 한편 사회복지서비스의 질적 향상을 위한 관리기법으로서 복지 QC 활동에 관한 사례연구(박경일, 2000)는 앞선 연구들이 대부분 기관평가의 모형들을 제시한 반면에 품질보증의 구체적인 관리기법을 소개한 실증적 사례로 좋은 연구라고 볼 수 있다. 이상의 대부분 선행연구들의 가장 큰 공통점은 '인증제' 성격을 띄고 있다는 점이다. 비록 일부 선진국에서 국가가 운영하는 인증기구도 있지만, 대개의 경우 인증제는 실제로 정부와 같은 국가기관보다는 민간차원에서 스스로 관리되는 경우가 훨씬 더 많다. 또한 이 접근방법은 국가기관에서 전면적으로 사용하기에는 너무 비용이 많이 들고 객관성이나 공정성을 최대한 확보하기에는 많은 어려움이 있다는 점이다. 비용이나 객관성문제로 미국을 비롯한 OECD 국가들이 채용하지 않는 점을 주목할 필요가 있다. 우리와는 달리 이들 국가들은 단순하면서도 객관성을 확보할 수 있는 '수행평가'(Performance Measurement)를 선택하고 있다. 수행평가에 대한 연구는 국내 사회복지계에서는 거의 전무한 실정이다. 최근 김통원(2000)이 수행평가를 언급하였지만 아직 개념적인 수준을 소개하고 있는 정도이다. 결국 사회복지관련 평가에서의 국내연구들이 인증제 중심으로 편중되어 있으며 외국의 연구동향 및 사회복지현장에서의 평가제와는 상당한 차이점을 보이고 있다.

현재 보건복지부에서 시행중인 평가방식이 인증제 성격(보건복지부, 1999)을 뚜렷

이 보이고 있기 때문에 특히 운영 및 관리상의 문제들이 대두되며 그 중에서도 심사 인력과 관련한 객관성 및 공정성에 대한 시비는 계속 일어나리라 예상된다. 이외에도 인증제는 정부에서 수많은 기관들에 대한 평가비용을 모두 부담하기에는 엄청난 비용과 노력이 든다. 왜냐하면 종합적인 조직의 품질성을 측정해야 하므로 평가항목들이 많아서 제대로 평가하려면 최소한 2~3일간 기관실사를 해야 하며 심사인력들의 교육을 철저히 시켜야 하기 때문이다. 이러한 비용과 운영 및 관리부담 때문에 보건복지부와 서울시의 평가가 잘못하면 졸속으로 행해지게 된다. 그 때는 그야말로 '형식적인 평가' 혹은 '평가를 위한 평가'로 전락할 위험성이 있다.

평가주체가 정부이고 그 평가비용을 정부에서 부담을 한다면, 수많은 사회복지기관을 평가해야 하므로 비교적 단순 명료하고 간편한 평가체계를 개발하여야 한다(DTFMS, 1993). 하지만 우리의 평가체계는 평균 평가문항수가 120개가 넘는 방대한 평가체계인 인증(심사)제도를 채택하고 있다. 인증심사제는 외부 민간평가기구가 평가주체가 되며, 인증을 원하는 사회복지기관들이 모든 비용을 자신들이 부담하면서 평가받는 것을 의미한다. 우리는 평가주체가 정부이면서도 수많은 피평가기관의 막대한 비용까지 부담하는 방식은 선택상의 오류를 범하는 것이라고 볼 수 있다.

3. 사회복지시설평가 현황과 문제점

1) 평가 실시현황

현재 진행되고 있는 우리나라의 사회복지 시설평가의 현황은 중앙부처인 보건복지부에서 전국규모로 하고 있는 평가제와 서울특별시 자치단체에서 진행하고 있는 평가제로 크게 나눌 수 있다. 본 연구에서는 전국적 규모인 보건복지부 평가만을 연구대상으로 하였다. 보건복지부에서 실시한 1999년 평가는 총 사회복지시설 1,267개 시설 중에서 장애인 복지관 36개 기관, 정신요양시설 59개 시설 등 총 95개(총시설대비 7%) 시설에 대해서 1999년 11월에 평가를 실시하였다. 다음 해인 2000년도 평가는 아동시설, 노인시설, 여성시설, 장애인시설, 부랑인시설, 사회복지관등 총 522개 시설(41%)을 대단위 전국적 규모로 실시하였다(〈표 2〉 참조). 이를 각 시설의 종류별

로 살펴보면 다음과 같다. 1999년도에는 장애인복지관과 정신요양시설에 대해서 평가를 실시하였고, 각각 36개, 59개의 시설을 평가함으로써 시설대비 평가실시율은

〈표 2〉 사회복지기관수 대비 평가실시 현황

시설종류		대상기관	평가원료기관	
		1999. 10. 31 기준	1999년	2000년
장애인복지관		54 ¹⁾	36 (67%)	-
정신요양		67 ²⁾	59 (88%)	-
아 동	영아	26		28 (100%)
	유아	211		-
	직업보도	5	-	-
	교호시설	6		-
	자립지원	13		-
	일시보호	10		-
소 계		271	-	28 (10%)
노 인	무료양극	92		-
	무료요양	94	-	60 (64%)
	실비 및 유료	43		-
	소 계	229		60 (26%)
여 성	모자보호	37		37 (100%)
	모자자립	3		2 (67%)
	모자일시	6	-	5 (83%)
	미혼모	8		8 (100%)
	선도보호	11		9 (82%)
	성폭력보호	6		-
소 계		71	-	61 (86%)
장애인	지체	37		-
	시각	10		-
	청각·언어	14		-
	정신지체	59	-	54 (92%)
	중증요양	68		-
	근력			-
소 계		188	-	54 (29%)
부랑인	성인부랑인	35	-	34 (97%)
	아동부랑인	9		-
	소 계	44	-	34 (77%)
사회복지관		343	-	285 (83%)
계		1,267	95 (7%)	522 (41%)

주: 1), 2)는 1998년 말 자료임.

67%, 88%를 나타냈다. 2000년도에 실시한 평가의 경우는 1999년도에 실시한 장애인복지관과 정신요양시설을 제외한 총 6개 종류의 시설에 대해 평가가 이루어졌다. 먼저 아동의 경우에는 단지 영아시설에만 28개의 시설에 실시되어 영아시설에 대한 시설대비 평가실시율은 100%를 나타냈지만 전체적인 아동시설에 대해서는 271개의 시설 중에 28개의 시설이 실시되어 분야들 중에 가장 낮은 실시율인 10%를 나타냈다. 노인시설은 총 229개의 시설 중에 무료요양시설만 60개를 실시함으로써 전체적으로 26%의 실시율을 보였다. 여성의 경우에는 성폭력 보호시설을 제외하고 관련시설 모두에 대해서 평가를 함으로써 분야 중에 가장 높은 실시율인 86%를 나타냈다. 장애인시설의 경우에는 정신지체시설 59개 중 54개를 실시하여 전체 188개의 시설에 대해서는 29%의 실시율을 나타냈다. 부랑인 시설은 44개 중 성인부랑인시설 34개에 대해서 평가를 실시하여 전체적으로 77%의 실시율을 나타냈고, 마지막으로 가장 많은 시설인 사회복지관의 경우에는 343개의 시설 중에 285개를 실시함으로써 83%의 평가실시율을 보였다. 이상의 결과로 1999년에는 총 1,267개의 시설에 대해 95개의 시설을 평가함으로써 7%의 평가실시율을 보였지만 2000년도의 경우에는 522개의 시설을 평가함으로써 6배가 증가한 41%의 실시율을 보였다. 따라서 1999년과 2000년도를 합한 총 평가실시기관은 617개로써 전체 사회복지기관 대비 약 48%의 총 실시율을 보였다.

2) 평가의 문제점과 분석

(1) 평가에서 나타난 문제점

지난 2년간 보건복지부의 평가에서 나타난 다양한 문제점들을 보건사회연구원(변재관 외 10인, 2000)에서 발행한 보고서를 통해서 간접적으로 종합하였다. 비록 매우 제한적이지만 피드백 차원에서 지적하고 있는 보고서 내용들과 이를 보완하기 위해서 핵심정보제공자와의 면담에 의해서 다음과 같이 3개 영역으로 주요 문제점들을 취합하였다.

① 관리상의 문제점

-전문 평가인력의 부족

- 전문 평가인력들의 일정관리문제
- 평가인력의 처우개선 문제
- 평가 전문기술의 축적 불가능
- 보건사회연구원의 업무과중 및 이중업무
- 평가인력에게 교육을 충실히 시키지 못함

② 평가시행상 문제점

- 기준 부재로 자의적 해석
- 정성적 자료의 평가에서 주관성 개입
- 시설별로 상이한 평가영역 및 평가항목
- 작은 기관들에 불리한 평가내용
- 표준화된 최소한의 기준의 부재로 지나친 긍정적 평가
- 실적수치의 표준편차가 지나치게 높음
- 상충되는 평가지표들
- 평가지표를 평가자가 다르게 이해
- 각 개별시설에 해당되지 않는 평가지표들
 목표달성수치의 비적절성
- 현장평가 팀원간의 격차조정 보완부족
- 평가지표에서 시설별 특수성을 반영 못함
- 프로그램 질적인 측면에서 심도 있는 평가가 미흡
- 정량적 평가항목들이 지나치게 많음

③ 평가대상의 문제점

- 시설종사자들의 추가 업무부담
- 서류조작 등 정직성 문제
- 거주자조사와 종사자조사에서 사실과 다르게 응답
- 이용자 만족도 조사의 신뢰성 부족
- 입소자의 지적 능력부족으로 응답 불가능
- 서류 및 문서화 미비
- 서식 등이 통일되어 있지 않음

(2) 문제점 분석

이상의 3가지 영역에서 나타난 문제점들을 우수한 평가체계의 기본 요소들인 표준성, 측정가능성, 통제성, 충분성, 비교성, 신뢰성, 비용효과성이라는 7가지 기준에 의해 분석하면 다음과 같다(APQC, 1999).

① 표준성 미비

제대로 평가를 하려면, 먼저 평가자 입장에서는 평가체계가 잘 구축되어 있어야 하며, 이를 사용하는 평가인력도 잘 훈련되고 평가를 합리적으로 실시할 수 있는 충분한 예산도 확보되어야 한다. 피평가기관의 입장에서도 서류 등 문서화 작업이 되어 있으며 정직하게 평가에 임해야 한다. 하지만 현재 우리는 표준 매뉴얼도 제대로 구축되어 있지 않는 상태이다. 현행 평가체계에서 발생하는 대부분의 평가시행상의 문제들은 표준성 미비에서 비롯된다. 자의적 해석, 주관적 판단, 지나친 긍정적 평가 등의 문제는 우리의 평가체계가 기준이 되는 문서화된 '표준 매뉴얼'(standards manual)⁴⁾이 있어야 하는데 제대로 준비가 되어있지 않기 때문이다(CARF, 1998; 1994). 보건복지부 시설평가체계가 1999년 6월에 평가위원회가 구성되어 불과 6개월만에 만들어 그 해 11월에 예비평가를 실시하여 표준 매뉴얼을 제대로 구비하지 못하고 시작했다고 할 수 있다. 더욱이 2년이 지난 지금도 이러한 미비점을 보강하려는 진지한 노력을 찾아보기가 힘든 것 같다. 전국규모의 평가실시를 통해서 얻어지는 평가 전문기술이나 피드백이 계속 축적하여 '표준 매뉴얼'의 완성도를 높여가야 한다.

② 측정가능성 부족

정성적 평가에서 주관성이 개입하거나 서비스의 질을 심도 있게 측정하는 것을 기

4) 인증제에서 사용하는 표준 매뉴얼은 다음과 같은 7개 부분으로 구성된다. ① 인증기준: 전체 인증심사의 기준이 되는 것으로 주의깊게 잘 숙지해야 한다. ② 표준: 현장 심사과정에서 구체적으로 항목별 수준을 제공한다. 표준은 두 부분으로 구성된다. 하나는 기관의 관리와 운영이고 다른 하나는 프로그램에 관한 내용이다. ③ 해설서: 표준 적용의 일관성을 극대화하기 위해서 만들었다. 표준의 적용은 문서 확인과 현장확인 둘 다 포함한다. ④ 용어풀이: 표준 매뉴얼에서 사용된 용어의 정의와 내용을 설명한다. ⑤ 서류 색인: 조사와 인증과정을 위해서 필요한 서류의 리스트를 나타낸다. 이는 평가자로부터 심사된 정보를 설명하는 것뿐 아니라, 효과적인 프로그램을 지원하고 유지하는 데 필요하다. ⑥ 색인(목차): 평가항목, 표준 번호, 그리고 페이지 번호를 복사라 한다. ⑦ 인증 정책과 절차: 심사의 인증과정에서의 시행절차를 기술한다.

피하는 것은 바로 측정가능성(*measurability*)의 부족 때문이다. 평가시 계량화가 불가능하거나 질적인 측면에서의 평가가 요청되는 경우에는 비계량지수를 신중하게 설정해야 한다(Martin & Kettner, 1996; Chwardt, 1990). 예를 들면 서비스 질의 경우, 클라이언트가 사회복지서비스를 적절하게 제공받을 때와 부당하게 대우받을 때를 LOF(*Level of Functioning Scales*) 혹은 BARS(*Behavior Anchored Rating System*) 같은 측정기법을 사용하여 서비스의 차이점을 계량화할 수가 있어야 한다. 아직도 비계량지수를 익숙하게 설정할 만큼의 평가전문가가 흔치 않으며, 시설별로 평가안을 따로 만들어, 전문적 지식을 활용하는 기회를 가질 수 없었다. 측정가능성이 없는 평가항목 지표들을 계속 사용하면 평가현장에서 혼란과 불신이 생기는 것은 자명하다.

③ 통제 불가능

사회복지기관들을 비교시에 수량적으로 나타난 결과만을 가지고 단순 비교하는 것은 매우 위험하다. 즉 사회복지기관의 설립연도, 재정도, 위치, 크기, 설립목적 등 그 기관의 고유성(*locality*)이 평가체계에 반영되지 않는다면 많은 기관들이 평가결과를 가지고서 공정하지 않다고 주장할 수 있다. 대부분의 고유성이 종사자의 의지라도 통제가 불가능한 경우가 많은데도 평가항목으로 비중 있게 평가된다면 피평가기관으로서는 억울하게 당하는 꼴이 된다. 그래서 평가항목을 선정하는 기준에 사회복지기관들이 통제할 수 있고 납득하게 만드는 통제가능성(*controllability*)을 강조하게 된다. 만약에 사회복지기관에서 아무리 노력해도 통제가 불가능한 환경적 의구요소들을 평가항목으로 계속 고집하게 되면 부익부 빈익빈현상이 나오게 되며 평가에 따른 동기 부여도 힘들게 된다.

④ 충분성 부족

평가지표항목끼리 상충되는 경우는 피해야 하겠지만, 만약에 있으면 충분한 근거를 제시하고 그들을 상황에 따라서 구분할 수 있게끔 충분한 설명을 해야한다. 이러한 충분성(*sufficiency*)은 평가의 내용이 불분명하거나 항목간 구분이 모호하여 중복 평가되는 것에도 적용이 되며 만약에 충분한 근거를 찾지 못하면 배제시켜야 한다. 충분성 부족은 앞에서 언급한 표준성 부족과도 관련성을 가지게 된다.

⑤ 비교 불가능

보건복지부가 평가체계를 처음 구축할 때부터 시설별로 따로 진행하는 분할주의 방식을 취함으로써 시설별로 상당히 다른 평가지표들이 나타나게 되었다. 물론 시설별 특성을 감안한 것도 있겠지만 비교가능성(*comparability*) 라는 관점에서는 상당한 취약성을 가지게 되었다. 자원의 재배분이나 정책결정에 비교성은 상당히 중요한 의사결정의 근거가 됨에도 불구하고 그 가능성을 약화시켰다. 원칙적으로 평가항목은 계속성을 유지하여 안정성을 확보함으로써 연도별 시설별로 갱신정도의 비교가 가능하도록 해야한다.

⑥ 신뢰성 부족

평가인력이 부족하고 교육을 충분히 시키지 않거나, 평가자가 지나치게 긍정적 평가를 하거나, 투명성이 없거나, 임의적인 평가를 하면 평가체계 자체에 대한 신뢰성(*reliability*)을 잃게 된다. 정확하고 일관된 정보수집을 위해서 인증심사를 하는 전문가에게 어느 정도의 자격을 사전에 요구하며 이들에게 평가심사와 관련한 소정의 교육과정을 이수하게 한다. 평가인력이 기준을 잘못 적용하거나 평가자가 평가방법이 부족하며 능력이 떨어짐으로 인해 신뢰성을 얻지 못하는 경우도 있다. 평가상의 문제점이 평가체계 자체에도 있을 수가 있지만 그것을 심사하는 평가인력에 의해서 공정성이나 객관성이 훼손되어 신뢰성을 잃는 경우도 많다.

⑦ 비용 과대

우리의 평가체계가 많은 사회복지기관들을 대상으로 인증제와 같은 방대한 평가방식을 추구함으로써 평가인력을 관리하기도 힘들고 정부가 계속 부담하여야 하는 재정적 부담도 만만치 않다. 벌써 평가인력에 대한 처우가 문제점으로 지적되는 실정이고 보면, 과도한 평가비용 때문에 평가를 제대로 할 수 없을 수도 있다. 형식적인 평가가 되지 않으려면 평가에 소요되는 제비용들이 합리적으로 산정되어 예산이 확보되어야 한다.

4. 미국 평가체계와 제한점

1) 미국의 평가체계

최근 선진국에서 공공기관이나 프로그램을 평가할 때 주로 사용하는 평가방법이 수행평가(Performance Measurement)이다. 수행평가는 기존의 전통적인 평가방법인 성과중심(Outcomes-Oriented Approach) 접근보다 더 확장된 개념을 가진 평가방법으로, 상당한 기간동안 주된 평가방법으로서 앞으로 활용될 것으로 전망되고 있다(Alberta, 1998; US GAO, 1998).

수행평가중심의 미국의 사회복지기관평가는 접근하는 방식에 있어서 전반적인 공공부문의 개혁과정에서 종합적으로 이루어지고 있다. 공공부문의 개혁을 위한 수단으로서 연방정부와 주 정부 등에 대한 전반적인 평가가 시작되었으며 사회복지기관에서도 같이 적용하게 되었다. 특히 GPRA(Government Performance and Results Act)법 제정과 GASB(Governmental Accounting Standards Board) SEA보고서(Service Efforts And Accomplishments)의 요구는 수행평가를 더욱 더 촉진시키는 계기가 되었다(김통원, 2000).

(1) GPRA 내용

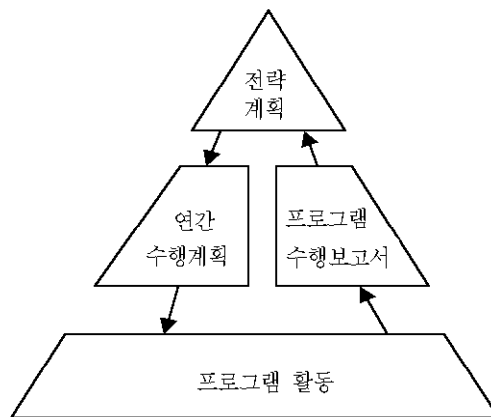
1993년에 수행평가를 시행하고 보고토록 하는 법안 - <행정수행과 그 결과에 대한 법률>(GPRA: Government Performance and Results Act) - 이 통과되었다. 이 법안의 제정을 통하여 모든 연방기관과 사회복지기관을 포함한 연방, 주, 지방으로부터 자금을 지원 받는 기관으로 하여금 행정업무의 수행기준과 목표에 대한 계획을 세우고 예산지출분야에 대한 성과와 수행정도를 보고하도록 규정하고 있다. 이 법률에 따라 1997년부터 모든 미국 연방기관으로 하여금 수행기준과 목표에 대한 전략계획을 수립하도록 규정하고, 1999년 회계연도부터 모든 정부기관이 수행평가를 실시하고 성과를 보고하도록 규정하고 있다. 이 법은 미국 공공부문에서 수행평가가 체계화되도록 하는 데 많은 영향을 주었다(Epstein, 1992).

아래의 통합주기 <표 3>에서 보는 바와 같이 우선 전략계획(strategic plan)을 세워야 한다. 전략계획은 기관의 설립목적을 달성하기 위해서 최소 5년간의 계획으로 이

루어지고, 인력, 자본, 정보 등 기타 자원들을 가지고 기관의 목적을 어떻게 달성할 것인지에 대한 기술을 포함한다. 또한 최소한 3년마다 계획을 재조정하고, 복직달성에 영향을 주는 주요한 외부요인에 대해서도 기술해야 한다(DTFMS, 1993). 전략계획이 세워지고 나면 프로그램 예산과 활동에 관한 연간 프로그램 수행 계획서를 제출해야 한다. 이는 GPRA가 중앙정부 예산관리처(*Office of Management Budget*)에 제출하도록 요구한다. 수행계획서는 전략 계획에서 도출된다. 세운 활동들은 목표지향적이어야 하고, 양적이고, 측정 가능한 것이어야 한다. 전략계획서는 프로그램 목적달성에 필요한 과정, 기술, 인적, 물적 자원, 정보 다른 자원들에 대한 기술도 포함되어야 한다. 그리고 수행 및 성과를 평가는 척도들이 마련되어야 한다.

연간 수행계획서에 의해 프로그램 활동이 이루어진다. 프로그램이 이루어진 후 연간 프로그램⁵⁾ 수행보고서를 작성해야 한다. 프로그램 수행보고서는 수행목적과 실제로 달성된 프로그램 결과를 비교하고, 프로그램 목표달성에 대하여 검토하고 기술해

〈표 3〉 통합주기



자료 : Department of the Treasury Financial Management Service. 1993. "Performance Measurement Guide," *Government Publications Section*, p.35.

5) 한국은 공공부문에서 기관중심으로 정부가 예산(보조금)을 주지만 미국의 경우는 대부분 프로그램중심으로 정부자금이 나간다. 최근에 우리 정부의 각 부서에서 프로그램 사업중심으로 운영단체를 모집하는 것처럼, 우리도 점차로 공공자금이 프로그램/사업위주로 집행될 것이다.

야 한다. 수행이 잘 이루어지지 않았을 경우 그 이유를 설명하고, 수행정도가 계획의 목표와 일치하지 않을 경우는 이유를 보고서에 설명한다. 수행보고서는 다시 프로그램 전략을 계획하는 데 환류(feedback)를 제공한다. 즉 프로그램 목적 및 목표달성, 프로그램 결과, 수행이 안된 이유 등에 관한 종합적 분석 및 정보는 새로운 프로그램을 계획하고 전략을 세우는 데 도움이 된다(Martin, 2001; Congress of USA, 1993).

(2) SEA 보고서의 중요성

GPRAB보다 약간 먼저 1991년에 미국의 정부회계표준위원회(GASB)는 Service Efforts And Accomplishments(SEA) 보고서를 발표하여 서비스나 프로그램의 효율성, 질적 향상, 효과성 측면을 중심으로 보고하도록 하는 기준을 제시했다(Martin, 2001; 김통원, 2000). 새로운 SEA 기준은 과거 정부에 제출하는 회계보고가 수량적 측면과 생산성 측면을 강조한 단편적 성격에서 새로운 변화를 의미하는 것이다. 미국의 정부회계표준위원회(GASB)는 특히 지난 몇 년 동안 정부기관들만이 아니라 공공자금을 사용하는 모든 사회복지기관에도 의무적으로 SEA(Service Efforts & Accomplishments) 보고서를 요구하고 있다. 다시 말해서 미국의 모든 사회복지기관들은 정부의 공공자금을 지원 받는 만큼 수행평가의 결과에 대해서 GASB에서 요구하는 SEA 보고서를 제출해야하는 의무를 가지게 되었다. 연방정부 혹은 주와 지방정부의 지원을 받는 민간 사회복지서비스 기관들이 의무적으로 수행평가 결과의 수집과 보고에 참여해야 한다면, 이에 따른 영향은 매우 크게 된다. 왜냐하면 미국의 사회복지기관들에 대한 평가는 결국 SEA 보고서로 집약되기 때문이다(Martin, 2001; 김통원, 2000).

2) 미국평가체계의 한계점

비록 미국평가체계는 우리의 평가체계에 비해서 간단하고 명확하여 객관성 확보와 비용측면에서 탁월하지만, 한계점을 가지고 있는 것도 부인하기 힘들다. 왜냐하면 측정 가능한 평가항목만으로 구성되는 핵심수행측정지표(Key Performance Indicators) 위주로 운영하기 때문이다. 다음은 수행평가보고서에서 나타나는 7가지 한계점들에 대한 것이다(GASB, 1998; Martin & Ketter, 1996).

첫째, 미국 수행평가체계는 몇 가지 핵심수행 측정지표만을 포함하고 있어, 여러

각도에서 종합적으로 측정하지 못하는 문제가 있다. 따라서 개별서비스, 집단서비스, 프로그램, 기관의 수행에 대해서 다양한 각도에서 포괄적으로 측정하는 측면이 부족하다.

둘째, 미국 수행평가체계는 수행을 증진시키는 방법이나 서비스나 활동 요인 등 외부요인들과 업무를 개선시키는 구체적인 개선책 등이 결과 보고서에 나타나 있지 않다. 앞에서 언급한 것처럼 불과 10~20개 정도의 핵심수행 측정지표만을 포함하므로 이러한 정보들에 대해서 매우 제한적일 수밖에 없다.

셋째, 미국 수행평가체계는 제공된 서비스가 사용된 과정(*process*)에 관한 정보를 포함하지 않고 있다. 즉 미국평가체계는 성과에 관한 정보를 설명하는 보고서로는 유용하나, 프로그램의 산출(*outputs*)과 성과(*outcomes*)를 이끌어 내는 과정(*process*)과 서비스 전략에 관한 평가는 미국평가체계 범주에 나타나 있지 않다. 특히 평가과정은 측정하기가 용이하지 않으므로 처음부터 제외를 하고 있다고 해석된다.

넷째, 미국 수행평가체계는 목적과 목표의 달성을 항상 적절히 측정하고 있다고 볼 수 없다. 예를 들어, 범죄예방 목적을 달성하는 데 있어 범죄율이 이를 잘 측정하고 있다고 믿고 있으나, 범죄율의 감소가 실제로 범죄예방을 달성하는 데 적절한 평가항목인지는 많은 의문점이 남는다. 즉 평가하려는 내용과 실제로 사용되는 평가항목이 잘 연계되지 않는 경우도 있다. 이러한 한계점은 우리의 평가체계에서도 흔히 볼 수 있는 측정상의 오류이며 의도된 정보를 제대로 포착하지 못하는 타당성 그리고 관련성에서의 문제들이다(APQC, 1999).

다섯째, 미국 수행평가체계에서 성과(*outcomes*)를 항상 측정할 수 있는 것은 아니다. 왜냐하면 아무리 중요한 성과라도 측정할 수 없는 경우가 있기 때문이다. 즉, 객관성을 확보하기 위해서 측정 가능한 것만 측정하기를 요구한다. 지나친 객관성의 추구는 정성적인 자료에 대한 기회현상까지 보이고 있다. 반면에 우리의 평가체계는 정성적인 자료를 무리하게 조작화하는 경향을 보이는 것과는 상당히 대조적인 현상을 보이고 있어 흥미롭다.

여섯째, 미국 수행평가체계는 기관의 책무성을 평가하는 데 필요한 모든 자료를 충분히 제공하고 있지 않다. 인증심사보다 훨씬 적은 평가항목으로는 충분한 정보를 수집할 수 없기 때문이다. 미국의 수행평가에 사용되는 평가항목의 수와 우리의 평가항목의 수는 작게는 5배에서 많게는 10배정도의 차이를 보이고 있어, 정보의 절대적

인 양에서 많은 차이를 가지고 있다.

마지막으로 GPRA에 의한 평가가 처음으로 시행된 것은 1997년으로 아직 초기단계이고 충분한 검증이 안되었다고 주장한다. 하지만 유사한 평가접근이 OECD 여러 나라에서 이미 오래 전부터 시행되어 왔으며 그 효용성은 어느 정도 검증되었다고 할 수 있어 이러한 주장이 지나친 신중함을 나타내고 있다.

5. 한국과 미국 평가체계의 특성비교

보건복지부에서 시행한 평가체계와 미국 사회복지평가 체계의 주요 내용들을 비교하였다. 평가체계를 비교하는 틀은 PUMA(Public Management, 1999)에서 사용한 OECD 국가간에 평가체계를 비교한 것을 사용하였다. PUMA에서는 6개 영역에서 비교하나 본 연구에서는 결과중심의 관리를 빼고 준비성을 첨가하여 총 6개 영역에서 비교하였다.

1) 준비성

미국의 GPRA(*Government Performance and Results Act*) 평가체계의 경우는 1993년에 법이 제정되고 처음으로 시행되기는 1997년이며 전면적으로 실시한 것은 2000년부터이다. 반면에 한국에서는 사회복지현장에서 평가를 최소한 3년간 유예를 요청했지만(이청자, 2000), 서울시는 2개월만에 그리고 보건복지부는 6개월만에 첫 평가를 강행하였다. 앞에서 지적한 것처럼 우리는 인증제에서 가장 중요한 '표준 매뉴얼'을 제대로 완성하지 못한 상태로 사회복지현장에서 평가가 시작되었다. 오히려 사회복지계보다 정부측에서 제대로 준비없이 평가를 시작하였다고도 볼 수 있다. Heisenberg(2000)의 불확실성 원리(*Uncertainty Principle*) 효과에 의하면, 사람들이 평가 및 측정과 관련한 이야기를 시작하면서부터 이미 사람들의 행동에서 변화가 오기 시작한다는 원리이다. 실제적인 평가를 하기도 전에 이미 평가를 한 만큼의 간접적인 효과를 가져온다는 것이다. 보통 전면적인 평가를 위해서 예비평가, 준비평가, 본 평가 등으로 3년 정도의 시간을 주는 것은 이러한 원리와 같은 맥락으로 볼 수 있

다. 준비기간이 충분하면 자발적인 참여 역시 자연스럽게 끌어올 수 있고 시행착오도 줄일 수 있었을 것이다.

2) 평가의 목적과 접근

미국 평가의 궁극적인 목적은 사회복지기관들이 유연한 조직으로의 전환을 추구하는 내부관리를 목표로 하는 것이며, 외부적 목표는 성과에 대한 책임성 확보를 추구하고 있다. 비록 효율성을 강조하고 있지만 예산절감은 부차적인 목표로 되어 있다(CFOC, 1997). 우리도 마찬가지로 시설운영의 개선과 서비스의 질 제고를 유도한다고 표방한다. 하지만 미국처럼 예산절감을 평가의 목표로 두지는 않았다. 미국평가의 근거가 미국 연방정부 모든 부서들과 지방정부들 그리고 공공자금을 받는 사회복지기관들까지 모두 적용시키는 GPRA와 GASB의 SEA 보고서 등 법률에 기초를 두고 추진되는 것에 비해 우리는 사회복지사업법에 근거를 두고 보건복지부 산하에 있는 사회복지기관만을 대상으로 평가를 시행한다. 다시 말해서 미국은 평가대상에 모든 공공분야를 포함하며 동일한 평가체계를 적용하지만 한국은 단지 사회복지기관들만 적용하고 있다. 평가체계의 적용범위가 엄청나게 차이가 있지만, 두 나라 모두 위로부터의 주도(*Top-Down*)로 평가가 시작되었다. 한국이 준비성 없이 평가를 시작했지만 사회복지계에서 독자적인 평가체계를 먼저 시작한 점은 선도적인 것이라 평할 수 있다.

3) 평가체계

앞에서 이미 언급하였지만, GPRA 법에 기초하여 사회복지기관을 비롯한 모든 공공기관들을 평가하는 미국의 평가체계는 우리나라의 평가체계와 매우 다르다. 무엇보다도 평가항목이 적으며 측정가능성을 강조한다. 반면에 우리의 평가는 일종의 인 증심사 제도를 채택하고 있어 평가항목이 100~150개 정도로 미국의 평가항목보다 보통 10배 정도 많다. 또한 측정하기가 용이하지 않은 정성적 평가항목도 적지 않다. 평가항목이 많고 정성적 평가도 하려면, 그만큼 평가에 소요되는 시간과 비용이 많이 들며 객관성 확보를 위해서 평가심사원들도 잘 교육해야 한다는 것을 의미한다. 많은

평가항목을 하루만에 평가를 하는 경우 잘못하면 ‘형식적인 평가’로 끝날 수 있다. 미국의 평가체계는 단순히 성과와 실적을 강조하는 결과중심의 접근을 하고 있는 반면에, 우리의 평가체계는 조직의 질(*organizational quality*)을 살펴보기 위해서 시설과 환경, 운영 및 인력, 서비스 및 인권, 그리고 지역사회와의 관계 등 하위영역들에 대한 품질(*quality*)을 다양한 측면에서 평가한다. 다시 말해서 미국의 평가체계는 단순 명료하며 객관성 확보에 우위를 가지지만, 우리의 평가체계처럼 얻어지는 정보가 풍부하지 못하다. 하지만 앞에서 지적한 것처럼 우리 평가체계는 ‘표준 매뉴얼’의 미비로 그리고 무엇보다도 ‘평가체계의 운영 및 관리의 부실’로 인하여 평가체계의 완성도와 효용성이 매우 떨어진다.

4) 서비스의 품질

미국은 서비스의 품질에 대한 기준을 제시하며 서비스에 대한 만족도 조사를 실시한다. 우리도 서비스의 품질에 대한 평가항목들이 비교적 다양하게 제시되고 있으며 서비스 만족도 조사를 장애인 시설을 제외한 대부분의 기관에서 실시하고 있다. 실제로 미국의 평가체계보다도 서비스의 품질 측정 측면에서 한국이 우위에 있다고 할 수 있다. 단지 우리의 경우 서비스의 적절한 표준(*standards*) 혹은 기준을 제시하지 않고 있다. 과연 어느 정도가 적절한 수준인지를 설정해 주는 표준(*standards*)이 없어 그 적정성을 가늠하기가 힘들다. 비록 최선이 아니고, 최저수준을 지향한다 해도 표준을 각 개별 평가항목과 더불어 제시하여야 한다.

5) 내부평가/검토

내부관리적인 측면에서 한국의 평가체계는 미국의 평가체계보다 훨씬 우위에 있다. 한국의 평가체계는 평가를 준비하는 과정이나 실제 시행과정 그리고 평가 후에 획득되는 정보의 질과 양이 미국에 비해서 엄청나게 많기 때문이다. 평가가 학습 지향적(*teaching-oriented*) 관점을 내포하게 되어 지속적인 개선을 위한 수단이 된다. 평가결과가 나쁘지만 않다면 우리의 사회복지기관들은 무료로 인증심사 수준의 서비스를 정부로부터 받고 있다고 보면 된다. 본 연구는 정부가 평가의 주체가 되어 시행하

는 외부평가라는 측면에서 보기 때문에 한국의 평가체계에 대해서 비교적 비판적이지만, 한국의 평가체계를 내부평가 또는 내부관리라는 측면에서 보면, 품질과 과정을 중시하는 사회복지서비스의 본질에 부합되는 것으로 긍정적으로 볼 수 있다.

6) 결과의 공개와 활용

미국에서는 평가의 결과에 대해서 공개의 원칙을 강조하여 구체적으로 목표대비 결과에 대해서 명확하게 연간백서로 발행한다. 우리도 평가결과를 보고서로 발행하고 있지만 많은 정보들이 제한적으로 공개되고 있다. 미국은 또한 공개된 정보를 가지고 사업 예산 편성 과정에 결과를 반영하며 개인보수와 조직단위의 성과만영 성과급에까지 연계시킨다. 한국도 인센티브 제도를 운영하고 있다고 볼 수 있으나 개인보수나 조직단위의 성과급에까지 영향을 주기에는 어렵다. 왜냐하면 미국의 평가체계는 핵심수행측정지표를 중심으로 명확하게 측정하기 때문에 객관성확보가 쉽지만 우리의 평가체계는 아직도 완성도면에서 떨어져 평가결과를 적극적으로 공개하고 보상체계와 직접연계하기에는 공정성이나 신뢰성 측면에서 평가주체인 정부도 상당히 부담스러워 하고 있기 때문이다.

6. 국내 사회복지 평가체계의 향후 전략

지금까지 보건복지부가 시행하고 있는 사회복지기관 평가의 현황과 문제점, 그리고 미국의 주요 평가체계 내용들과 세한점을 세시하면서 두 나라의 평가체계들을 장단점들을 비교하였다. 이러한 사실에 입각해서 향후 우리의 평가체계의 개선을 위해서 단기전략과 장기전략으로 나누어서 제시할 수가 있다. 본 연구에서 단기전략이란 현행 평가체계를 그대로 유지하면서 구체적으로 보완을 하는 것이고, 장기전략은 현행 평가체계의 주체가 정부가 되어서 계속 운영하기에는 근본적으로 문제가 있으므로 새로이 판을 짜는 것을 의미한다. 먼저 단기적인 관점에서 전략들을 살펴보면 다음과 같다.

1) 단기전략

(1) 평가전담기구의 설립

한국의 현행 사회복지기관 평가체계의 관리 및 운영상에 있어 부실한 문제를 극복하기 위해서는 전문적인 관리를 위한 평가전담기구가 필요하다. 보건사회연구원이 평가업무의 관리적인 측면에서 별도의 기관설립을 제의한 것처럼 지금과는 다른 새로운 운영방식을 모색해야 한다고 생각한다.

(2) 표준 매뉴얼의 완성

평가자의 자의적 해석, 주관적 판단, 지나친 긍정적 평가 등의 문제는 표준성 미비에서 비롯된다. 앞에서 지적한 것처럼 우리의 평가체계는 기준이 되는 문서화된 '표준 매뉴얼'(standards manual)이 필요하다. 비록 우리의 사회복지평가체계가 내부관리라는 측면에서 유용한 방식이지만 그리고 표준 매뉴얼의 보강 등 아직 그 완성도를 높이기 위한 노력들이 필요하다.

(3) 평가심사원 교육 및 전문화

앞에서 언급한 것처럼 평가자에 대한 교육이나 전문성이 부족하여 지나치게 긍정적 평가를 하거나, 투명성이 없거나, 임의적인 평가를 하게 되면 평가체계에 대한 신뢰성을 잃게 된다. 따라서 인증심사를 하는 심사원에게 어느 정도의 자격을 요구하고 이들에게 소정의 교육과정을 이수하게 함으로써 평가기준을 적용시 정확성과 신뢰성을 가지게 해야 한다.

(4) 충분한 예산확보

우리의 평가가 비용이 많이 드는 인증제 방식을 추구하고 있으면 그 비용도 현재 정부가 계속 부담하는 것으로 되어 있다. 따라서 정부의 재정적 부담도 만만치 않다. 수많은 기관들을 제대로 평가하기 위해서는 평가에 소요되는 제비용들이 합리적으로 산정되어 예산이 확보되어야 한다. 그렇지 않으면 과도한 평가비용 때문에 형식적인 평가가 될 수 있다.

2) 장기전략

본 연구자는 현행 평가체계는 고비용 부담과 공정성 및 객관성 확보에 대한 시비로 인하여 계속 유지하기는 힘들 것으로 예상한다. 다시 말해서 현행 평가제를 활성화하여 온전한 평가체제로 개선시키는 작업은 구조적으로 어려움이 있다고 본다. 따라서 장기적인 관점에서 우리나라의 평가체계의 재정립을 위해서 다음과 같은 몇 가지 제안을 하고자 한다.

(1) 상호보완관계

장기적으로 우리 정부도 현재와 같은 인증제방식의 평가제도보다 미국처럼 수행평가중심으로 체도를 바꿀 가능성이 많다. 동시에 현행 인증제 방식은 민간 인증기관에 넘겨 자율적으로 운영할 가능성이 크다. 향후 한국의 사회복지기관 평가체계가 정부에서 시행하는 전면적인 수행평가방식과 사회복지기관이 자율적으로 획득하는 인정제로 이원화된다면, 각기 성격이 다른 두 개 평가방식의 장점들을 보완적으로 활용할 수 있어 사회복지기관과 클라이언트 모두에게 이익이 될 수 있다. 정부입장에서도 성과중심(결과중심)의 단면적인 평가와 더불어 인증과 관련한 기관들에 대한 포괄적인 정보를 활용하여 사회복지기관평가를 다면적으로 평가함으로써 평가결과 활용에 신뢰성을 확보할 수 있다. 사회복지기관도 다면적 평가로 일방적인 평가로부터 비롯되는 위험성을 분산할 수 있어 좋다.

(2) 과학성 추구

앞에서 정부의 정책적 책임을 강조하였지만 평가와 관련한 정책만이 아니라 모든 정부의 정책이 보다 더 신중하고 과학성을 추구하여야 한다. 미국에 경우 정책을 수행하기 앞서 반드시 실증적인 연구를 선행하여 충분히 검토를 한 후에 시행하기 때문에 정책적 실패를 사전에 예방한다. 정책이 한번 실패하면 국민의 세금 낭비는 물론이고 사회복지계의 경우처럼 본연의 서비스 업무에게 부정적 영향을 미칠 수도 있다.

(3) 결과활용의 극대화

만약에 평가가 평가를 위한 평가로 끝나고 형식적인 평가로 끝난다면 평가로부터

얻어지는 정보는 그대로 사장되고 만다. 실제로는 평가를 받는 것이 목적이 아니고 얻어지는 정보를 가지고 대내외적으로 적극 활용하는 것에 그 목적이 있다. 평가체계의 완성을 통해서 제대로 평가해서 얻어지는 평가결과는 그 활용에 있어 신뢰성을 확보할 수 있다. 평가의 결과는 한마디로 '정보'이다. 평가의 결과정보가 정확하고 객관성을 확보하고 있으면 피평가자들도 신뢰하게 되고 정부에서도 과감하게 인센티브제를 활용할 수 있다. 평가결과의 활성화는 사회복지기관이나 정부에 막대한 반대급부를 제공하게 한다.

(4) 정치성 배제

평가는 논리성과 합리성에서 출발한다. 평가체계의 개발이나 시행과정 그리고 그 활용이 합리적 원리에 바탕을 두고 있다. 평가제운영은 합리적 원리에만 충실해야지 다른 정치성은 배제해야 한다. 특히 평가가 본연의 목적 이외는 사용되지 말아야 한다. 예를 들면 평가를 사회복지기관을 통세하는 수단으로 활용하거나 국가 예산을 줄이는 방법으로 사용하는 경우는 곤란하다.

(5) 평가문화의 정착

아직까지 우리의 평가문화는 제대로 정착되어있지 않다고 판단한다. 왜냐하면 사회복지기관들 중에 평가에 대한 편견 혹은 불신을 가지고 있으며 서류를 조작하거나 부정직하게 응답을 하는 경우도 종종 목격이 되기 때문이다. 또한 평가체계를 만드는 주체들도 평가에 대한 기본적인 이념과 원리를 잘 이해하지 못하고 있거나 왜곡된 시각을 가지고 있기도 하다. 지나치게 평가를 상이나 벌을 주는 것(*Punishment & Reward System*)으로 여기거나 수량적인 평가항목들만이 객관성이 보장되는 것으로 오해하는 이도 많다. 그래서 온전한 평가를 시행하기에는 아직 '전체적인 분위기'가 제대로 성숙되어 있지 않다고 생각한다. 우리는 이미 평가의 시대에 살고 있다. 사회복지기관들도 평가문화(文化)의 정착을 위해서 적극적으로 노력할 필요가 있다. 종국적으로 평가문화의 정착이란 종전의 규제나 타율적인 행정중심에서 비롯되는 기존 가치나 행동규범이 몰락하고 사회복지전문가의 자율성을 바탕으로 클라이언트 중심, 질 중심, 전문성 중심으로 평가와 관련된 새로운 가치들이 도래함을 의미한다.

•참 고 문 헌•

- 기획예산처. 2000. 《성과주의예산 2001년 시범사업 추진 지침》, 기획예산처.
- 기획예산위원회. 1999. 《성과주의예산제도 추진방안》, 기획예산위원회.
- 김응석 외 5인. 1997. 《아동복지수용시설 운영평가 및 개선방안》, 한국보건사회연구원.
- 김통원. 2000. “미국 공공부문 사회복지서비스 평가체계.” 《사회복지》 통권 제144호, 한국사회복지협의회. pp. 186~194.
- _____. 1998. “사회복지 수용시설 평가제의 도입과 대비전략: Self-Assessment 결과를 중심으로.” 《1998 추계학술대회 자료집》, 한국사회복지학회. pp. 123~144.
- _____. 1997(a). 《사회복지관 평가의 문제점과 개선방안》, 한국복지관협회.
- _____. 1997(b). “사회복지시설의 표준 평가지표 개발에 관한 연구.” 《1997년 춘계학술대회 자료집》, 한국사회복지학회. pp. 127~163.
- 김통원·김혜란. 1998. “무료노인입소시설 소속기관평가.” 《한국노년학》 제18권 3호(통권28호), 한국노년학회. pp. 155~165.
- 박경일. 2000. “복지 QC활동에 관한 사례연구.” 《한국사회복지학》 통권40, 한국사회복지학회. pp. 97~130.
- 변재관 외 10인. 2000. 《사회복지시설 평가지표 개발 및 평가결과》, 한국보건사회연구원.
- 보건복지부. 1999. 《사회복지시설 평가(인증) 추진계획》, 보건복지부.
- 이변송·김진욱·함진식. 1996. 《서울시 사회복지관 단순평가모형 개발》, 서울시정개발연구원.
- 이정호·김성이·김통원. 1999. 《사회복지시설 평가의 길잡이》, 한국사회복지협의회.
- 이청자. 2000. “사회복지시설 평가 실시에 대한 제언.” 《사회복지》 통권 제144호, 한국사회복지협의회. pp. 102~113.
- 최성재. 1998. 《서울시 사회복지시설 운영실태조사 및 사회복지시설 평가지표 개발》, 서울시.
- 최일섭 외 4인. 1998. 《사회복지관 평가모형 개발연구》, 한국사회복지관협회.
- 최재성. 1998. “장애인복지관 평가프로그램 개발.” 《장애인복지관 기능정립 및 평가프로그램 개발연구에 대한 공청회 자료집》, 한국장애인복지시설협회.
- 황성철·강혜규. 1994. 《사회복지관 운영평가 및 모형개발》, 한국보건사회연구원.
- Alberta. 1998. Government Performance Measures/Benchmarking. on World Wide Web: <http://www.financenet.net/financenet/start/topic/perf.htm>
- APQC. 1999. *Establishing Performance Measures*. American Productivity & Quality Center.
- Brinkerhoff, R., & Dressler, D. 1990. *Productivity measurement: A guide for managers and evaluators*. CA: Sage, Newbury Park.
- CARF. 1998. *Behavioral Health Standards Manual*. Commission on Accreditation of

- Rehabilitation Facilities.
- CARF. 1994. *Standards Manual for Organizations Serving People with Disabilities*. Commission on Accreditation of Rehabilitation Facilities.
- Cassam, Emlyn., & Gupta, Himu. 1992. *Quality Assurance for Social Care Agencies*. Longman Group UK Ltd.
- Chief Financial Officers Council GPR Implementation Committee. 1997. *Integrating Performance Measurement into the Budget Process*. Chief Financial Officers Council GPR Implementation Committee.
- Clarke, John, Gewirtz, Sharon., & McLaughlin, Eugene. 2000. *New Managerialism, New Welfare*. Sage with association with The Open University.
- Congress of the United States of America. 1993. Government Performance and Results Act of 1993. on gopher://pula.financenet.gov:70/00/docs/legis/gpra93.gop
- Department of the Treasury Financial Management Service. 1993. *Performance Measurement Guide*. Government Publications Section.
- Dickens, Paul. 1994. *Quality and Excellence in Human Service*. England, Wiley.
- Epstein, P. D., Epstein and Fass Associates, 1992. "Get Ready: The Time for Performance Measurement Is finally coming!" *Public Administration Review*, 52(5). October.
- Ernest, R. House., & Kenneth, R. Howe. 1999. *Values in evaluation and social research*. Calif.: Sage, Pub info Thousand Oaks.
- Gore, A. 1993. *Creating a government that works better and costs less: Report of the national performance review*. Washington, DC, Government Printing Office.
- Governmental Accounting Standards Board of the United States. 1998. Performance Measurement Information Clearinghouse for Governments. on the World Wide Web: <http://www.retgers.edu/Accounting/raw/gasb/seagov/home.htm>
- Government Applications: The Government Performance & Results Act of 1993. Performance Measurement and Outcome Evaluation in Government. on the World Wide Web: <http://mgdsl.ise.vt.edu/sthepvon/measure/govt.htm>
- Gunther, John., & Hawkins, Frank. 1996. *Total Quality Management in Human Service Organizations*. Springer Publishing Company.
- Heisenberg, Werner Karl. 2000. Heisenberg Uncertainty Principle. on the World Wide Web: <http://physics.hallym.ac.kr/education.oregon/imamura/208/jan271hup.html>
- Martin, Lawrence L., & Kettner, Peter M. 1996. *Measuring the Performance of Human Service Programs*. Sage.

- Martin, Lawrence L. 2001. *Financial Management for Human Service Administrators*. Allyn and Bacon.
- _____. 1993. *Total quality management in human service organizations*. CA: Sage, Newbury Park.
- OECD. 1998. In Search of Results: Performance Management Practices. on the World Wide Web: <http://www.oecd.org/puma/management/pac/pubs/pmp97/overview.htm>
- Osborn, D., & Gaebler, T. 1992. *Reinventing government*. MA: Addison-Wesley, Reading.
- PUMA. 1999. Public Management. on the World Wide Web: <http://www.oecd.org/puma/management/pac/index.htm>.
- Rossi, P., & Freeman, H. 1993. *Evaluation: A systematic approach*. CA: Sage, Newbury Park.
- Royse, D., Thyer, E. A., Padgett, D. K. & Logan, T. K. 2001. *Program Evaluation*. Brooks/Cole.
- United States General Accounting Office. 1998. *Performance Measurement and Evaluation: Definitions and Relationships*, GAO.

A Comparison of Social Welfare Evaluation Systems between Korea and the United States

Kim, Tong-Won

(Sung Kyun Kwan University Professor)

It has been 5 years since evaluation systems were introduced to the social welfare field in Korea. More than half of social welfare agencies (617 out of 1,267) were evaluated in the past two years. Most areas of the agencies were the target of an extensive evaluation in 2000. As “the Evaluation Era” has come, problems were raised especially regarding the Accreditation system of the Ministry of Health and Social Welfare. More specifically, the problems were related to operation and management of the evaluation systems including the issue of evaluators’ objectivity and fairness. In addition, evaluation costs have been too high to the government. The evaluation systems of OECD nations including the United States were quite different. These nations have used the Performance Measurement in order to secure the objectivity and the fairness. Although the quantity of information was limited compared to that of the Accreditation system, these nations have taken the lead in the governmental evaluation. In this context, this study compared social welfare evaluation systems between Korea and the United States. It presented the current status and problems of the Accreditation system of the Ministry of Health and Social Welfare, discussed the American Performance Measurement and its limitations, and compared the strengths and weaknesses of the two evaluation systems. In addition, this study suggested the strategies of Korean evaluation systems in the long term as well as in the short term.