

평가방법론 측면에서 본 정부업무평가

이 광희*

〈目 次〉

- I. 서론
- II. 평가방법에 대한 이론적 고찰 및 분석의 틀
- III. 정부업무평가에 대한 방법론적 분석
- IV. 방법론적 측면에서 본 정부업무평가 개선방안
- V. 결론

〈要 著〉

이 글은 평가목적에 따라 구분될 수 있는 지표방식과 설계방식, 평가초점을 따라 나뉘는 과정평가와 결과평가는 평가방법론적 분석틀에 의거하여 우리나라의 정부업무평가 제도를 살펴보고 있다. 정부업무평가는 과정에 초점을 두고 있는데, 정부활동의 총체적 결과를 보여줄 수 없으며 부처의 특성을 반영하지 못하는 등의 문제가 제기되고 있다. 또 점검·평정을 목적으로 한 지표위주의 평가모형을 선택하고 있으나, 부처의 책임성 확보나 평가결과의 학습이라는 측면에서 한계를 나타내고 있다. 따라서 부처의 고유사업에 대해서는 과정과 결과를 연계하면서도 분석을 목적으로 설계모형을 적용하고, 지원업무에 대해서는 과정(특히 집행과 산출)에 초점을 두는 지표방식을 적용하는 것이 효율적이다.

【주제어: 평가방법, 정부업무평가, 과정평가, 결과평가, 정량평가, 정성평가】

I. 서 론

심사분석으로 시작된 정부업무평가제도는 기관평가, 통합정부업무평가 제도를 거치면서 점차 체계화되고 있다. 평가주체도 다양해지고 평가대상도 확대되었으며, 여러 가지로 분화되어 운영되는 평가제도들이 하나로 통합되어가는 추세도 나타나고 있다. 이처럼 우리나라의 평가제도는 다양한 실험들을 통해 제도

* 한국행정연구원 국정평가연구센터 소장(khlee@kipa.re.kr)

적으로 성숙되고 있다.

그러나 이러한 발전에 조응하는 평가방법론의 성숙은 아직 더딘 것으로 생각된다. 통합정부업무평가제도가 출범하면서 계량적 평가와 질적 평가에 대한 혼란된 인식이 나타나고 있다. 자체평가결과를 기관별로 종합 평정하는 과정에서 나타나는 계량화 과정을 계량적 평가방법으로 인식하는 오류가 쉽게 발견된다. 나아가 기관평가는 정성평가방법을 사용하고 현행 통합정부업무평가는 정량평가방법에 의존하고 있다고 하면서, 정성평가 도입을 강조하는 주장도 심심치 않게 제기되고 있다. 이처럼 평가방법론을 둘러싼 논의가 계속되고 있으나 평가방법론에 관한 연구는 많지 않은 편이다.

본 연구는 이러한 문제인식에서 출발하고 있다. 평가방법론에 대한 체계화를 시도하여 평가목적에 적합한 평가방법을 제시하고자 한다. 평가목적과 평가방법의 조응은 평가의 일관성을 유지하고 비용 대비 평가실익을 극대화하는데 기여할 것이기 때문이다. 그리고 평가목적에 적합한 평가방법을 정부업무평가제도에 적용시켜 봄으로써 정부업무평가제도에 대한 개선방안을 모색하고자 한다.

II. 평가방법에 대한 이론적 고찰 및 분석의 틀

평가방법이 무엇이냐에 대해서는 많은 논의가 있어 왔다. 대체로 어떤 자료를 활용하여 어떤 방식으로 평정하느냐에 따라 정량적 평가와 정성적 평가로 구분하는 것으로 이해하고 있는데 이는 평가방법론을 협소하게 해석한 것이라 본다. 평가방법의 의미를 좀 더 넓게 확장한다면 평가의 목적이나 초점, 또는 평가주체 등에 따라 평가유형을 분류하는 것도 포함한다고 해야 할 것이다. 왜냐하면 평가의 초점에 따라 평가의 주요 대상이 달라지기 때문이다. 예를 들어 과정에 초점을 두느냐 결과에 초점을 두느냐 하는 문제는 평가를 어떻게 해야 하는가라는 방법과 긴밀히 연계될 수밖에 없다.

이 글에서는 두 가지 기준에 의한 분류를 살펴보고자 한다. 첫째는 평가목적 또는 성격에 따라 구분되는 지표방식에 의한 평가와 설계모형에 의한 평가이다. 이 분류 하에서 자료수집, 분석과정 및 추론방법에 의해 차이나는 정량평가와 정성평가가 다루어질 것이다. 둘째는 평가의 초점을 어디에 두느냐에 따라 나뉘는 과정평가와 결과평가의 구분이다.

1. 지표방식 vs 설계모형

첫째 분류는 지표방식에 의한 평가와 설계모형에 따른 평가의 구분인데, 이러한 구분은 평가의 목적에 따른 것이라 볼 수 있다. 정부업무평가기본법 제2조

에 따르면 “평가는 일정한 기관·법인 또는 단체가 수행하는 정책·사업·업무 등에 관하여 그 계획의 수립과 집행과정 및 결과 등을 점검, 분석, 평정하는 것”이라고 정의되고 있다. 여기서 점검이라 하면 정책이나 사업이 계획대로 잘 추진되고 있는지를 점검하는 것이고, 분석이라 하면 정책이나 사업의 성과가 있다면 그 원인이 무엇인지를 분석하는 것이며, 평정이라 하면 정책이나 사업의 성과가 우수하거나 그렇지 못한 기관을 갈라보는 것이라 하겠다.

평가의 목적 또는 주요 성격이 점검이나 평정에 있다면 점검해야 할 사항이나 평정요소 등을 지표화하는 방식을 취해야 할 것이다. 반면에 평가의 목적이 정책(또는 사업) 성과의 인과적 분석에 있다면 독립변수와 종속변수 선정 등 인과적 분석에 합당한 설계모형을 개발해야 할 것이다. 지표에 의한 방식은 평가의 목적이 주로 점검이나 평정에 있으므로 조직 관리에 적합하다고 볼 수 있다. 비전-임무-전략-성과목표-성과지표로 이어지는 성과관리는 이러한 전형을 잘 보여준다. 설계에 의한 평가방식은 개별 사업에 대한 평가방식으로 매우 유용하다. 예를 들어 우유급식프로그램, 기초생활대상자 및 차상위계층에 대한 자활사업 등 다양한 사업을 실시하고 그 성과를 분석하는데 많이 활용된다.

먼저 설계모형에 의한 방식부터 살펴보자. 설계모형에 의한 평가는 전통적인 사회과학 방법론에 입각해 있으며, 실험설계 등을 활용하는 정량적 방법과 사례분석 등을 활용하는 정성적 방법으로 나눌 수 있다. 계량적 방법이거나 질적 방법이나의 문제는 사회과학 방법론에서도 오래된 논쟁거리 중의 하나라고 할 수 있다. 계량적 연구는 구체적 현상에 대한 정량적 측정에 근거하여 특정 사례들을 일반화시키거나 인과적 가설을 검증하는데 목적을 두고 있다. 그리고 다른 연구자들에 의해서도 쉽게 반복 시행할 수 있는 측정과 분석 방법을 사용하고자 한다. 이와 달리 질적 연구는 폭넓은 접근방법을 활용하지만 계량적인 측정이나 통계분석 등에 의존하지 않으며, 사례분석, 심층면접, 역사적 맥락 분석 등을 활용하고 있다(King, Keohane, & Verba 1994: 3~4).

평가영역에서는 양적 방법론이 먼저 발달하였다고 볼 수 있다. 대규모 사회적 공작이 진행되었던 “위대한 사회(The Great Society)” 시대는 프로그램 평가의 필요성을 촉발시켰고 당시 사회과학의 주류 패러다임이었던 행태주의 방법론이 평가영역에 본격적으로 도입되었다. Rossi and Freeman(1985: 34)에 따르면 프로그램이 미친 사회적 영향(impact)을 평가하기 위하여 실험(experiment) 방법을 적용한 D. Campbell(1969)의 논의가 양적 평가방법의 효시라고 지적하고 있다. 1930년대 이래 자연과학에서 도입되어 사회과학을 지배해왔던 실험적 평가모형들은 계량적이고 객관적인 측정을 통해 관측 가능한 현상에 초점을 맞추었다.

1980년대에 행태주의에 대한 비판이 가해지면서 평가영역에서도 양적인 평가

방법에 대한 공격이 가해졌다. 과학의 기본적인 불확실성과 실험방법의 불충분성으로 인해 사회현상의 복잡성을 파악할 수 없다는 것이다. 사회과학 지식이 프로그램을 설계하는데 부적절했으며, 과학적 방법이 프로그램의 효과성(effectiveness)을 평가하는데 한계를 보였다는 것이다. 그리하여 계량화(quantification)보다는 이해(understanding)를 중시하는 질적 방법이 부각되었다(Cook 1985: 22-29). 이러한 질적 방법론에는 참여관찰, 심층면접 등이 제시되고 있다.

두 가지 평가방법론의 차이를 비교해보면 <표 1>과 같다. 정량적 평가방법은 집합자료와 서베이 자료 등 계량화된 자료를 이용하여 정책이나 프로그램의 효과를 실험 등을 통해 파악함으로써 정책이나 프로그램에 내재해있는 이론 또는 정책의 가치를 검증하는 방법론이라 하겠다. 이와 달리 정성적 평가방법은 심층면접이나 사례분석 등을 통해 정책이나 사업이 가져오는 의도된 효과뿐만 아니라 의도하지 않은 효과까지 파악함으로써 이론의 빌굴 또는 정책 가치의 발견에 기여하는 방법론이라 할 수 있다. 한편 최근에는 정량적 평가방법과 정성적 평가방법을 혼합하여 사용할 데 대한 요구가 높아지고 있다(Greene et al 2001). 자료 수집에 있어서 계량적인 자료뿐만 아니라 질적인 자료까지 포함하고, 실험 설계 외에 참여관찰 등을 수행하는 등 정량적 평가방법과 정성적 평가방법을 혼합한 평가들이 매우 유용하다고 지적되고 있다.

<표 1> 정량적 평가방법과 정성적 평가방법 비교

	정량적 평가방법	정성적 평가방법
자료 수집	계량화된 자료 (집합자료, 서베이자료)	질적 자료 (면접자료, 관찰자료 등)
분석 과정	실험 (진실험, 준실험 등)	심층면접, 사례분석 등
추론 방법	연역적 (가설 검증)	귀납적 (이론 발견)

정량평가냐 정성평가냐와 관련하여 김현구(2003)는 과학적 연구와 평가적 추론으로 구분하고 있다(<표 2> 참조). 과학적 연구에 입각한 평가를 분석적 평가라 하고 평가적 추론에 입각한 평가를 판단평가로 구분하고 있다. 과학적 연구라 함은 실험 등 정량적 평가방법에서 활용되는 패러다임을 뜻한다. 반면 평가적 추론에는 질적 방법을 포함한 다양한 접근법이 포함되고 있다. 여기서 분석적 평가는 정책에 내재해 있는 가치를 검증한다는 점에서 연역적이며 과학적

방법을 활용한다는 점에서 정량적 평가방법으로 볼 수 있다. 판단 평가는 다양한 요인들을 고려하여 정책의 가치를 판단한다는 점에서 귀납적 방식을 취하고 있으며, 질적 방법 등의 활용을 명시하고 있어 정성적 평가 패러다임으로 분류할 수 있을 것이다.

〈표 2〉 분석적 평가와 판단 평가 비교

	목적(기능)	방법
분석적 평가	정책의 가치 입증	과학적 분석
판단 평가	정책의 가치 판단	질적 방법, 기술통계, 전문가적 판단, 정치논리·행정절차·법적규범 등 다양한 요인 고려

* 출처: 김현구(2003: 60)에서 재작성

지표방식에 의한 평가는 평가의 효율성을 높이는 장점이 있다. 지표를 개발하여 적용하므로 평가모형을 설계하거나 과학적인 인과관계를 모색하는 번거로움이 없으며, 평가수행에 필요한 인력이나 비용을 최소화시킬 수 있다. 또한 정책에 대한 평가보다 기관에 대한 평가를 할 경우, 하나의 지표를 많은 평가대상에 공통적으로 적용하므로 기관의 업무 수행에 대한 점검 및 평정에 편리하다 (김주환 2003)

지표방식에 의한 평가를 할 경우 지표가 계량화되었느냐 아니냐에 따라 정량지표와 정성지표로 구분하기도 한다. 예컨대 지표에 의한 평가결과가 3, 2, 1 등 수치로 표현되는 경우 계량화된 것으로 보는 반면 평가결과가 우수, 보통, 미흡 등의 방식으로 나타나면 정성지표로 이해하는 것이다. 그러나 이런 지표에 의한 구분을 정량평가와 정성평가로 그대로 적용시키는 것은 잘못이다. 정량평가는 계량화된 자료를 많이 수집하지만 실험설계 등 과학적 분석방법을 활용하여 정책(또는 사업)에 내재해 있는 이론을 연역적으로 추론한다는 점이 중요하다. 지표가 계량화되었든 그렇지 않든 간에 지표방식은 점검이나 평정의 성격이 강한 평가에서 많이 활용되지만, 실험설계 등 과학적 분석이 요구되는 평가에서는 적용하기 어렵다. 즉 지표가 계량적인 것과 설계모형의 정량적 방법이 정량적인 것과는 차원이 다른 문제이다.

결론적으로 지표방식과 설계방식의 분류 및 정량평가와 정성평가의 구분을 혼동하지 말아야 한다. 점검 및 평정의 성격을 지닌 평가에서는 지표방식이 효과적이고, 분석이 필요한 평가에서는 설계모형에 의해야 할 것이다. 지표방식에서 지표가 정량적이거나 정성적이냐에 따라 구분될 수 있지만, 이러한 구분이 설

계방식에서 연역적 성격을 지니는 정량적 평가방법과 귀납적 성격을 지니는 정성적 평가방법의 분류와는 차원이 다르다는 점에 주의해야 한다.

2. 과정평가 vs 결과평가

다음으로 평가의 초점을 어디에 두느냐에 따라 과정평가와 결과평가로 나누어볼 수 있다. 조직 관리를 연구한 학자들에 의하면 정부관리(government management)의 방법은 크게 과정 지향적 관리와 결과 지향적 관리로 나뉜다 (Halachmi 1982; Wholey 1983; Halachmi & Bouckaert 1996). 이에 따라 조직의 성과를 평가하는 방식도 과정에 초점을 두느냐 결과에 초점을 두느냐로 분류해 볼 수 있다.

이와 관련하여 국내외 평가학자들의 유형분류를 살펴보자. 예컨대 Scriven에 따르면 형성평가(formative evaluation)와 총괄평가(summative evaluation), Suchman에 따르면 노력평가, 효과평가, 성과의 충분성 평가, 능률성평가, 과정평가 등으로 나뉘며, Nachmias는 과정평가(process evaluation)와 영향평가(impact evaluation)로 구분한다(김명수 2000: 101-104). 또 Rossi & Freeman은 프로그램 개념화와 디자인 평가, 집행 모니터링, 프로그램 유용성/utility에 대한 평가로 나누고 (Rossi & Freeman 1985: 38-46), 이진주 외(1996)는 정책분석, 과정평가, 총괄평가로 분류하고 있다. 이러한 분류들은 약간씩 차이가 나지만, 대체로 형성-집행-산출-결과라는 정책 사이클 중 어디에 초점을 두느냐에 따라 평가유형을 구분한다는 점에서 공통점을 가진다.

이처럼 부분적으로 차이를 보이는 유형화를 요약한다면 과정에 초점을 두는 과정평가와 결과에 중점을 두는 결과평가는 나눌 수 있을 것이다. 과정평가는 정책이나 사업이 이미 수립된 계획에 따라 집행된 정도를 평가하는 집행모니터링으로 협소하게 정의할 수 있으나 정책의 형성에서부터 산출에 이르기까지의 모든 과정을 살펴보는 평가로도 확대하여 이해할 수 있다. 그러나 과정평가를 산출까지 포함할 정도로 넓게 해석한다 해도 결과까지 포괄하기는 힘들 것이다. 이에 반해 결과에 초점을 두는 평가는 정책이 형성되고 집행되는 과정보다는 그 산출물이 사회적으로 어떤 영향을 가져왔는지, 어떤 효과가 있었는지를 평가하는 것이라 하겠다.

<표 3> Chen의 평가 유형 분류

		평가 기능 (Evaluation functions)	
		개선(Improvement)	평정(Assessment)
사업 단계	과정	과정개선에 초점을 둔 평가 <i>Process-improvement evaluation</i>	과정 평정에 초점을 둔 평가 <i>Process-assessment evaluation</i>
	결과	결과 개선에 초점을 둔 평가 <i>Outcome-improvement evaluation</i>	결과 평정에 초점을 둔 평가 <i>Outcome-assessment evaluation</i>

* 출처: Chen(1996: 123)

과정평가와 결과평가로 구분하는 것은 첸(Chen 1996)의 연구에서도 지적된 바 있다(<표 3> 참조). 그는 형성평가(formative evaluation)와 총괄평가(summative evaluation)로 구분하고 있는 스크리븐(Scriven)의 분류를 제한적이라고 비판하면서, 보다 포괄적인 분류 모형으로 평가의 기능과 평가의 초점을 두 축으로 하는 4가지 모형을 제시하고 있다. 그러나 첸이 제시한 4가지 평가유형이 베타적으로 구분될 수 있는지는 명확하지 않다. 결과를 개선하는데 초점을 둔 평가(Outcome-improvement evaluation)는 과정평가(Process evaluation)와 연관될 수밖에 없는데 실제 평가과정에서 양자 간의 분리가 가능한지 의문이다. 이런 점에서 보면 과정 개선에 초점을 둔 평가(Process-improvement evaluation)와 결과 평정에 초점을 둔 평가(Outcome-assessment evaluation)로 유형화하는 것이 타당하지 않을까 생각된다.

과정평가는 2차 대전 이후 서구 사회가 높은 경제적 성장을 누리고 국가가 다양한 사회복지 서비스를 제공하였던 “큰 국가”的 패러다임에 적합한 평가였다. “많을수록 좋다”, “대량생산과 대량소비” 등 투입이 많을수록 산출이 증대할 것이라는 신념하에 경제적·사회적 시스템의 비용에 대한 깊은 고려없이 예산증대, 인력증원, 그리고 서비스 확대가 이루어졌다. 이와 같은 공급자 중심의 행정에서는 투입 및 과정에 대한 통제가 중시될 수밖에 없다. 따라서 이러한 관점에 입각한 과정평가는 가치중립적이며 목표달성을 초점을 맞추게 된다.

한편 국가실패 현상을 지적하며 등장한 신공공관리(NPM)는 공공부문의 효율성에 대한 새로운 관심을 촉발시켰다. 또한 대량생산경제시스템에서 포스트포디즘으로의 전환이 이루어지면서 삶의 질 문제 등이 부각되었다. 따라서 단순히 양적 증가에 초점을 맞추었던 공급자 위주의 행정에서 서비스의 품질에 대한 관심이 중요해지는 수요자 중심의 행정으로의 변화가 진행되었다. 결과에 대한 강조는 앤고어 당시 대통령의 NPR(Nation Performance Review)과 오스본과 개블

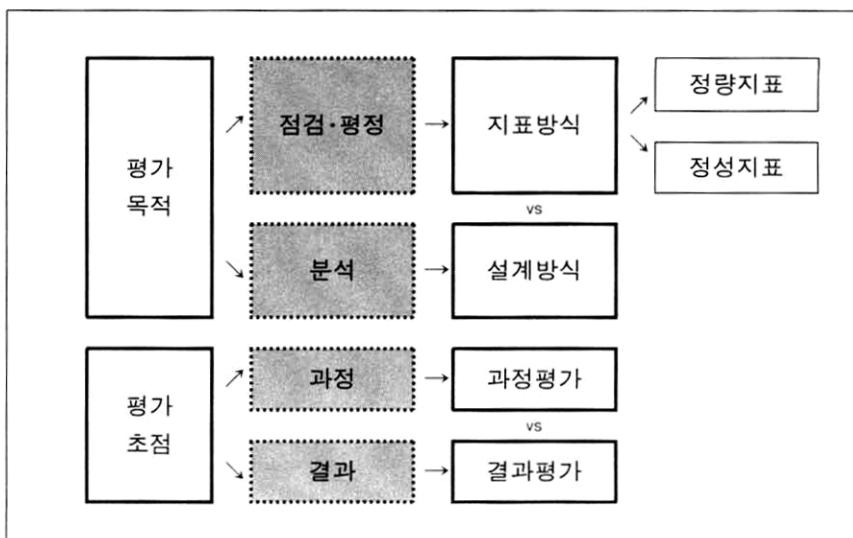
러의 정부재창조(Reinventing Government)에서 정점을 이룬다. 이와 같은 결과에 초점을 둔 평가는 과정평가와는 달리 목표달성을 위주에서 벗어나 가치가 개입된 평가를 지향하게 된다.

과정에 초점을 둘 것이냐 아니면 결과에 초점을 둘 것인가의 문제는 평가의 목적과 긴밀히 연관되어 있다. 자동차 사고에서 운전자가 생존할 수 있는 자동차의 능력을 측정하는 것은 차량의 연비가 효율적인지 아닌지를 결정하는데 도움을 주지는 않을 것이다. 즉 성과를 측정하고 평가하려는 목적에 따라 과정에서의 효율성 및 적절성에 초점을 둘 것인지, 과정이야 어떻든 궁극적인 효과나 결과에 초점을 둘 것인지를 결정해야 한다(Halachmi & Bouckaert 1996; 4~5).

3. 분석 틀

지금까지 평가방법에 대한 이론적 고찰을 크게 두 가지로 나누어 살펴보았다. 평가목적에 조응하는 지표방식과 설계방식, 평가의 초점에 어디에 두느냐에 따른 과정평가와 결과평가의 구분이다. 이러한 논의를 토대로 정부업무평가에 대한 분석 틀을 제시해보면 <그림 1>과 같다. 우선 평가목적을 점검·평정과 분석으로 구분하여 살펴보고 이러한 목적에 맞게 지표방식 또는 설계방식으로 되어 있는지를 검토한다. 그리고 지표방식의 경우 정량지표인지 정성지표인지 살펴본다. 다음으로 평가의 초점을 어디에 두느냐에 따라 과정평가인지 결과평가인지 구분하도록 한다.

<그림 1> 분석 틀



III. 정부업무평가에 대한 방법론적 분석

우리나라의 평가 제도를 역사적으로 개괄해보면 <표 4>와 같다. 즉 1960년대에 시작된 심사분석제도를 평가의 효시라고 볼 수 있으며, 그 후 경제기획원의 심사분석 및 행정조정실의 정책평가와 심사평가, 그리고 국무조정실이 주관한 정부업무평가 및 통합국정평가체계로 이어진다. 제도명칭이나 주관부처에서 약간의 변동이 있었는데, 평가대상이나 평가내용의 기준에서 보면 크게 4단계로 구분할 수 있다. 즉 심사분석-정책평가-기관평가-통합정부업무평가라고 하겠다.

<표 4> 우리나라 정부업무평가제도 개괄

시기	1961-1981	1981-1994	1990-1994	1994-1998	1998-2005	2006-
제도	심사분석	심사분석	정책평가	심사평가	정부업무 평가 (기관평가)	통합정부 업무평가
주관 부처	기획조정실	경제기획원	행정조정실	행정조정실	국무조정실	국무조정실
4단계 구분	심사분석		정책평가		기관평가	통합정부 업무평가

한편 평가대상이나 평가내용의 기준으로 하여 4단계로 나눌 때 각 단계별 평가방법론상의 특징은 <표 5>가 잘 보여준다. 각 단계별로 평가대상이 점점 확대되어 오고 있으나, 평가방법론상의 초점은 여전히 과정 지향적이라 할 수 있다. 그리고 평가목적 또는 성격에 의하면 인과적 분석보다는 점검·평정에 중점을 두는 지표방식에 의한 평가방식을 고수하고 있다.

<표 5> 평가제도의 평가대상 및 평가방법

제도	심사분석	정책평가	기관평가	통합정부업무평가
평가 대상	주요사업 (부처) 중점사업	주요국정 중심과제	주요정책 기관역량 만족도	자체평가 (주요정책, 재정사업, 조직, 인사, 정보화) 특정평가 (관리대상 업무, 특정시책, 고객만족도)
평가 방 법	지표방식 (점검) 정성지표	지표방식 (점검) 정성지표	지표방식 (점검·평정) 정성지표	지표방식 (점검·평정) 정량지표
	초점	과정	과정	과정

심사분석 단계에서는 과정에 초점을 둔 전형적인 집행평가의 성격을 가진다. 부처가 수행하고 있는 주요 정책 또는 사업별로 질적인 자료와 양적인 자료에 의존하여 문제점이나 성과를 파악해나가는 귀납적 방법을 부분적으로 사용하고 있다. 하지만 이러한 집행평가는 각 부처의 주요 정책이나 사업이 계획된 대로 추진되고 있는지를 점검하는 성격이 강하여 지표방식에 의한 평가가 이루어지고 있다.

정책평가 단계에서는 정책이나 사업의 효과성 등을 파악하고자 한 점에서는 진도분석 위주의 심사분석과 구분되고 귀납적 성격이 강화된 측면이 있지만, 여전히 과정에 초점을 두고 있으며 지표방식을 고수하고 있다.

기관평가 단계에서도 정책의 효과성이 부분적으로 강조되고 있지만 계획-집행-성과라는 과정 중심적 평가방식이 지속되고 있다. 또한 만족도 조사 등 양적 자료 수집이 부분적으로 체계화되었으나 기관별 평정이라는 성격이 강화되어 지표방식의 평가방법이 계속 사용되고 있다.

통합정부업무평가단계에서는 평가대상이 상당히 확대되지만 계획-집행-산출/결과-활용 등 단계별로 평가하는 방식은 여전하다. 질적 자료뿐만 아니라 서베이자료, 집합자료 등 계량화된 자료를 많이 활용하고 있으나, 정책이나 사업, 또는 업무별로 단계에 따른 과정평가 및 지표방식에 의한 평가방법에 여전히 의존하고 있다. 또한 목표달성을 점검이 강조되면서 정책평가 및 기관평가에서 존재하였던 귀납적 성격이 오히려 약화되었다. 다만 정성지표 위주였던 기관평가 단계에 비해 정량지표로 바뀐 점이 큰 차이점이라고 할 수 있다.

이상의 결과를 토대로 정부업무평가의 방법론에 대한 문제점을 지적해보면 크게 두 가지로 제시할 수 있다. 하나는 과정평가의 편향이며, 또 하나는 정량평가에 대한 오해이다.

1. 과정평가의 편향과 한계

심사분석-정책평가-기관평가-통합정부업무평가로 이어져온 정부업무평가는 종양행정기관을 평가대상으로 하여 부처의 주요 정책 및 사업에 대한 평가에 중점을 두어 왔으며, 과정에 초점을 맞춘 정성적 평가방식을 취하여 왔다. 예컨대 계획-집행-성과-환류 단계에 따라 각 단계의 적절성, 효율성 등을 파악하고자 한다. 평가지표가 계량화되었다고 하는 통합정부업무평가의 자체평가와 특정평가도 근본적으로 이러한 틀에 입각하고 있다.

그러나 과정에 초점을 둔 평가는 정부업무의 기술적 측면에 대한 평가로 치우치게 되고 나아가 가치중립적인 평가의 경향이 강하게 나타나게 된다. 즉 평가를 잘 받기 위한 전략적 행위로 인해, 달성하기 어려운 목표는 배제되거나 설정된 목표 수준이 하향 조정되는 등의 문제가 나타나게 된다. 또한 계획-집행-

산출로 이어지는 단계에 초점을 맞추다 보니 정부활동의 총체적 결과를 잘 보여줄 수 없다는 점도 문제로 제기되고 있다.

예컨대 통합정부업무평가의 자체평가에서 재정경제부가 90점, 외교통상부가 70점을 받았다면 이러한 점수 차이를 어떻게 해석해야 할까? 이러한 점수에 근거하여 한국정부의 경제정책은 우수하고 외교능력은 그렇지 못하다고 해석하는 것이 타당할 것인가? 국가의 경제발전을 추진하고 외교력을 향상시키기 위한 정부의 노력에 대한 평가와 그러한 노력의 결과 실제 나타난 성과에 대한 평가는 다른 차원에서 이루어져야 할 것이다. 각 부처가 추진하는 정책이나 사업에 대해 일상적으로 점검하고 평가해야 할 과정적 요소가 존재하므로 과정평가의 필요성은 충분히 인정이 된다. 그러나 과정에 초점을 둔 평가가 결과의 측면에서 부처의 우수성을 파악할 수 있는 평가모형까지 담당할 수는 없다.

또한 과정평가는 피평가기관의 수용도에 부정적인 영향을 주고 있다. 과정에 초점을 둔 통합정부업무평가체계는 모든 부처에 공통적으로 적용할 수 있는 평가지표를 활용하고 있기 때문에, 부처의 특수성을 강조하는 피평가기관들이 불만을 제기하고 있다. 예컨대 통일부, 외교통상부, 국방부 등 애초 계획 수립과정에서 예상했던 것과는 다른 돌발 상황에 의하여 정책이나 사업이 수시로 변경되는 부처의 경우, 과정지표에 의한 평가방식에 의하면 낮은 점수를 받을 가능성이 커진다. 또 이러한 과정지표에 의한 점수가 정부의 외교능력, 국방능력에 대하여는 어떤 판단의 근거도 제시하지 못하고 있다. 이러한 점들로 인해 피평가기관들은 자신들의 성과에 대한 평가가 제대로 이루어지지 못하고 있다는 불만을 가지고 있다.

〈표 6〉 기관평가와 통합정부업무평가의 영역 비교

기관평가 (2004)	주요 정책	기관역량		고객만족도
	각 부처별 주요 정책과제	혁신추진, 인사행정 효율성 조직운영 효율성, 정보화, 정보공개, 국정홍보, 공직기강 확립		주요정책과제, 민원행정서비스, 기관행정이용자
통합정부 업무평가 (2006)	자체평가		특정평가	
	주요정책 재정성과	인사, 조직, 정보화	혁신관리, 정보공개, 정책홍보관리, 대통령지시사항, 규제개혁, 청렴도, 법적 의무 · 권장사항 및 법제업무	주요정책, 민원행정서비스

* 출처: 국무조정실(2004; 2006a)

그리고 계획·집행·산출·결과·환류라는 단계별 평가 방식을 인사나 조직, 정보 공개나 홍보 등의 부문에까지 모두 적용하는 것이 타당한가라는 의문이 제기되고 있다. <표 6>이 보여주는 것처럼 기관평가와 정부업무평가 영역을 비교해보면 두 평가 모형은 상당히 유사하다. 기관평가 모형에서의 주요정책과제 평가는 자체평가의 주요정책 및 재정성과 평가로 볼 수 있으며, 기관역량평가는 자체평가의 조직, 인사, 정보화, 특정평가의 혁신관리, 정책홍보관리, 정보공개 등의 평가와 유사하다. 기관평가의 이론적 토대를 제시하고자 하는 김현구(2003)는 기관평가 모형을 정책총합모형으로 제시하고 있다. 그에 따르면 조직론적 개념 틀과 정책학적 실행논리를 통합하여 정책총합모형을 제시하고 있다. 여기서 정책은 기관의 고유 업무에 관한 실질정책(substantive policy)뿐만 아니라 이를 지원하는 관리정책(management policy)까지 포함하는 종합적 개념으로 제시된다 (2003: 67). 즉 기관역량을 뜻하는 다양한 요소들도 관리정책으로 범주화하여 정책평가와 기관평가를 등치시키는 구도라 하겠다. 이러한 관점에서 본다면 통합 정부업무평가도 기본적으로 정책총합모형으로 볼 수 있다.

그런데 이러한 정책총합모형에 의해 인사와 조직이 자체평가 부문에 포함되었다고 볼 수 있지만, 실제 소요된 평가노력에 비해 평가실익이 크지 않다. 즉 두 부문의 경우 정례적으로 년 단위의 평가를 통해 모든 부처가 인사 및 조직 관리 전반을 계획, 집행, 산출/결과 및 평가결과 활용도 등으로 구분하도록 되어 있는데, 이는 곧 매년 동일한 평가내용이 다루어질 수 있다(공병천 2007). 또 인사나 조직 관리의 경우 자체의 정책순환 단계에 따른 성과 도출보다는 인사와 조직 부문에서의 역량 개선이 실제 고유 사업에서의 성과 개선으로 나타나는 것이 바람직하다(행정자치부·한국행정연구원 2007: 119). 따라서 인사나 조직 관리의 경우 점검·평정의 성격을 분명히 하여, 계획 대비 실적 위주로 평가하는 방식이 효율적이라고 볼 수 있다.

2. 정량평가에 대한 오해

통합정부업무평가의 지표체계가 계량화된 방식으로 이루어져 있다고 하여 이를 정량적 평가방법이라고 주장하는 것이 타당한가? 앞에서 살펴본 대로 정량적 평가방법이 가지는 근본적인 관점은 정량적인 자료를 활용하지만 분석 및 추론의 과정이 정책의 가치 또는 정책에 내재해 있는 이론을 검증하는데 초점을 두는 연역적 방식을 취하는데 있다. 그런데 통합정부업무평가가 이러한 관점을 가지고 있다고 하기는 힘들다. 왜 그런지 구체적으로 살펴보도록 하자.

2006년도 정부업무평가 지표에 따르면, 자체평가가 50점, 특정평가가 50점으로 구성되어 있다(국무조정실 2006). 자체평가의 경우 주요정책 25점, 재정성과 8점, 인사 5점, 조직 5점, 정보화 7점으로 구성되었다. 그리고 주요정책평가의

경우 계획단계에서 계획수립의 적절성이 11점, 성과계획의 적정성이 11점으로 구성되고, 집행단계에서 시행과정의 효율성이 15점, 시행과정의 적절성이 11점으로 구성되며, 산출/결과단계에서 목표달성을 40점, 그리고 평가결과 활용이 15점으로 구성되고 있다(<표 7> 참조).

<표 7> 주요정책과제 평가지표(2006년도)

평가항목	평가지표	배점
계획 수립의 적절성	계획수립 시 사전조사, 의견수렴 등을 충실히 이행하였는가	7
	정책목표 달성을 위한 정책수단은 적절하게 구비하였는가	4
성과 계획의 적정성	성과목표 및 성과지표는 적정하게 설정하였는가	6
	성과목표치 및 검증방법은 적정한가	5
시행 과정의 효율성	과제는 일정계획에 맞추어 추진하였는가	7
	과제추진을 위한 자원은 효율적으로 집행하였는가	5
시행 과정의 적절성	시행과정에서 행정여건·상황변화에 적절히 대응하였는가	7
	시행과정에서 관련기관·정책과의 연계 및 협조체제를 구축하여 운영하였는가	4
목표 달성도	당초 설정한 목표는 달성하였는가	40
평가결과 활용도	평가결과를 공개하고 학습에 활용하였는가	5
	평가결과를 개인성과 및 차기계획에 반영하였는가	10

* 출처: 국무조정실(2006a)

이처럼 통합정부업무평가지표는 정량화되었지만 평가방법이 정량적인 것이 아니다. 통합정부업무평가의 성격이 점검 및 평정이라는 기관평가의 틀을 벗어나지 못하고 있다는 점에서 지표방식에 의한 평가방법을 계속 고수하고 있지만, 기관평가와 다른 점은 기관별 평정의 방식이 정성지표에서 정량지표로 바뀐 것이다. 즉 기관별 종합평정 과정이 보다 체계적으로 정량화되었다.

자체평가의 세부평가지침에 따르면 주요정책이나 재정성과, 조직 및 인사관리 업무 등에 대하여 자체평가위원들이 정성적으로 평가¹⁾하는 방식을 취하고 있다. 예를 들어 <표 8>은 2006년 자체평가영역의 주요정책과제부문의 계획수립의 적절성 평가항목 중 계획내용의 충실성을 평가하는 지표에 대한 세부착안사항을 보여주고 있다.

1) 물론 이 때의 평가도 부처의 과제들의 수행 과정을 점검하고 그 결과를 평정하는 성격을 지닌다.

〈표 8〉 주요정책평가 세부착안사항

평가지표	측정방법	측정기준	배점
계획수립 시 사전 조사, 의견 수렴 등을 충실히 이행하였는가 (7)	<p>▶ 과제별 측정방법 〈평가요소〉 ④ 의견수렴절차 이행(이해관계자, 전문가 등) ⑤ 관련 통계현황 및 사례에 대한 조사 실시 ⑥ 환경, 갈등, 성차별 등 관련 영향분석 실시 ⑦ 관련 정책과의 연계·협조방안 등을 사전협의</p> <p>* 위 요건 중 해당사항이 없는 경우 이를 소명 하는 경우에만 요건충족으로 간주</p>	<p>① ④⑤⑥⑦의 기준을 모두 충족 ② ④의 기준을 충족하고, ⑤⑥⑦의 기준 중 2개항만 충족 ③ ④의 기준을 충족하고, ⑤⑥⑦의 기준 중 1개항만 충족 ④ ④의 기준만 충족 또는 ⑤⑥⑦의 기준만 충족 ⑤ ⑤⑥⑦의 기준을 2개항만 충족 ⑥ ⑤⑥⑦의 기준을 1개항만 충족 ⑦ ④⑤⑥⑦의 기준을 모두 미충족</p>	① 7 ② 6 ③ 5 ④ 4 ⑤ 3 ⑥ 2 ⑦ 0

④ 의견수렴절차 이행(이해관계자, 전문가 등)

- 의견수렴 방법은 E-mail, 공청회, 토론회, 간담회, 협의회, 설문조사 등 넓게 인정
- 증빙자료로 의견수렴을 위해 작성한 공문 및 의견수렴결과를 정리한 문서 제시

⑤ 관련 통계현황 및 사례에 대한 조사 실시

- 통계현황이나 사례 조사를 실시하고 조사결과에 대한 명확한 근거자료가 있는 경우 인정 (통계 현황이나 사례조사에 대한 근거자료 제시할 경우 반드시 출처 등 명시)

⑥ 환경, 갈등 등 관련 영향분석 실시

- 환경·교통·재해 등에 관한 영향평가법 등 법, 시행령, 시행규칙과 같은 법령에 규정된 영향 분석의 의무가 있는 경우 뿐만 아니라, 법에 의무가 없어도 정책 환경이나 정책효과 등에 대한 분석을 실시한 경우 인정

⑦ 관련 정책과의 연계·협조방안 등을 사전협의

- 관련 정책은 부처내부와 부처외부(관계부처 등)를 포함한 넓은 개념으로 봄
- 다만, 동 사업과 정책의 관련성이 인정되는 경우를 말함
- 아울러, 명백히 관련되는 정책이 있는 경우에도 연계·협조방안을 강구하지 않은 경우에는 동 평가요소를 이행하지 않은 것으로 봄

* 의견수렴 등을 단위과제별로 실시한 경우 이행과제를 대표할 수 있는 단위과제를 분석했거나, 뚜렷한 대표과제가 없는 경우 50%이상의 단위과제에 대해 의견수렴 혹은 환경 분석 등을 실시하면 이행과제 전체가 요건을 충족한 것으로 인정

* 출처: 국무조정실(2006b)

이 지표는 계획 수립 시 사전조사, 의견수렴 등을 충실히 이행하였는가를 측정하는 지표인데, 자체평가 대상과제별로 의견수렴절차 이행여부, 관련 통계현황 및 사례에 대한 조사 실시 여부, 환경·갈등·성차별 등 관련 영향 분석 실시 여부, 관련 정책과의 연계·협조 방안 등의 사전 협의 여부 등을 판단하여 점수를 부여하도록 되어 있다. 그리고 과제별 평균점수가 해당 부처의 계획내용 충

실성 평가지표 점수가 된다. 이처럼 해당 부처의 지표별 점수는 정량화된 방식으로 부여되지만, 실제 점수를 부여하기 위한 과제별 평가 작업은 자체평가위원들이 부처가 제시한 다양한 자료들을 토대로 점검·평정하는 수준에 그치고 있다.

다시 말하면 자체평가위원들은 부처의 주요 정책을 평가할 때, 정책의 효과성을 판단하고 이론 검증을 위한 인과적 추론이나 정책의 가치를 입증하는 설계모형의 평가방식을 사용하고 있지 않다. 더욱이 기관평가 모형에서 다양한 질적 자료들로부터 정책의 의도한 효과(및 의도하지 않은 효과)를 파악하려 했던 귀납적 성격이 통합정부업무평가에서는 오히려 약해졌다. 단계별로 분절화된 평가가 이루어짐으로써 하나의 정책 또는 사업별로 분석과 추론이 이루어지지 못하고 있으며 결국 기관별로 정책단계의 평정점수만을 계량화된 방식으로 보여주고 있을 뿐이다. 이처럼 현재의 통합정부업무평가는 방법론 측면에서 과정에 초점을 둔 지표방식을 계속 고수하고 있다.

물론 통합정부업무평가의 지표가 계량적이라고 해서 잘못되었다는 주장은 아니다. 점검·평정의 성격이 강한 평가에서 지표방식의 평가가 설계모형에 입각한 평가보다 더 효율적일 수 있으며, 이 경우 정성지표에 비해 정량지표가 보다 간편하고 세련된 체계를 갖출 수 있다. 다만 정량지표를 사용한다고 해서 점검·평정의 성격이 강한 평가가 인과적 분석을 목적으로 하는 평가로 등치되는 것은 아니라는 점을 인식해야 한다. 즉 점검·평정의 성격을 지닌 정부업무평가에서 정성적인 지표체계가 정량적인 지표체계로 바뀌었다고 해서 평가방법론의 전환이 이루어졌다고 주장하는 것은 곤란하다.

IV. 방법론 측면에서의 정부업무평가 개선방안

평가목적에 따른 지표방식과 설계모형의 구분, 평가초점에 따른 과정평가와 결과평가의 분류라는 틀을 이용하여 정부업무평가에 대한 분석을 바탕으로, 방법론 측면에서의 정부업무평가 개선방안을 모색해볼 필요가 있다.

먼저 평가대상에 따라 평가방법을 달리 적용할 필요가 있다. 앞서 살펴보았듯이 정부업무평가의 대상은 크게 각 부처의 고유사업과 지원 업무로 나누어볼 수 있다. 부처의 고유 사업에 대한 평가의 목적은 점검·평정보다는 사업의 전반적인 추진 과정 및 그 결과에 대한 분석에 초점을 맞추어야 한다. 그 이유는 이러한 분석이 다음의 두 가지 측면과 관련되기 때문이다. 하나는 정책이나 사업의 성과가 무엇인지를 규명하여 공개함으로써 해당 부처의 책임성을 명확히 하는 것이다. 다른 하나는 성과에 대한 인과적 분석을 통해 새로운 학습의 기회를

가진다는 점이다. 따라서 분석을 목적으로 하는 평가는 연역적이든 귀납적이든 설계 방식에 입각한 평가를 수행하는 것이 바람직하다. 현재와 같은 지표방식에 의한 평가는 기관별 평정을 목적으로 하는 평가에서는 효율적이지만, 평가를 통해 생성된 정보나 결과가 책임성과 학습이라는 측면에서는 별로 도움이 되지 않고 있다. 자체평가위원회의 자체평가, 정부업무평가위원회의 확인점검 등 많은 인력과 비용을 들이며 생성된 정보가 부처에 대한 종합평정을 위한 점수 도출의 근거로만 사용된다면 안타까운 일이 아닐 수 없다.

관리정책 또는 관리대상업무로도 불리며 기관역량에 속하는 지원업무의 경우 분석적 성격보다는 점검·평정의 성격이 강하다. 이러한 업무는 고유 사업처럼 부처가 독자적으로 사업계획을 수립하여 추진하기보다는, 주관기관의 일정한 지침이나 기준 하에 진행되는 경우가 많다. 예컨대 조직의 경우 행정자치부가 매년 “정부조직관리지침”을 수립하여 부처에 제시한다. 그리고 실제 조직부문의 평가지표를 보면 정부조직관리지침에 의거한 평가지표들이 다수 포함되어 있다. 따라서 고유 사업처럼 정책순환의 단계별 평가나 성과분석에 초점을 두기보다는, 지침에서 제시한 기준이 잘 지켜지는지, 해당되는 실적이 있는지 등을 점검하고 평정하는 것으로 충분하다. 따라서 지원업무에 대한 평가는 지표방식을 사용하는 것이 효율적이다. 이상의 내용을 요약하면 <표 9>와 같다.

<표 9> 평가대상과 평가방법

평가대상	목적(성격)	평가방법	
고유 사업 또는 정책	분석	설계방식	과정-결과 연계
지원 업무	점검·평정	지표방식	과정(집행·산출)

<표 9>에서 제시된 내용을 정부업무평가에 적용시켜보면, 자체평가와 특정평가의 부문 간 조정방안을 도출할 수 있다. 우선 분석적 성격이 강조되는 고유사업 평가의 경우 자체평가의 주요정책과제, 재정사업, 특정평가의 특정시책 평가가 해당된다. 그리고 점검·평정 성격과 조응하는 지원업무의 경우 자체평가의 조직, 인사, 정보화, 그리고 특정평가의 관리대상업무가 해당된다(<표 10> 참조). 이렇게 보면 자체평가와 특정평가 부문 모두 분석적 성격이 필요한 평가와 점검·평정 성격이 적용되어야 할 평가가 섞여있다.

<표 10>에서 제시된 분류에 따라 정부업무평가의 부문을 조정하는 방안을 제시하면 다음과 같다. 우선 자체평가의 5가지 부문 중 주요정책 및 재정성과와 조직, 인사, 정보화를 분리하는 것이다. 부처의 고유사업에 대한 평가인 주요 정

책 및 재정성과는 자체평가의 부문으로 온존시키되, 지원 업무에 속하는 조직, 인사, 정보화는 특정평가의 관리대상업무와 합치는 것이다. 그런데 지원업무에 대한 평가를 자체평가에 두느냐, 특정평가에 두느냐의 문제가 제기된다. 정부업무평가가 자체평가를 중시한다는 점에서 지원업무에 대한 평가를 자체평가에 두는 것도 좋을 것이다.

〈표 10〉 평가방법에 따른 정부업무평가영역

평 가 방 법	성격	분석 (설계모형)		점검·평정 (지표 방식)	
		과정·결과		과정(집행·산출)	
정부업무 평가 영역	주요 정책, 재정성과	다수 부처 관련 시책, 국민관심도 높은 시책, 사회적 파급효과 큰 시책	조직, 인사, 정보화	혁신, 홍보, 규제, 법적 의무·권장 사항	
	자체평가	특정평가 (특정시책)	자체평가	특정평가 (관리대상업무)	

그러나 지원업무에 대한 평가가 점검·평정의 성격이 강하다는 점을 감안하면 상급기관이 주관하는 특정평가에 포함시키는 것이 효율적이다. 즉 지원업무에 대한 평가는 모든 부처의 기관역량을 일정 수준이상 향상 또는 유지하려는 목적에서 이루어지므로 상급 기관에서 평가하는 방식으로 할 때 실효성이 높다. 또 인사·조직·정보화가 특정평가에 포함될 경우, 평가기관의 입장에서 보면 인사나 조직 부문에 대한 자체평가업무 부담이 줄어들게 되고 자체평가위원회의 활동도 주요 정책이나 재정사업에 집중하게 할 수 있는 이점이 생긴다. 실제 부처의 자체평가위원회 활동에서 주요 정책이나 재정사업이 큰 비중을 차지하고 있다. 그리고 부처별로 작성하는 성과관리계획에서 조직과 인사를 반영한 정책목표와 성과목표가 제시되도록 하고 있는데, 이것이 실제 부처의 고유사업과 중복되거나 유기적으로 연계되지 못하는 문제점이 발생하고 있다. 따라서 조직과 인사를 자체평가에서 제외한다면 이러한 문제점을 해소하는데도 도움을 줄 수 있다.

한편 자체평가의 주요정책, 재정성과와 특정평가의 특정시책은 분석적 성격이 강한 평가 부문이다. 그러나 이 평가부문들의 성격이 유사하다고 해서 하나의 영역으로 합치는 것은 곤란한 측면이 있다. 자체평가의 주요정책이나 재정성과는 부처의 고유사업에 대한 평가이지만, 특정평가의 특정시책은 다수 부처 관련 시책 또는 현안이 되는 시책이다. 따라서 부처의 고유사업에 대한 평가는 자체평가의 논리가 적용될 수 있지만, 다수 부처가 관련된 시책의 경우 특정 부처

의 자체평가보다는 상급기관이 중립적 입장에서 평가를 수행하는 것이 바람직하다.

이상의 내용을 정리해보면 <표 11>과 같다.

<표 11> 정부업무평가 부문 조정방안

자체 평가 중시 방안	자체평가		특정평가	
	고유사업	지원업무	특정시책	고객만족도
	주요 정책, 재정성과	조직, 인사, 혁신, 정보화, 홍보, 규제 등	다수 부처 관련 시책, 현안 시책 등	주요 정책, 민원행정 서비스
평가 실효성 중시 방안	자체평가		특정평가	
	고유사업	지원업무	특정시책	고객만족도
	주요 정책, 재정성과	조직, 인사, 혁신, 정보화, 홍보, 규제 등	다수 부처 관련 시책, 현안 시책 등	주요 정책, 민원행정 서비스

V. 결 론

우리나라의 평가제도는 부처별 평가모형을 채택하고 있다. 기관평가와 통합 정부업무평가의 패러다임에서 나타나는 부처별 평가는 부처의 고유사업 및 지원업무를 평가대상으로 하면서 계획-집행-산출이라는 과정에 초점을 두고 있으며 목표달성을 점검하는 수준에서 이루어지고 있다. 또 평정결과의 서열화라는 부담 속에서 가치중립적 평가로 치우치는 경향이 있다.

이러한 문제점과 관련하여 평가방법론적 측면에서의 개선방안을 다음과 같이 제시하였다. 우선 부처의 고유사업 및 특정시책에 대해 기관별 종합평정을 위한 지표방식보다는 설계모형에 입각한 분석적 평가방식을 적용하는 것이다. 이를 위해 고유사업에 대한 평가결과를 점수로 환산하여 서열화하는 방식을 지양하고, 부처에 대한 자율성을 최대한 보장하는 것이 필요하다. 다음으로 부처의 지원업무에 대한 평가는 점검·평정의 성격이 강하므로 상급기관에 의한 지표방식의 평가를 적용하는 것이 바람직하며, 필요하다면 서열화를 적용하는 것도 가능하다.

참고문헌

- 공병천. (2007). 중앙행정기관 자체평가의 주요 이슈별 현안 및 과제. 「계간 감사」, 여름호: 64-69.
- 국무조정실. (2004). 「2004년도 정부업무 평가지침」.
- 국무조정실. (2006a). 「2006년 하반기 정부업무평가 설명회 자료집」.
- 국무조정실. (2006b). 「주요정책과제평가 세부착안사항」.
- 국무조정실·정책평가위원회. (2002). 「정부업무평가백서」.
- 김명수. (2000). 「공공정책평가론」. 서울: 박영사.
- 김주환. (2003). 우리나라 정부기관의 평가방법론의 문제점과 개선방안. 「한국정책분석 평가학회 학술대회 발표논문」.
- 김현구. (2003). 정부업무 기관평가의 이론적 논고. 「한국행정학보」 37(4): 57~78.
- 노화준. (2003). 「정책평가론」. 서울: 법문사.
- 이진주 외(1996) 「정책평가를 위한 새로운 모형」. 서울: 나남.
- 행정자치부·한국행정연구원. (2007). 「2006년도 조직부문 자체평가 결과분석 및 지표 개선 연구」.
- Behn, R. D. (2003). Why Measure Performance? Different Purposes Require Different Measures. *Public Administrative Review* 63(5): 586-606.
- Brudney, J. L. and England, R. E. (1982). Urban Policy Making and Subjective Service Evaluation: Are They Compatible? *Public Administration Review* 42(2): 127-135.
- Chen, H. T. (1996). A Comprehensive typology for program evaluation. *Evaluation Practice*, 17 (2): 121-130.
- Clarke, A. & R. Dawson. (1999). *Evaluation Research - An Introduction to Principles, Methods and Practice*, London: Sage.
- Cook, T. D. (1985). Postpositivist Critical Multiplism. in R. L. Shotland and M. M. Mark (eds) *Social Science and Social Policy*, 21-62. Thousand Oaks, CA: Sage.
- Cook, T. and C. Reichardt (1979). *Qualitative and Quantitative Methods in Evaluation Research*, Beverly Hills, California: Sage Publication.
- Greene, J. C., Benjamin, L., & Goodyear, L. (2001). The Merits of Mixing Methods in Evaluation. *Evaluation* 7(1): 25-44.
- Halachmi, A. (1982). Purpose and Perspective in Evaluation: A Case Study. *Administrative Science Review* 12(2): 1-16.
- Halachmi, A. & G. Bouckaert (1996). Performance Appraisal and Rubik's Cube. in

- A. Halachmi & G. Bouckaert (eds) *Organizational Performance and Measurement in the Public Sector*, Westport, Connecticut: Quorum Books.
- King, G., Keohane, R., & Verba S. (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton University Press.
- Nachimias, D. (1979). *Public Policy Evaluation: Approaches and Methods*. New York: St. Martin's Press.
- Rossi, P. H. & Freeman, H. E. (1985). *Evaluation: A Systematic Approach*, Beverly Hills, California: Sage Publication.
- Wholey, J. S. (1983). *Evaluation and Effective Management*, Boston: Little Brown and Co.
- Wholey, J. S., K. E. Newcomer, and Associates (eds.) (1989). *Improving Government Performance*. CA: Jossey-Bass.

Abstract

Evaluation Methodology and Government Performance Evaluation

Kwanghee Lee

This article analyzes government performance evaluation through the framework which consists of indicator mode versus design mode and process evaluation versus outcome evaluation. Evaluation where government performance evaluation focuses on the policy process was not able to show the entire government outcome or consider the character of the department. The indicator mode for monitoring and assessing limits departmental accountability and learning about performance. It is efficient to apply the design mode of analysis to department programs and to apply the indicator mode to focusing the process on management work.

[Key words: evaluation method, Government performance evaluation, process evaluation, outcome evaluation, quantitative evaluation, qualitative evaluation]