

# 평가의 등급화

- 무상원조 프로젝트 등급화를 중심으로 -

이 지 영 (ODA평가실 평가전문관)

## 목 차

1. 들어가며
2. 개발사업 평가 등급화 개요
3. 선진원조기관(JICA) 등급제 사례
4. KOICA 사업평가 등급화
5. 맺으며

## 1. 들어가며

무상원조사업의 등급화. 지금까지 수원국의 상환의무가 없는 무상원조는 대부분 수원태세를 갖춘 개발도상국가에 대한 ‘경제적 또는 물질적 도움’ 정도로 인식되는 것이 일반적이었다. 하지만 무상원조 규모가 점차 증가하면서, 체계적인 성과관리틀(Performance Management Framework) 안에서 소위 ‘좋은 사업’과 ‘좋은 사업’으로 인지할 수 있는, 즉 사업 간의 비교할 수 있는 평가수단이 필요하단 의견이 대두되기 시작하였다. ‘좋다’, ‘성공적이다’ 또는 ‘효과적이다’의 의미를 개발협력의 컨텍스트(context)안에서 규정하기 위해 평가의 기준이 제시되어야했고, 그 결과 여러 가지 사업성과를 관리할 수 있는 방안들이 마련되었다. 그와 맞물려, 공정하고 객관적인 평가방법이 마련되어감에 따라 평가기능을 강화하기 위한 하나의 수단으로 공평한 평가지표를 선정하여 기준에 따른 평가 점수나 등급을 부여하였다.

대다수의 선진원조기관이 유상원조사업 평가 수행에서 등급을 사용하는 것이 당연시 되는

상황에서, 무상원조 역시 정성적인 평가와 함께 평가의 등급(Rating) 또는 점수(Scoring) 즉, 정량적인 방법을 사용하여 평가하도록 명시하고 있다. 그리고 프로젝트 수준의 사업 평가<sup>1)</sup> 경우에는 등급제가 보편적이거나, 프로그램 수준의 사업 평가<sup>2)</sup> 등급은 일반적이지 않은 것이 원조 사업 평가의 실정이다(Sasaki, 2012). 그리고 프로젝트 수준의 평가라 하더라도 사업의 평가를 점수화 하는 수준에 머무르는 경우도 있고, 점수화와 함께 등급을 제시하는 기관도 있다. 물론 사업 등급화에 대한 선진원조기관들의 회의론이 있음에도 불구하고, 그것을 잠재울 설득력 있는 논리를 찾지 못해 지금까지 정성적인 평가와 함께 등급제를 동시 시행해오고 있다.

KOICA 역시 평가의 등급제가 필요한 것인지에 대한 논의가 꾸준히 있어왔고, 등급제를 도입하려는 시도와 노력들이 있었지만 무상원조사업의 등급제가 과연 필요한 것인지에 대한 당위성 및 평가등급제도 노력 대비 활용도가 낮다는 의견이 지배적이며 지금까지 확실한 제도로 자리매김을 하지 못했던 실정이었다. 하지만 지속적으로 평가의 책임성 확보를 강화해야한다는 국제적 원조사업 평가흐름에 맞추어 평가의 명확성을 높일 수 있는 방안을 마련하고자 올해부터 프로젝트 사후평가 대상으로 평가를 점수화하여 점수범위별로 등급제를 시행하고 있다.

## 2. 개발사업 평가 등급제 개요

‘평가를 등급화 한다’는 것은 문자 그대로 평가수행에서 등급을 부여하는 과정을 의미한다. 여기서 등급이란 다양한 방법으로 표현 될 수 있는데, A, B, C 등의 알파벳 문자로 나열을 할 수 있고, 좋다, 나쁘다, 보통이다 등의 기호의 정도를 표현할 수도 있으며, 매우 만족한다, 만족한다, 불만족한다 등의 만족도로 표시할 수도 있다. 이렇게 다양한 방법으로 표현할 수 있는 등급의 전제는 그 등급이 나온 근거가 필요하다는 것이다. 일반적으로 개발협력 사업의 평가에서는 각 등급이 어떤 근거로 표현되었는지 평가기준에 따라 점수를 측정하고, 측정된 점수의 구간을 정해 그 구간에 맞는 등급을 3, 4, 또는 5 구간 등급으로 나누는 방법이 사용된다. 일반적으로 평가의 등급제라 하면 평가의 점수와 등급을 동일시하는 경우가 대부분이다.

하지만 사업을 평가하여 점수를 정하는 것과 평가의 등급을 부여하는 것은 다른 개념이다.

1) 프로젝트 평가란, 구체화된 자원 및 추진일정 안에서 특정 목표달성을 위해 설계된 개별적 개발사업에 대한 평가로 때로는 좀 더 광범위한 프로그램의 틀 안에서 이루어질 수 있음

2) 프로그램 평가란, 전지구적, 지역별, 국가별, 분야별 개발 목표 달성을 위한 일련의 개발사업에 대한 평가임. 개발프로그램은 분야, 주제 및 지역을 교차하는 다양한 개발활동을 포함함

평가형태는 랭킹(ranking), 등급제(grading), 점수제(scoring), 비율화(apportioning) 4가지로 구분이 된다(Scriven, 2007). 이 논문에 따르면, 등급제의 가장 일반적인 등급분류는 A부터 F까지로 미국에서 가장 보편적으로 사용하고, 영국에서는 우수, 매우 좋음, 좋음, 보통, 미흡으로 구분하여 사용한다고 한다. 평가의 정량적인 방법으로 가장 일반적인 점수제는 점수화가 가능하다면 가장 좋은 평가등급이 가능하다는 장점이 있지만, 평가가 점수화로 가능하지 않은 상황이나 모호한 기준으로 점수를 부여할 수 있다는 환경적 제약요인이 있을 수 있다. 그리고 이러한 등급제와 점수제의 장점을 혼용하여 점수가 전체의 80~100% 사이에 속하는 사업은 등급 A를 부여하는 등의 평가의 등급에 점수를 기반으로 각 점수의 범위 안에 등급을 부여하는 점수제바탕의 등급제를 사용하는 경우가 있다. 이처럼 어떤 제도를 도입할 때 긍정적 및 부정적 효과가 있듯이, 평가등급 도입 시에도 이와 같은 긍정적 및 부정적 요인을 분석해 볼 수 있다.

## 가. 평가등급의 긍정적 요인

평가결과의 등급화는 평가 기준에 따라 가치를 표준화할 수 있다는 장점이 있다. 아래 <표 1>에서와 같이, 기준 1에 대해 각 가치에 해당하는 정의에 따라 기준 1의 가치 1, 2, 3, ... 등을 결정한다. 개발협력사업 평가의 경우에는 기준이 DAC 5대 기준인 적절성, 효과성, 영향력, 효율성, 지속가능성이 될 수 있고, 가치는 매우만족, 만족, 보통, 미흡이 될 수 있다. 예를 들어, 한 사업이 적절성의 기준 평가결과로 가치값 A를 받았다고 하면, 그 사업은 A 등급의 가치를 부여받았다고 볼 수 있다. 그리고 기준이 동일하다는 전제하에 그 가치는 어느 사업이나 동일하다. 그래서 평가시스템의 가치기준이 체계적으로 구분 되어있고, 가치를 표준화한다면 대내외적으로 사업평가결과 등급의 명확성을 높일 수 있다.

〈표 1〉 평가등급의 기준에 대한 가치 정의

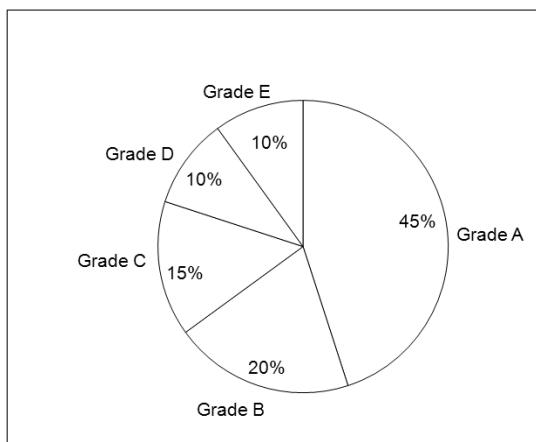
	기준 1	기준 2	기준 3	...
가치 1	정의	정의	정의	...
가치 2	정의	정의	정의	...
가치 3	정의	정의	정의	...
...	...	...	...	...

그리고 사업 등급체계의 도입은 간단한 숫자 혹은 글자로 되어있어 성과를 쉽게 이해할 수 있다. 앞서 언급한 바와 같이 평가등급은 A, B, C 등과 같은 알파벳 혹은 매우 좋음, 좋음 등의 만족도 등으로 표현할 수 있는데, 사업을 잘 알고 있는 관련 이해관계자들은 한 단어로 표현된 평가결과가 없다하더라도 사업의 평가결과에 대해 이해가 쉽지만, 그 외의 대다수는 평가결과의 기술로서는 사업이 성공적 이었는지 혹은 아니었는지에 대한 가늠이 쉽지 않다. 그러므로 사업의 성과를 쉽게 인지할 수 있기 위해서는 정성적으로 기술한 평가보다 평가의 등급화가 보다 효과적이다.

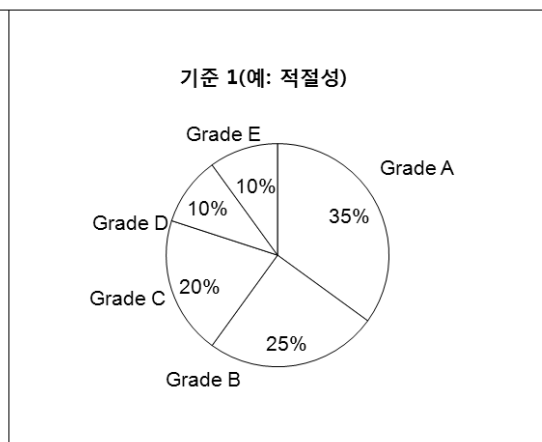
또한 평가 등급의 결과로 향후 성공 및 실패요인 혹은 개선점을 파악하기에 매우 유용하다. 가령, <표1>의 A 등급 사업이 가치등급의 정의를 바탕으로, 어떤 기준과 근거로 A가 되었는지에 대한 성공 요인을 쉽게 분석할 수 있게 되고, E 등급의 사업이 어떤 기준과 근거로 E가 되었는지에 대한 실패 요인 및 향후 개선점을 파악하기에 활용도가 높다.

더하여, 등급을 활용한 통계작업에도 효과적으로 사용될 수 있다. 개발사업 평가의 경우, <그림1>과 같이 적절성, 효과성 등의 기준별로 전체 사업에 대한 평균 등급, 즉 A등급 사업, B등급 사업 등이 각각 몇 프로인지 제시가 될 수 있고, <그림2>와 같이 각 기준별로도 그룹핑을 하여 몇 프로가 적절성에서 A가 나왔는지에 대해 분석 자료를 도출해낼 수 있다.

〈그림 1〉 전체사업 평가가치별 비율



〈그림 2〉 평가기준별 비율



## 나. 평가등급의 부정적 요인

일반적으로 평가등급을 제시할 때, 기술적 평가내용이 함께 제시되는 경우가 많다. 그 이유는 평가등급의 제시만으로는 풍부한 평가정보를 제공해 줄 수 없기 때문이다. 특히 개발협력 분야에서 발주되는 개별사업은 동일한 사업이 없고, 지역별 분야별로 다양한 분야의 사업을 포괄하고 있어 평가등급이 이러한 다양성에 대한 정보를 모두 제공해주지 못하는 한계가 있다. 그리고 명확한 등급체계가 이루어지고, 가치가 동일하게 부여된다 하더라도 다양한 평가자에 의해 평가가 이루어지므로 점수가치를 객관화하기 힘든 경우가 많다. 그리고 사업 착수 전, 수행단계, 종료 등의 단계가 각각의 사업마다 다르기 때문에, 평가의 시기에 따라 점수가 달라질 수 있는 가능성도 존재한다. 그리고 등급가치 기준이 확실하다 하더라도, 다양한 평가자에 의해 평가가 이루어지므로 평가자마다 해석이 다를 수 있어 본래 의도한 평가등급제도의 목적을 이루지 못하는 경우도 있다.

그리고 좋은 평가등급에 목적을 두어 사업성과의 질에 중점을 두지 않을 수 있다. 일반적인 시험의 평가나 여타 다른 사업들은 질문에 대한 답이 명확하여 성과의 질에 대한 판단이 차치하는 경우가 있지만, 개발사업 평가 특히, 무상원조사업은 소프트웨어 사업이 대부분이어서, 평가를 수행할 때 사업성과의 질에 대한 고려가 필수적이다. 그래서 사업수행의 성공여부를 판단할 때 평가등급이 주요한 판단기준이 된다면, 평가등급이 부여되는 각 기준의 정의에 맞추어 사업을 수행하기 때문에 평가등급이 A가 나왔다고 하더라도, 본연의 사업수행 목적대로 운영이 되지 않아 사업의 효과를 저해할 수 있다는 가능성도 배제할 수 없다.

이와 같은 평가의 질 부분만이 아니라, 국가별로도 사업의 규모 차이가 나고, 외부적 환경 요소들이 모두 같을 수 없으므로, 등급으로 사업을 비교한다는 것이 무의미할 수 있다. 또한 평가점수가 향후 발주하는 사업에 영향을 주게 된다면, 이러한 사업평가결과와 등급의 비교로 인해 잘못된 판단을 초래할 가능성이 있다. 그러므로 이러한 평가등급의 부정적 요소를 어느 정도 해소를 하기 위해 사업평가등급을 제시할 때, 기술적인 평가방법을 더하여 정성적인 평가를 수행하는 것이 그 이유이다.

### 3. 선진원조기관(JICA)의 등급제 사례

평가 등급제는 World Bank나 ADB와 같은 선진원조개발은행을 주축으로 하여 대다수의 국가 프로그램 평가에 도입되어 지금까지 사용되어오고 있다. 하지만 수행된 사업을 통한 교훈 및 책임성 증진이라는 국제개발협력사업 평가의 목적을 이루기 위해서는 평가등급제도가 필요하지 않다는 회의론이 선진원조기관 사이에서 계속 논의되고 있어, 선진원조기관마다 평가제도를 운영하는 방식이 다른 것이 현실이다. 그리고 서두에 언급한 바와 같이 프로젝트형 사업보다는 국가형 프로그램 사업에 평가등급이 도입되는 경우가 대부분으로, 회의론이 있음에도 불구하고 사업을 관리 측면에 중점을 두어 평가등급제를 유지해오고 있다. 아래 <표1>은 국내외 선진원조기관의 등급제 현황으로 국내에서는 유상원조기관인 EDCF에서 프로젝트 사후평가에 평가등급을 적용하여 사용하고 있고, 해외 원조기관에서는 World Bank, ADB, JICA 등이 평가등급을 적용하고 있음을 알 수 있다.

그리고 평가등급을 적용하고 있는 선진원조기관 중에서도, 평가등급의 기준을 적용함에 있어 ADB와 같이 DAC 5대 기준을 준용하는 기관과 World Bank와 같이 효과성과 영향력을 묶어 4대 기준을 적용하여 평가등급을 산정하는 기관으로 나눌 수 있다. 효과성과 영향력을 분리하는 경우는 프로젝트형 사업이 직접적으로 효과가 나타나는 경우와 향후 그 사업에 대한 영향이 구분 되어야한다는 논리이고, 효과성과 영향력을 분리하지 않은 경우는 중기 및 장기 성과의 구분이 모호한 경우가 많아 구분하지 않는다는 논리이다.

특히 KOICA는 프로젝트형 사업의 사후 평가를 주로 수행하고 있어, 무상원조 프로젝트형 사업의 사후평가 등급제를 사용하고 있는 JICA의 사례를 자세히 살펴보고자 한다.

JICA의 원조유형은 기술협력(Technical Cooperation), ODA 차관(ODA Loan), 무상원조(Grant Aid)로, 2008년 이후 통합 평가체계를 도입하여 사후평가를 수행하고 있다. 특히 프로젝트의 사후평가에 평가등급을 산정하여 평가결과의 명확성을 높이고 있다. 평가대상은 2억엔 이상의 프로젝트를 대상으로 시행한다. 평가기준은 DAC 5대 기준 중에 효과성과 효율성을 함께 평가하여, 적절성, 효율성, 효과성/영향력, 지속가능성으로 평가한다. 각 기준별로 평가점수를 3점 척도로 나누어 산출하고 있으며, 전체 평가등급은 A(Highly Satisfactory), B(Satisfactory), C(Partially Satisfactory), D(Unsatisfactory) 4단계로 나누어져있다.

〈표 2〉 국내외 선진원조기관의 등급제 현황

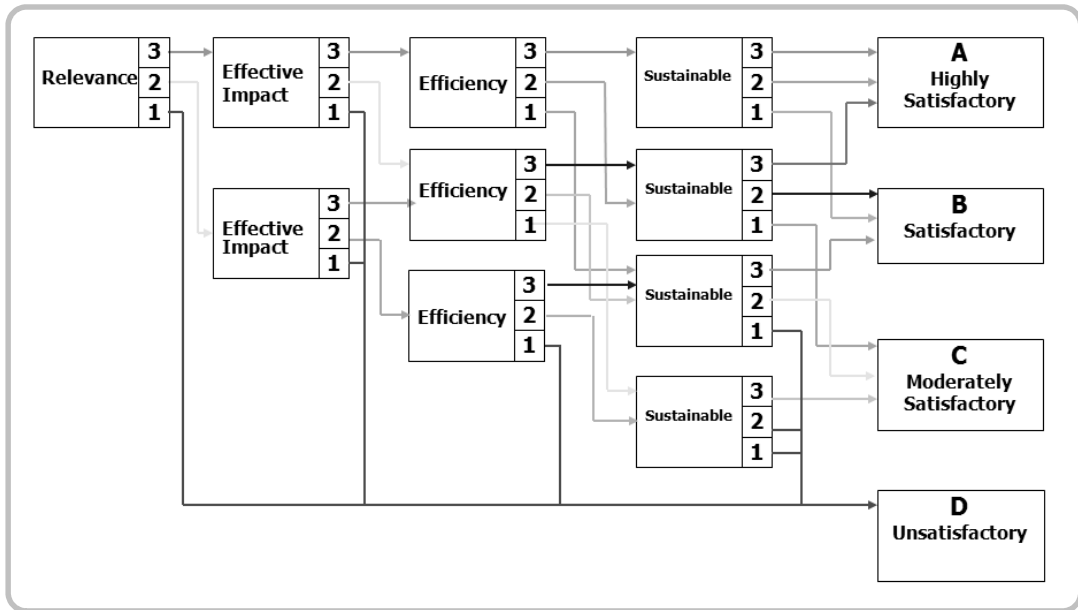
구 분	기 관 명	평가등급제 상세내용
국내 기관	EDCF	<ul style="list-style-type: none"> <li>• 사후평가에 평가등급 적용</li> <li>• 평가등급(4단계): 매우 성공적, 성공적, 부분 성공적, 미흡</li> </ul>
국외 기관	World Bank	<ul style="list-style-type: none"> <li>• 1970년대부터 평가등급체계 도입</li> <li>• 영향력이 큰 사업에 대한 관리지표로써 활용</li> <li>• 평가등급(5단계): 매우 만족, 만족, 부분 만족, 불만족, 매우 불만족</li> </ul>
	ADB	<ul style="list-style-type: none"> <li>• WB와 함께 사업 평가등급체계 도입</li> <li>• 평가등급(4단계): 매우 성공적, 성공적, 부분 성공적, 성공못함</li> </ul>
	CIDA	<ul style="list-style-type: none"> <li>• 양자원조기관 중 유일하게 프로그램 평가등급 적용</li> <li>• 분야별 평가등급은 있으나, 전체 프로그램 평가등급 없음</li> <li>• 평가등급(5단계): 매우 만족, 만족, 부분 만족, 불만족, 매우 불만족</li> </ul>
	JICA	<ul style="list-style-type: none"> <li>• 사후평가에 평가등급 적용</li> <li>• 사업의 평가등급을 전적으로 신뢰하기 보다는 참고용으로 사용</li> <li>• 평가등급(4단계): A(매우 만족), B(만족), C(부분 만족), D(불만족)</li> </ul>
	GIZ	<ul style="list-style-type: none"> <li>• 책임성 측면에서 효과적인 지표로 활용하기위해 도입</li> <li>• 평가등급(6단계) : 1=매우만족, 6=매우 불만족</li> </ul>

\* 자료출처 : Ryo Sasaki, 2012, 수정

JICA는 Rating Flowchart라는 독특한 평가등급 산출방식을 사용하여 각 기준에 해당하는 점수를 따라 A, B, C, D의 등급을 결정하게 된다. 가장 주요한 특징은 적절성부터 효과성, 효율성, 지속가능성 순으로 flowchart가 구성되어, 적절성에서 1점을 받은 사업은 D등급으로 평가 결과를 결정한다. 그리고 적절성에서 2, 3점을 받더라도 그 다음 기준인 효과성에서 1점을 받은 사업 역시 D등급으로 판정하여, 평가결과에서 적절성과 효과성의 비중이 큼을 알 수 있다.

2011년도에 수행된 사후 평가는 107개 프로젝트형 사업이 대상이었고, 그 중 ODA차관은 51개 사업, 무상원조는 36개 사업이고, 기술협력부분은 20개 사업이었다. 대부분이 아시아, 중앙아메리카 및 남아메리카 지역에서 수행된 사업이었고, 분야는 도로, 다리, 전력, 수력, 그리고 교육이었다. 총 사업 중 A등급은 37%(39개 사업), B등급은 39%(42개 사업), C등급은 16%(17개 사업), D등급은 8%(9개 사업)를 차지하였고, A와 B등급은 전체의 75%로 대부분의 사업 만족도가 높은 것으로 나타났다. 각 기준별로 3점을 받은 비율을 살펴보면, 적절성은 전체 프로젝트의 95%, 효과성/영향력은 66%, 효율성은 30%, 지속가능성은 44%였다. 그리고 2점을 받은 비율은 적절성 5%, 효과성/영향력은 28%, 효율성은 58%, 지속가능성은 45%로 나타났다.

〈그림 3〉 JICA 평가등급 산출방법



\* 자료출처 : JICA, 2012

## 4. KOICA 프로젝트 평가 등급제

KOICA의 원조유형은 무상원조(Grant Aid)로 대부분이 교육 및 보건을 주요 분야로 한 소프트웨어 사업을 수행하여왔기에 사업수행의 교훈과 향후 개선점 등을 내용으로 하는 정성적 평가를 중점적으로 수행하여 왔다. 물론 국제개발사업의 평가 흐름에 발맞추고자 2007년 종료평가를 대상으로 평가 등급제를 시범적으로 시행한 적이 있었으나, 사회 인프라 구축 등의 사업이 대부분인 유상원조와는 달리 사업의 효과가 사업 종료 시에 나타나지 않는 보건 및 교육 분야가 많아 평가의 등급을 산정하는 것이 효과적이지 않다는 지적이 있어 평가등급제도의 실효성을 거두지 못했다.

그리고 평가 종류도 사전평가, 중간평가, 종료평가, 사후평가, 영향평가 등으로 다양하여, 모든 평가를 대상으로 등급제를 수행하는 것이 바람직한지, 아니면 사업의 효과성이 어느 정도 가시화될 수 있는 사후평가를 대상으로 등급제를 시행할 것인지에 대해 논의가 계속되어왔다. 현재 KOICA는 프로젝트형 사업 평가에서는 사후평가를 대상으로 하는 사업에 한정하여



등급제를 수행하고 있고, 2013년도 기준으로 교육, 보건, 공공행정, 농림수산, 산업에너지 분야에서 19개 사후평가를 대상으로 평가등급을 산정할 계획에 있다. 그리고 사후평가를 대상으로 하는 사업인 점을 감안하여, DAC 5대 평가기준에서 효과성과 영향력을 하나로 간주하고, 4대 평가기준인 적절성, 효과성/영향력, 효율성 및 지속가능성에 대한 점수를 산출하여 점수 대별로 그에 해당하는 등급을 결정할 예정이다.

〈표 3〉 KOICA 평가방법 개선 전후 비교

구 분	기 준 (~'12)	개선 후
평가대상사업	사후평가 대상사업	
평가방법	정성적 평가	정성적 평가 + 등급제
평가항목 및 기준	〈5대 기준〉 적절성, 효과성, 효율성, 지속가능성, 영향력	〈4대 기준〉 적절성, 효과성/영향력, 효율성, 지속가능성

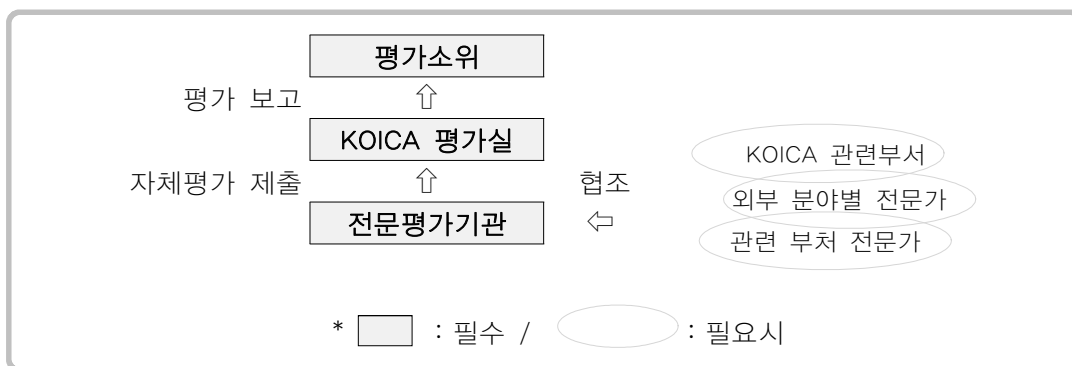
평가 등급의 산출 방법은 각 평가기준별 1,2,3점으로 총 12점 만점이며, 총점이 11점에서 12점사이의 등급은 매우 성공적(very successful), 9점에서 10점 사이는 성공적(successful), 7점에서 8점 사이는 부분 성공적(partly successful), 6점 이하는 미흡(unsuccessful)로 등급을 산정한다.

〈표 3〉 KOICA 평가 등급 산출 방법

평가 기준		정의		
		3	2	1
적절성	• 수원국의 개발필요 및 목표, 전략 일치 정도 등	매우 적절	부분 적절	부적절
효과성/영향력	• 의도한 산출물, 목적, 목표를 달성한 정도 • 사업을 통해 상황이 변한 정도(외부효과 제외) 등	계획된 효과 80%이상	계획된 효과 50~80%	계획된 효과 50%이하
효율성	• 사업이 효율적으로 운영되는지 여부 • 최적의 시나리오로 제한된 자원으로 최대의 결과를 달성했는지 여부	효율적 (100% 이내)	부분 효율적	비효율 (150%이상)
지속 가능성	• 사업이 수원국 우선순위 부합 여부 • 효과적 운영을 위한 제도 및 조직 여부 • 사업지원 종료 후 결과를 유지할 잠재력 등	매우 지속가능	문제점 개선으로 지속가능	지속불가능

KOICA의 평가 등급체계는 평가의 공정성을 위해 평가전문기관을 입찰 선정하여 실시하고 KOICA는 외부 평가기관의 정상·등급 평가에 관여하지 않는다. 평가전문기관은 사업 등급과 관련하여 필요하다고 판단될 경우에 KOICA 관련 부서 및 외부 분야별 전문가, 관련 부처 전문가의 의견자문을 구할 수 있으며 상세내용을 최종 결과보고서에 수록한다.

〈그림3〉 KOICA 평가 등급 산출 방법



## 5. 맺으며

무상원조기관인 KOICA의 프로젝트형 사업평가에서의 평가등급제 시행은 평가의 책임성 확보차원에서 큰 의미를 둘 수 있다. 지금까지 소프트웨어 개발사업에 중점을 두어 효과적인 사업의 수행을 하기 위해 과정평가와 성과평가를 하였던 것에 더하여, 평가의 등급제는 사업평가 결과의 명확성을 높일 수 있는 방안으로서 역할을 할 것으로 기대되고 있다. 개발도상국에 원조를 함에 있어 단기적 성과가 나는 사업뿐만이 아니라 경제·사회적으로 지속가능한 개발을 이룰 수 있는 기초사업, 특히 보건, 교육 등의 사업을 수행함에 있어 평가 점수체계의 도입은 분야를 좀 더 쉽게 이해할 수 있는 척도로서의 역할을 할 것이다. 그러므로, 무상원조라는 KOICA 사업의 특성을 고려하여 앞서 살펴본 선진 원조기관의 평가등급체계 선례를 참고하여 KOICA만의 독자적인 평가등급체계를 계속적으로 개발해 나가야 할 필요가 있다.

특히 개발협력사업 평가 가이드라인에 일반적인 평가 질문사항들이 열거 되어있지만, 전체 분야에 공통적으로 평가할 수 있는 평가항목이 없으므로, 보건과 교육분야 사업을 평가할 때 평가팀에서 수행할 평가 필수 질문들이 상이하여 전 분야 사업에 획일적인 지표를 바탕으로

평가하기가 어려운 실정이다. 이번 2013년도 19개 사후평가 시범사업을 통해 등급화가 가능하다고 판단되는 분야에 대해 공통평가지표를 표준화하여 평가등급을 수행하는 것도 향후 고려해야 할 사안 중 하나이다. 그리고 DAC 5대 평가 기준 이외에도 기타 범 이슈에 관한 평가 등급화가 쉽지 않으므로, 그에 대한 평가지표의 표준화가 이루어질 수 있는지에 대한 가능성 여부를 지속적으로 논의해야 한다.

## 참고문헌

한국국제협력단, 2008, 개발협력사업평가 가이드라인

한국국제협력단, 2007, 사업평가 가이드라인

EDCF, 2011, EDCF 평가매뉴얼

JICA, 2012, Annual Report

JICA, 2011, Annual Report

Michael Scriven, 2007, The Logic of Evaluation.

Ryo Sasaki, 2012, An In-Depth International Comparison of Major Donor Agencies: How Do They Systematically Conduct Country Program Evaluation?