

국방R&D사업 자체평가시스템 메타평가

김순영*[†] · 하규수**

*국방기술품질원

** 호서대학교

A Study on the Meta Evaluation for Defense R&D Programs

Soon Yeong Kim*[†] · Kyu Soo Ha**

*Dept. of Technology Evaluation, Defense Agency for Technology and Quality.

**The Graduate School of Venture, Hoseo University.

This study is the result of meta evaluation for the self evaluation of defense R&D programs in Korea by using meta evaluating indicators. The overall meta evaluation result of defense R&D programs gained 74.3 points out of 100, so it was evaluated as 'Good'. But it demonstrated that further improvement for overall system of defense R&D programs evaluation is required. And especially, it demonstrated that more alternatives are necessary in order to improve the utilizations and the feedbacks of evaluation results. The evaluation context component gained 80.2 points out of 100, so it was evaluated as 'Very Good'. The evaluation input component gained 73.1 points out of 100, so it was evaluated as 'Good'. The evaluation process component gained 74.8 points out of 100, so it was evaluated as 'Good'. And the evaluation outcome component gained 69.0 points out of 100, so it was evaluated as 'Good'. Basic model of meta evaluation was derived from the literature review and brain storming. And this meta evaluation model was determined by adopting the result of experts who performed evaluations for defense R&D programs in recent years. The reliability of components and items was verified by Cronbach's α coefficient. It was over 0.6 in evaluation components and items. And the reliability of evaluation context was 0.877, that of evaluation input was 0.755, that of evaluation process was 0.755, that of evaluation output was 0.755 respectively. From the analysis, it is attempted to identify possible problems and to find out the ways of improvements related to the self evaluation system of defense R&D programs. The ultimate objective of this study is to manage the programs effectively and improve the reliability and the objectiveness of the defense R&D programs.

Keywords : Defense R&D Programs, Performance Evaluation, Self Evaluation, Meta evaluation

1. 서 론

정부에서는 지속적인 연구개발투자를 통해 미래의 성장 동력을 끊임없이 확충해 나아가고 기술경쟁력을 키우기 위해 많은 노력을 하고 있다. 그리고 2009년도 우리나라 국가 연구개발예산은 12조 원을 넘어서는 등 연구개발

예산이 크게 증가하고 있는 추세다. 또한 최근의 연구개발사업이 대규모화, 장기화됨에 따라 성과중심의 평가의 필요성이 커지고 있고, 한편으로는 국가연구개발사업의 책무성과 효과성을 제고해야 한다는 목소리가 점차 높아지고 있다. 따라서 정부에서는 국가연구개발사업의 효율성과 생산성을 제고하기 위하여 2005년 12월에 연구개발성과평가

법을 제정하여, 정부 각 부처는 소관 연구개발사업에 대한 조사·분석·평가를 실시하고 있다(과학기술부, 2005).

또한 2006년 2월 방위사업청이 개칭하면서 국방연구개발사업에 대해 성과중심의 평가시스템이 독립적으로 처음 실시됨에 따라 이 평가시스템이 얼마나 효율적인가에 대한 관심도 높아지고 있다.

본 논문에서는, 국방연구개발사업 자체평가 시스템 측면에서 메타평가를 실시하여 문제점 및 개선방안을 도출 적용함으로써 국방연구개발사업을 보다 효율적으로 관리하고 연구성과에 대한 신뢰성과 객관성을 제고시키고자 한다.

2. 국방R&D사업 고찰

2.1 국방연구개발 활동의 특성

국방연구개발의 개념과 특성을 살펴보면, 국방 연구개발은 '대내·외적인 위협을 제거하기 위해 국방에 필요한 무기체계, 자동화 체계에 관한 과학적·기술적 조사, 연구, 개발 및 시험 등을 수행하는 일련의 과정'이라고 할 수 있는데, 그 범주는 국방력 형성에 관련된 전 분야를 망라한 것이라 하겠다. 계획된 예산 내에서 요구되는 첨단 무기체계 및 기술을 획득하고, 정보 노출을 방지하고, 군수지원 중단에 의한 작전수행 제한 방지를 그 목적으로 한다. 그리고 해외구매를 통한 무기체계의 획득은 연구개발을 통한 무기체계보다 많은 수명주기비용을 요구하게 된다. 또한 연구개발을 수행하면 미래 무기체계 개발 소요의 준비비용을 절감할 수 있고, 국외 도입 시에는 협상능력을 제고하는 그 바탕이 되기도 한다.

2.2 국방연구개발사업 형태

방위사업청(2006)에 따라 국방연구개발사업을 사업추진방법으로 구분하면, 방위력개선사업은 무기체계 및 핵심기술 연구개발사업과 무기체계 구매사업으로 구분한다. 그리고 국방과학기술 경쟁력제고 및 국가자원의 효율적인 활용을 위하여 정부 관련 부처와 공동으로 민·군 겸용기술 사업이나 국책사업을 추진할 수 있으며, 민·군 겸용기술 사업의 경우 사업비의 부담 및 추진절차·사업관리 등에 관한 사항은 민·군 겸용기술 사업촉진 법령에 따르고, 기타 국책사업에 관한 제반사항은 부처의 합의에 의해 제정된 별도의 법령에 따른다.

2.3 연구개발사업 평가시스템

우리나라에서 정부 예산으로 지원되고 있는 국가연구

개발 활동은 세부 평가를 실시하고 있다.

먼저, 부처 자체평가는 연구개발사업을 수행하는 중앙행정기관은 당해 연도 특정평가 대상이 아닌 사업 중착수 후 1년 이상 경과된 사업에 대해서는 사업특성에 따른 성과중심의 자체평가를 실시한다. 중앙행정기관들은 평가대상 사업별로 전략목표와 성과목표를 설정한 후, 기획재정부의 표준성과지표를 활용하여 사업특성을 반영한 성과지표를 자율적으로 설정하여 평가를 실시한다. 그리고 특정평가는 기획재정부의 심층적인 성과평가 기능을 강화하기 위하여 주요 사업에 대한 평가를 실시하는 것이며, 상위평가는 자체평가 성과목표/지표의 적절성 및 자체평가 방법의 객관성·공정성 등을 점검하는 것이다.

3. 국방R&D사업 메타평가 모형 분석

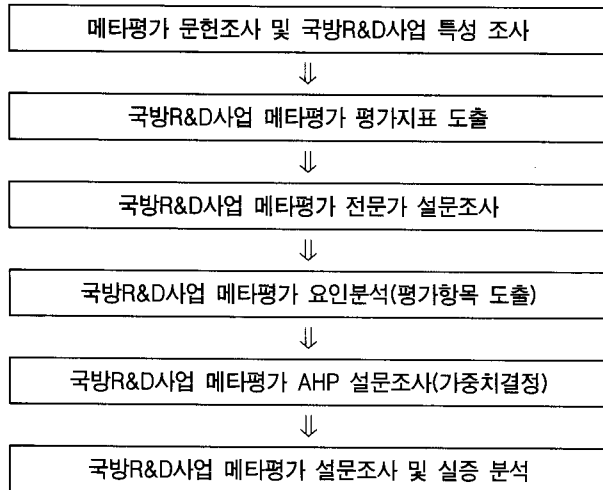
3.1 메타평가의 개념

메타평가(meta evaluation)는 평가에 대한 평가로 정의되고 있다. Cook and Gruder(1978)는 메타평가를 무엇을 해야 하는지에 관한 진단적 환류(diagnostic feedback)와 조언을 통해 평가자들이 평가목표를 달성하는 것을 돕기 위한 수단으로 보았다(Yi Chan Goo, 2003). Stufflebeam(1981)은 메타평가를 두 가지 역할로 분류하고 있는데, 평가의 가이드로서의 형성적(formative) 메타평가와 평가의 강점·약점에 대한 보고서로서의 총괄적(summative) 메타평가를 들고 있다. Larson and Berliner(1983)는 의사결정자들에게 평가에 내포된 정보의 질에 대한 독립적인 평가 제공, 평가자 선정에서의 가이드 제공, 다수의 독립적이지만 유사한 프로그램으로부터 연구지식 획득, 평가 자체 검토 수단 제공 등을 위해 메타평가가 필요하다고 하였다. Chelmsky(1985)는 메타평가를 통해 프로그램의 효과에 대한 정보를 제공함으로써 책무성과 정책형성을 유도하며, 알려지지 않은 연구의제를 제공함으로써 정책집행을 실현할 수 있다고 하였다. 그리고 Worthen, Sanders, Fitzpatrick(1997) 등은 메타평가의 목적이 평가의 잠재력을 부응하도록 돕는데 있다고 하였다.

3.2 국방연구개발사업 메타평가 연구 모형

국방연구개발사업의 메타평가 지표를 개발하기 위하여 먼저 메타평가 선행연구를 통하여 질문항목을 도출하였다. 선정된 질문항목들은 국방연구개발사업 자체평가의 전문 평가위원들의 의견을 반영하여 더욱 정제화 하였다. 그리고 메타평가 요소별 도출된 평가지표에 대한 평가항목 선정을 위해 요인분석을 실시하였다. 또한 평가항목별 가중치 배점을 객관적으로 도출하기 위해 계층

분석적 의사결정기법(AHP : Analytic Hierarchy Process)을 적용하였다. 따라서 본 논문에서는 체계적인 의사결정방법으로 국방연구개발사업의 자체평가의 목적에 맞는 메타평가 지표를 개발하여 실증 연구를 위한 것으로, 연구 모형 및 절차는 <그림 1>과 같다.



<그림 1> 연구모형 및 절차

3.3 국방연구개발사업 메타평가의 필요성

메타평가의 등장은 평가결과의 활용과 직접적인 관련이 있다(Stevenson, 2000). 평가 결과의 활용은 평가준비에서 시행과 종료 이후의 조치 등 평가과정 전반을 통하여 얻은 정보들을 진행 중이거나 향후 추진할 사업이나 정책의 개발과 집행을 포함한 정책활동에 이용하는 행동을 말한다(Leviton and Hughes, 1981). 그런데 많은 평가 전문가들은 평가결과가 잠재적 이용자들의 결정에 별로 영향을 미치지 못함을 지적한다. 이를테면 평가결과가 정책결정에 영향을 미치지 못하고 있고, 사업 및 정책평가가 하나의 요식행위에 그치고 평가결과가 정책개선에 제대로 활용되지 못하고 있으며, 정책 담당자들이 평가결과를 등한시하거나 아예 사용하지 않는 경우가 빈번하며, 비록 사용한다고 할지라도 극히 저조하다고 한다(노유진, 2005). 이처럼 평가결과의 활용이 제대로 이루어지지 않는 이유에 대하여 학자들은 다양한 이유를 들고 있는데, 이것은 바로 평가활동 자체에 대한 체계적인 평가가 필요하다는 것을 함축하고 있다.

따라서 국방연구개발사업 평가의 메타평가란 공공정책영역인 국방연구개발사업에 대한 정부개입 과정, 특히, 사업평가 과정에서의 사후적 평가 작업, 즉 국방연구개발사업에 대한 평가활동의 모든 국면인 평가의 과정과 결과를 평가하는 것이라고 할 수 있다. 요컨대 국방연구개발사업 평가란 정부가 국가 차원에서 중장기적인 기술

경쟁력을 확보하고 산업의 전략적인 발전을 도모하는 등 국가목표를 달성하기 위한 연구개발 활동의 과정과 그것의 결과물을 검토하는 것을 말한다. 이러한 1차적인 국방연구개발사업의 평가가 제대로 수행되었는가를 다시 검토하고, 시행 3년을 맞는 현 시점에서 국방연구개발사업 자체평가 시스템의 제도개선을 추구하는 것이 본 국방연구개발사업의 메타평가를 수행하는 근본적인 필요성이라 할 수 있다.

3.4 메타평가에 대한 대내·외 사례 선행연구

메타평가 모형을 연구해온 주요 학자들의 견해는 다음과 같다.

Cook and Gruder(1978)는 어떤 경우에 평가의 질이 낮아서 평가결과의 활용이 저조하다면서, 이를 해결하기 위한 방법의 하나로 일차 평가자들이 평가자료를 분석하고 있는 동안에 평가자 이외의 사람들이 동시에 평가자료를 재검토하는 등 메타평가 표준을 7가지로 정리해 그 사용방법을 설명하였다.

Stufflebeam(1981)은 메타평가를 평가작업에 대한 안내와, 그리고 평가의 가치와 장점(worth and merit)을 공개적으로 보고하는 것으로 규정하여 전자를 형성적 메타평가(formative metaevaluation) 후자를 총괄적 메타평가(summative metaevaluation)로 분류하였다. 또한 메타평가의 기준으로 기술적 적절성(technical adequacy) 기준, 유용성(utility) 기준 및 효율성(efficiency) 기준으로 나누고 있다.

Larson and Berliner(1983)는 메타평가의 상위 구성요소로 평가투입 요소, 평가과정 요소, 평가결과 요소로 나누고 각각의 하위구성요소를 제시하고 있다.

Scriven(1991)은 메타평가를 위한 도구로 KEC(Key Evaluation Checklist)를 제안하였다. 그는 3가지 범주로 구분하여 체크리스트를 만들었는데, Part A는 기반(foundation)에 대한 체크리스트를, Part B에서는 하위평가(sub-evaluation)에 대한 체크리스트를, Part C에서는 결론(conclusion)에 대한 체크리스트를 작성하여 제시하였다.

Joint Committee(1994)가 교육평가를 위해 사용한 사업평가 표준들은 현재 미국의 교육평가뿐만 아니라 주요 사업평가의 표준들로 활용되고 있으며, 이러한 표준들을 토대로 실제적인 평가가 수행되고 연구되고 있다는 점에서 매우 유용한 평가표준이다. Joint Committee의 사업평가 표준들인 유용성, 실현 가능성, 정당성, 정확성의 범주들은 나름대로 사업평가의 표준으로서 모든 것을 총 망라하고 있다.

미국의 평가협회인 AEA(1995)는 평가를 수행하는 평가자의 차원에서 평가 가이드라인을 제시하였다. 이 평가 가이드라인은 체계적인 조사, 능력, 통합성/정직, 인

격존중, 공공책임 등 5개 영역으로 구분하고 있다.

OECD(1999)에서는 ‘Best Practice Guidelines for Evaluation’을 발간하여 OECD 국가들이 평가의 활용을 증진시키고자 할 때 고려해야 할 핵심적인 이슈와 실제에 대해 다루고 있다. OECD는 평가를 위한 최고의 가이드라인을 평가로부터의 수집 단계, 평가들의 조직화 단계, 효과적인 평가수행 단계의 3단계로 구분하여 제시하고 있으며, 이는 방법론적인 문제보다는 정부의 평가활동의 관리와 개인평가의 관리에 초점을 맞추고 있다.

Hong(2000)은 국가간 정부연구개발사업의 메타평가를 비교분석하기 위해 외국학자의 선행연구를 토대로 메타평가의 분석모형에 포함되는 상위 구성요소로 평가상황, 집행메카니즘, 평가결과, 결과활용의 네 가지 범주로 구분하고 이를 메타평가를 위한 구성요소로 활용하고 있다.

이찬구(2003)는 영국의 범부처간 과학기술평가제도를 메타평가하기 위한 상위구성요소에 대하여 종전의 연구와 같이 평가기준, 평가자원, 평가수행, 평가활용의 4가지 범주로 구성하고 각각의 하위구성요소와 세부 측정항목을 부분적으로 수정하고 있다.

박종수(2003)는 정보화사업 평가계획과 국가정보화 평가계획에 근거하여 실시된 정보화사업(외부평가에 한정)을 대상으로 이의 평가실태를 분석하기 위한 메타평가모형을 설계하고 이를 적용하였다. 그는 메타평가의 평가요소를 평가상황, 평가투입, 평가수행, 평가결과, 평가활용 등 5개 영역으로 구분하였다.

이형우(2005)는 국가 연구개발 사업 평가체제에 대한 메타평가를 하는데 선행연구에서 활용한 변수를 참조하면서 연구주체인 국가 연구개발사업의 특성을 고려하여

평가의 전 과정에서 필요한 구성요소와 세부 평가항목을 도출하였다.

류영수(2007)는 기술영향평가의 메타평가 모형 개발 및 적용에 관한 연구를 수행하였다. 그는 기술영향평가를 평가기획(planning), 평가투입(input), 평가수행(process), 평가결과(output), 평가활용(utilization) 등 5단계의 순환과정으로 파악하고, 이에 대한 세부 구성요소를 제시하였다.

김병철(2008)은 정부출연연구기관 평가에 대한 메타평가모형을 설계하고 적용하기 위한 연구를 수행하였다. 그는 경제·인문사회연구회 평가시스템을 평가하는데 있어서 평가환경(context), 평가투입(input), 평가수행(process), 평가활용(utilization)의 4단계로 구분하여 세부 평가항목을 도출하였다.

그리고 국방연구개발사업의 경우 평가지표는 ‘국가연구개발사업 자체평가매뉴얼(과학기술혁신본부, 2008)’에서 권장하는 계획요소(사업목적 및 내용의 타당성, 사업추진체계의 합리성), 집행요소(사업관리 및 사업집행의 적절성, 성과관리의 적절성), 결과요소(성과달성도, 핵심지표 ①, 핵심지표 ②, 핵심지표 ③, 일반지표 ①, 일반지표 ②), 활용요소(평가결과의 활용정도)로 구성되어 있다.

3.5 국방연구개발사업의 메타평가 지표 개발

국방연구개발사업의 메타평가 지표를 개발하기 위하여 먼저 메타평가 선행연구 및 브레인스토밍(Brain Storming)을 통하여 먼저 질문항목을 도출하였다. 도출된 질문항목들은 국방연구개발사업 자체평가의 전문 평가위원들의 의견을 반영하여 더욱 정제화 하였다. 즉, 제 3.4항에서 보는

<표 1> 국방연구개발사업 메타평가의 분석 모형

평가 요소	평가 지표
평가상황 (Context)	① 평가목적의 명확성, ② 평가목적의 합리성, ③ 평가목적의 국방분야 가치에 부합 여부, ④ 평가대상사업 분류의 적절성, ⑤ 평가대상사업 기준의 명확성, ⑥ 평가대상사업의 국방분야 특수성 고려 여부, ⑦ 평가관련 이해관계자 및 수요 파악 여부, ⑧ 평가의 법적·제도적 근거의 타당성, ⑨ 평가규정의 합리성, ⑩ 평가모형의 적절성
평가투입 (Input)	① 평가인력의 양적 충분성, ② 평가위원의 전문성 및 경험, ③ 평가대상을 고려한 평가위원 배정, ④ 평가위원 구성의 공정성, ⑤ 평가위원 사전 교육의 필요성 여부, ⑥ 평가조직의 제도적 독립성, ⑦ 평가조직의 평가결과 비관련성, ⑧ 평가조직의 평가기능 부합 여부, ⑨ 평가조직 인력의 전문성, ⑩ 평가자료의 양적 충분성, ⑪ 평가자료의 질적 우수성, ⑫ 평가자료 제공의 적시성, ⑬ 평가자료 활용의 용이 여부, ⑭ 평가예산 규모의 적정성, ⑮ 평가예산의 안정적 지원 여부
평가수행 (Process)	① 평가목적과 평가방법의 부합성, ② 평가방법의 효율적 수행, ③ 평가팀 구성의 균형 여부, ④ 평가분석방법의 사용 여부, ⑤ 평가지표의 양적·질적 적절성, ⑥ 평가지표의 평가대상사업 특성 반영 여부, ⑦ 평가지표의 측정 요소별 적절성, ⑧ 적절한 평가절차 수립 여부, ⑨ 평가절차의 일관성 여부, ⑩ 피평가기관의 의견 반영 여부, ⑪ 평가기간 및 시간의 합리성, ⑫ 특수 이해집단과의 연계성 여부, ⑬ 평가수행 중 내·외부의 압력 행사 여부, ⑭ 평가주관기관의 공정성, ⑮ 이해관계자간 의사소통 여부
평가효과 (Outcome)	① 평가결과 공개 여부, ② 피평가기관의 소명 기회 여부, ③ 평가보고서의 의사결정권자 보고 여부, ④ 평가보고서의 충실성, ⑤ 평가보고서의 적기 제공 여부, ⑥ 평가결과의 차년도 추진계획 반영 여부, ⑦ 평가시스템 개선에 평가정보 활용 여부, ⑧ 평가결과 활용 연계장치 구비 여부, ⑨ 평가결과의 평가예산 배정과 연계성, ⑩ 평가결과 보완·권고사항 이행 여부, ⑪ 성과부진 및 문제사업의 개선 여부, ⑫ 평가정보의 접근 가능성 및 정보시스템 구축 여부

바와 같이 기존의 선행연구들이 제시하고 있는 평가요소는 평가목적에 따라 다양하게 제시되고 분석되었으나, 그 요소들의 내용을 분석하면 공통점을 발견할 수 있다. 그것은 평가의 시작단계에서 평가의 종결까지 고려해야 할 요소들을 어떻게 정리하느냐에 따라 분류되는 것이기 때문이다.

본 연구에서는 위에서 살펴본 선행연구에서 활용된 변수를 참조하면서, 본 연구의 주제인 국방연구개발사업의 특성에 공통적으로 적용 할 수 있는 측면들을 고려하여 평가의 전 과정에서 필요한 평가지표를 도출하였다. 즉, 국방연구개발사업 평가시스템 개선을 위한 메타평가 모형은 평가상황요소(evaluation context), 평가투입요소(evaluation input), 평가수행요소(evaluation process), 평가효과요소(evaluation outcome) 등 4개 요소로 설정하였으며, 각 평가요소에 해당하는 세부 평가지표는 평가상황 요소에는 10개의 지표, 평가투입 요소에는 15개의 지표, 평가수행 요소에는 15개의 지표, 평가효과 요소에는 12개의 지표로 구성되어 총 52개의 평가지표로 결정되었다. 평가요소별 구성되는 세부 평가지표는 <표 1>과 같다.

4. 메타평가 요소별 요인분석 및 가중치 결정

4.1 요인분석의 개요

본 연구에서는 요인분석을 실시하는데 있어서 변수들의 상관관계를 이용하여 본래의 변수들이 갖고 있는 의미를 최대한 보존하면서 보다 적은 합성요인으로 요약하는 R-type 요인분석을 적용하였다. 요인의 추출방법으로는 주성분분석(principle component analysis)을, 회전방법으로는 직각회전의 베리맥스(Varimax) 방법을 사용하였다. 각 변수의 요인적재치(factor loading)는 모두 0.5이상인 변수만을 포함하였으며, 요인 추출과정에 있어서는 고유값(eigenvalue) 기준을 적용하여 1.0보다 큰 요인에 대해 요인화 하였다. 또한, 요인범주 내에서 설문문항들의 결합이 신뢰성(reliability)을 갖고 있는지 알아보기 위한 내적 일관성 검증은 Cronbach's α 값을 사용하였으며, 그 값이 0.6이상이면 신뢰성이 있다고 판단하였다.

요인분석을 위한 설문지는 각 문항 Likert 7점 척도 방식으로 작성되었으며, 설문은 국방연구개발사업 자체 평가에 참여하였던 전문가를 대상으로 총 53부의 설문지를 회수하여 결과 분석에 적용하였다.

4.2 평가상황요소 요인분석 결과

평가상황요소에 해당하는 10가지 평가지표에 대하여 요인분석을 실시해 본 결과, <표 2>에서와 같이 3가지 요인

<표 2> 평가상황요소의 요인분석 결과

요인명	평가 지표	Factor 1	Factor 2	Factor 3
평가 목적의 타당성	평가목적의 합리성 여부	.903	.166	.087
	평가목적의 명확성 여부	.902	.155	.287
	평가목적의 국방 가치 부합 여부	.849	.014	.341
평가 제도의 명확성	평가규정의 합리성 여부	.007	.939	-.163
	평가관련 이해관계자 및 수요 파악 여부	.209	.915	-.001
	평가의 법적·제도적 근거 유무	.113	.782	.359
	평가모형의 적절성 구비 여부	.510	.537	.406
평가 대상의 합리성	평가대상사업 선정기준의 명확성 여부	.329	.238	.843
	평가대상사업의 국방 특성 고려 선정 여부	.375	.176	.792
	평가대상 분류의 적정성 여부	.091	-.168	.783
고유값(eigenvalue)		2.924	2.787	2.477
분산설명비율(%)		29.240	27.869	24.773
누적분산설명비율(%)		29.240	57.108	81.881

으로 축약되었다. 요인 1에는 3개 평가지표가 묶여서 '평가목적의 타당성'으로 명명하였으며, 요인 2는 4개의 평가지표를 묶어 '평가제도의 명확성'이라는 요인으로 축약되었다. 요인 3은 3가지 평가지표를 '평가대상의 합리성'이라는 요인으로 명명하였다. 그리고 각 요인의 고유치는 최고 2.924부터 2.477까지 나타났고, 3개의 요인이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 81.881%로 나타났다.

4.3 평가투입요소 요인분석 결과

평가투입요소에 해당하는 15가지 평가지표에 대하여 요인분석을 실시해 본 결과, <표 3>에서와 같이 4가지 요인으로 축약되었다. 요인 1에는 5개 평가지표가 묶여서 '평가인력의 적정성'으로 명명하였으며, 요인 2는 4개의 평가지표를 묶어 '평가자료의 충실성'이라는 요인으로 축약되었다. 요인 3은 4가지 평가지표는 '평가조직의 적합성'이라는 요인으로 명명하였고, 요인 4는 예산관련 평가지표 2가지를 묶어서 '평가예산의 합리성'이라는 요인으로 명명하였다. 그리고 각 요인의 고유치는 최고 4.127부터 2.404까지 나타났고, 4개의 요인이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 88.381%로 나타났다.

<표 3> 평가투입요소의 요인분석 결과

요인명	평가 지표	Factor 1	Factor 2	Factor 3	Factor 4
평가 인력의 적절성	평가위원의 전문성 및 경험 유무	.970	-.164	-.019	.015
	사업규모 고려 평가위원 구성 여부	.882	.181	.001	-.109
	평가위원 구성의 공정성 여부	.849	.126	-.164	-.300
	평가위원의 양적 충분성 여부	.835	.323	.185	.309
	평가위원 사전 교육 실시 여부	.782	.385	-.089	.039
평가 자료의 충실성	평가자료의 양적 충분성 여부	.234	.940	-.153	-.039
	평가자료의 질적 우수성 여부	.242	.889	.016	-.081
	평가자료 활용이 용이한지 여부	.075	.885	-.365	-.270
	평가자료의 적기 제공 여부	.028	.758	-.011	.500
평가 조직의 적합성	평가조직의 평가기능 부합 여부	-.197	.093	.947	-.116
	평가조직과 평가결과의 비관련성 여부	-.097	-.278	.799	-.129
	평가조직의 제도적 독립성 유무	.411	.054	.754	-.376
	평가조직인력의 전문성 및 경험 유무	-.010	-.436	.748	.156
평가 예산의 합리성	평가예산의 안정적 지원 여부	.059	-.056	-.014	.961
	평가예산의 반영 여부	-.193	-.069	-.437	.868
고유값(eigenvalue)		4.127	3.649	3.081	2.404
분산설명비율(%)		27.513	24.300	20.543	16.025
누적분산설명비율(%)		27.513	51.813	72.355	88.381

4.4 평가수행요소 요인분석 결과

평가수행요소에 해당하는 15가지 평가지표에 대하여 요인분석을 실시해 본 결과, <표 4>에서와 같이 4가지 요인으로 축약되었다. 요인 1에는 4개 평가지표를 묶어 ‘평가절차의 적절성’으로 명명하였으며, 요인 2는 4개의 평가지표를 묶어 ‘평가과정의 공정성’이라는 요인으로 축약되었다. 요인 3은 4가지 평가지표를 ‘평가방법의 적합성’이라는 요인으로 명명하였으며, 요인 4는 3가지가 평가지표를 묶어서 ‘평가기준의 합리성’이라는 요인으로

<표 4> 평가수행요소의 요인분석 결과

요인명	평가 지표	Factor 1	Factor 2	Factor 3	Factor 4
평가 절차의 적절성	계획된 평가절차의 일관성 여부	.950	.181	-.215	-.006
	평가절차의 적절한 수립 여부	.938	.029	-.103	-.042
	피평가기관 의견 반영 유무	.904	-.286	-.047	.113
	평가위원회 합리적 운영 여부	.873	-.127	-.165	.159
평가 과정의 공정성	특수 이해 집단과의 연계성 유무	-.161	.969	.077	.096
	기간 중 내·외부의 압력 행사 유무	-.128	.923	.191	.068
	이해관계자간 의사소통 유무	.101	.812	.371	.300
평가 방법의 적합성	평가주관기관의 공정성 여부	.163	.749	.516	.293
	평가등급 산정의 적절성 여부	-.128	.120	.976	-.056
	평가기간 및 시간의 합리성 여부	-.188	.290	.887	-.014
	평가방법의 효율적 수행 여부	-.276	.583	.711	-.235
평가 기준의 합리성	평가방법의 평가목적 부합 여부	-.441	.307	.626	-.299
	평가대상사업 특성 반영 여부	.283	-.036	.193	.920
	평가지표의 양적·질적 적절성 여부	-.010	.180	-.172	.903
평가 기준의 합리성	평가지표의 측정요소별 적절성 여부	-.069	.339	-.278	.860
	고유값(eigenvalue)	3.847	3.824	3.314	2.778
분산설명비율(%)		25.649	25.492	22.096	18.518
누적분산설명비율(%)		25.649	51.141	73.237	91.755

명명하였다. 각 요인의 고유치는 최고 3.847부터 2.778까지 나타났고, 4개의 요인이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 91.755%로 나타났다.

4.5 평가효과요소 요인분석 결과

평가결과요소에 해당하는 12가지 평가지표에 대하여 요인분석을 실시해 본 결과, <표 5>에서와 같이 3가지로 축약되었다. 요인 1에는 5개 평가지표를 묶어 ‘평가결과의 신뢰성’으로 명명하였으며, 요인 2는 4개의 평가지표를 묶어 ‘평가결과의 환류성’이라는 요인으로 축약되었다. 요인 3은 3가지 평가지표를 묶어서 ‘평가결과의 활용성’이라는 요인으로 명명하였다. 각 요인의 고유치는 최고 3.889부터 2.373까지 나타났고, 3개의 요인이 전체평가지표를 설명하

<표 5> 평가효과요소의 요인분석 결과

요인명	평가 지표	Factor 1	Factor 2	Factor 3
평가 결과의 신뢰성	피평가기관의 소명기회 유무	.873	.287	-.043
	평가보고서 작성의 충실성 여부	.863	.363	.036
	평가결과의 적절한 공개 여부	.860	.120	.229
	피평가기관의 평가결과 수용 여부	.826	.175	.375
	평가보고서의 적기 제공 여부	.770	.144	.421
평가 결과의 환류성	성과부진 및 문제 사업의 개선 여부	.226	.892	.038
	보완·반영사항의 이행 여부	.277	.708	.412
	평가정보시스템 구축 및 접근 가능성 여부	.295	.600	-.039
	평가결과의 예산배정 연계성 여부	.083	.538	.183
평가 결과의 활용성	평가결과의 차년도 연구계획에 반영 정도	.191	-.047	.889
	평가시스템 개선 시 평가결과 활용 여부	.330	.387	.771
	평가결과 활용을 위한 연계장치 구비 여부	.023	.549	.640
고유값(eigenvalue)		3.889	2.680	2.373
분산설명비율(%)		32.406	22.337	19.775
누적분산설명비율(%)		32.406	54.743	74.518

는 비율인 설명분산의 누적계수는 74.518%로 나타났다.

4.6 신뢰도 분석 결과

본 연구에 사용된 평가항목에 대한 신뢰도는 Cronbach's α 값으로 분석하였으며, 그 결과는 <표 6>과 같다. 각 평가항목별 신뢰성 분석결과, 평가절차의 적정성 항목이 0.950으로 가장 높은 값을 보여주고 있고, 평가결과의 환류성 항목이 0.762로 가장 낮은 값을 보여주고 있다. 이는 메타평가를 구성하는 전체 평가항목 간의 Cronbach's α 값이 0.6 이상으로서 본 연구의 평가지표는 신뢰할 만한 수준이라고 할 수 있다. 따라서 요인분석 결과에 따른 평가지표를 그대로 사용하는데 큰 무리가 없는 것으로 나타났다.

4.7 평가항목별 가중치 결정

요인분석 결과를 토대로 선정된 평가요소 및 평가항목에 대하여 가중치 결정을 위한 AHP 설문조사를 실시

<표 6> 메타평가 항목별 신뢰도 검증결과

평가요소	평가 항목	Cronbach's α 값
평가상황	평가목적의 타당성	0.918
	평가제도의 명확성	0.860
	평가대상의 합리성	0.824
평가투입	평가인력의 적정성	0.927
	평가자료의 충실성	0.912
	평가조직의 적합성	0.854
	평가예산의 합리성	0.889
평가수행	평가절차의 적정성	0.950
	평가과정의 공정성	0.949
	평가방법의 적합성	0.940
	평가기준의 합리성	0.898
평가효과	평가결과의 신뢰성	0.933
	평가결과의 환류성	0.762
	평가결과의 활용성	0.822

<표 7> 메타평가 요소 및 항목별 가중치

평가 요소(가중치)	평가 항목(가중치)	최종 가중치
평가상황요소 (0.218)	평가목적의 타당성(0.380)	0.083
	평가제도의 명확성(0.279)	0.061
	평가대상의 합리성(0.341)	0.074
평가투입요소 (0.288)	평가인력의 적정성(0.302)	0.087
	평가자료의 충실성(0.189)	0.054
	평가조직의 적합성(0.324)	0.093
	평가예산의 합리성(0.186)	0.054
평가수행요소 (0.262)	평가절차의 적정성(0.249)	0.065
	평가과정의 공정성(0.230)	0.060
	평가방법의 적합성(0.263)	0.069
	평가기준의 합리성(0.259)	0.068
평가효과요소 (0.232)	평가결과의 신뢰성(0.357)	0.083
	평가결과의 환류성(0.324)	0.075
	평가결과의 활용성(0.319)	0.074

하였다. 설문조사는 국방연구개발사업 자체평가 수행 관련 기관, 평가위원 등 25명을 대상으로 실시하였으며, 그 중 일관성 지수(CR ≤ 0.1)를 만족하는 22명의 유효한 자료를 활용하였다. 각 평가요소별, 평가항목별 가중치는 <표 7>과 같다.

5. 국방R&D사업 자체평가의 메타평가 분석

5.1 메타평가 방법

메타평가의 분석방법에 있어서 제 1단계는 평가항목에 대한 평가를, 제 2단계는 평가항목별 가중치를 반영한 평가요소별 평가를, 제 3단계는 평가요소별 점수를 합산한 종합평가를 실시하였다.

평가지표는 ‘전혀 그렇지 않다(1점), 별로 그렇지 않다(2점), 보통이다(3점), 그렇다(4점), 매우 그렇다(5점)’ 등 Likert 5점 척도를 사용하였다. 먼저, 평가항목별 점수는 하위 평가지표들의 평균값을 도출한 후, 이를 평가항목별 가중치의 합에 평가척도 최고 배점인 5를 곱한 점수로 나누어 100점 만점으로 환산하여 산출하였다. 그리고 평가요소별 평가는 가중치를 반영한 각 평가항목별 점수를 합산한 후, 이를 평균하여 산출하였다. 마지막으로 종합평가는 산출된 평가상황, 평가투입, 평가수행, 평가결과 등 4개 평가요소의 환산점수를 합산한 후, 이를 평균하여 산출하였다. 평가점수 기준은 최고점인 100점을 만점으로 설정하고, 93점 이상이면 ‘탁월’, 93점 미만 80점 이상이면 ‘우수’, 80점 미만 60점 이상이면 ‘보통’, 60점 미만 40점 이상이면 ‘미흡’ 그리고 40점 미만인 경우에는 ‘불량’으로 평가하였다.

설문조사 대상은 방위사업청, 국방기술품질원, 국방과학연구소, 항공우주연구원, 개발업체, 평가위원 등 2007년부터 2009년까지 최근 3년 동안 국방연구개발 자체평가에 참여한 경험이 있는 총 230명을 대상으로 하였으며, 그 중 205부를 분석 자료로 활용하였다. 설문지 분석은 사회과학 통계프로그램인 SPSS 14.0을 이용하였다(이학식의, 2008).

5.2 종합평가 결과

국방연구개발사업 자체평가에 대한 메타평가 결과를 종합해 보면, <표 8>과 같이 100점 만점에 74.3점을 획득하여 ‘보통’의 수준으로 평가되었다. 메타평가 4개의 평가요소 별로는 평가상황요소가 가장 높은 80.2점으로 ‘우수’하게 평가되었다. 평가상황다음으로 높은 점수를 받은 요소는 평가수행 요소로 74.8점으로 ‘보통’으로 평가되었다. 그리고 평가투입요소는 73.1점으로 ‘보통’으로 평가되었으며, 평가효과요소가 제일 낮은 점수인 69.0점으로 나타나 ‘보통’으로 평가되었다. 평가효과요소가 다른 평가요소들에 비해서 평가점수가 많이 떨어져 있다. 특히, 평가결과의 활용성과 환류성에서 낮은 평가를 받았는데, 이러한 결과는 평가결과 활용을 위한 연계장치 구

<표 8> 메타평가의 요소별 총점

평가 요소		평균	가중치	가중치 반영점수	100점 환산점수	평가
평가 상황	평가목적의 타당성	4.073	0.083	0.338	81.5	우수
	평가제도의 명확성	4.038	0.061	0.246	80.8	
	평가대상의 합리성	3.911	0.074	0.289	78.2	
	계(평균)	(4.007)	0.218	0.874	80.2	
평가 투입	평가인력의 적정성	3.700	0.087	0.322	74.0	보통
	평가자료의 충실성	3.656	0.054	0.197	73.1	
	평가조직의 적합성	3.721	0.093	0.346	74.4	
	평가예산의 합리성	3.480	0.054	0.188	69.6	
	계(평균)	(3.639)	0.288	1.053	73.1	
평가 수행	평가절차의 적정성	3.680	0.065	0.239	73.6	보통
	평가과정의 공정성	3.973	0.060	0.238	79.5	
	평가방법의 적합성	3.740	0.069	0.258	74.8	
	평가기준의 합리성	3.589	0.068	0.244	71.8	
	계(평균)	(3.746)	0.262	0.980	74.8	
평가 효과	평가결과의 신뢰성	3.510	0.083	0.291	70.2	보통
	평가결과의 환류성	3.411	0.075	0.256	68.2	
	평가결과의 활용성	3.428	0.074	0.254	68.6	
	계(평균)	(3.450)	0.232	0.801	69.0	
총 점			1		74.3	보통

비가 미흡한 실정이며, 평가정보시스템 구축을 통한 접근 가능성이 미흡하다고 할 수 있다.

5.3 평가요소별 평가

5.3.1 평가상황요소

국방연구개발사업 자체평가의 평가상황요소에 대한 메타평가 결과는 <표 9>에서와 같이 100점 만점에 80.2점으로 ‘우수’한 것으로 평가되었다. 평가항목별로는 평가목적의 타당성이 4.073, 평가제도의 명확성이 4.038, 평가대상의 합리성이 3.911로 평균 4.007에 해당하는 비교적 높은 점수이다. 다른 평가항목들에 비해 평가대상의 합리성이 비교적 낮게 나타났는데, 이러한 평가결과는 국방연구개발사업 평가의 필요성 및 가치, 그리고 평가제도의 법적·제도적 근거와 관련 규정에 대해서는 높게 평가하는 반면, 국방분야의 특수성을 고려하여 평가대상을 분류해야 할 필요성이 있음을 보여주고 있다고 해석할 수 있다.

먼저, 평가목적의 타당성에서는 평가목적의 합리성 여부 4.15, 평가목적의 명확성 여부 4.10, 평가목적의 국방연구개발사업 가치 부합 여부 3.97 등으로 평균 4.073로

<표 9> 평가상황 요소의 메타평가 결과

평가 항목		평균	가중치	가중치 반영점수	100점 환산점수	평가
평가 목적의 타당성	평가목적의 합리성 여부	4.15	0.083	0.338	81.5	우수
	평가목적의 명확성 여부	4.10				
	평가목적의 국방연구개발사업 가치 부합 여부	3.97				
	평균	4.073				
평가 제도의 명확성	평가규정의 합리성 여부	4.27	0.061	0.246	80.8	우수
	평가관련 이해관계자 및 수요 파악 여부	4.02				
	평가의 법적·제도적 근거	4.17				
	평가모형의 적절성 구비	3.69				
평균	4.038					
평가 대상의 합리성	평가대상사업 선정기준의 명확성	4.02	0.074	0.289	78.2	보통
	평가대상사업의 국방분야 특수성 고려 선정	3.89				
	평가대상 분류의 적정성	3.82				
	평균	3.911				
총계	4.007	0.218	0.874	80.2	우수	

다소 높아 100점 만점에 81.5점을 얻어 '우수'하게 평가되었다. 평가제도의 명확성에서는 평가규정의 합리성 여부 4.27, 평가관련 이해관계자 및 수요 파악 여부 4.02, 평가의 법적·제도적 근거 유무 4.17, 평가모형의 적절성 구비 여부 3.69 등으로 평균 4.038로 다소 높아 100점 만점에 80.8점을 얻어 '우수'하게 평가되었다. 평가대상의 합리성에서는 평가대상사업 선정기준의 명확성 여부 4.02, 평가대상사업의 국방분야 특수성 고려 선정 여부 3.89, 평가대상 분류의 적정성 여부 3.82 등으로 평균 3.911로 100점 만점에 78.2점을 얻어 '보통'으로 평가되었다.

5.3.2 평가투입요소

평가투입요소에 대한 메타평가 결과는 <표 10>에서와 같이 100점 만점에 73.1점으로 '보통'으로 평가되었다. 평가항목별로는 평가인력의 적정성이 3.700, 평가자료의 충실성이 3.656, 평가조직의 적합성이 3.721, 평가예산의 합

<표 10> 평가투입 요소의 메타평가 결과

평가 항목		평균	가중치	가중치 반영점수	100점 환산점수	평가
평가 인력의 적정성	평가위원의 전문성 및 경험	3.93	0.087	0.322	74.0	보통
	사업규모 고려 평가위원 구성 여부	3.50				
	평가위원 구성의 공정성 여부	4.00				
	평가위원의 양적 충분성 여부	3.72				
	평가위원 사전 교육 실시 여부	3.35				
평균	3.700					
평가 자료의 충실성	평가자료의 양적 충분성 여부	3.75	0.054	0.197	73.1	보통
	평가자료의 질적 우수성 여부	3.66				
	평가자료 활용이 용이한지 여부	3.59				
	평가자료의 적기 제공 여부	3.62				
	평균	3.656				
평가 조직의 적합성	평가조직의 평가 기능 부합 여부	3.82	0.093	0.346	74.4	보통
	평가조직과 평가 결과의 비관련성	3.52				
	평가조직의 제도적 독립성 유무	3.74				
	평가조직 인력의 전문성 및 경험 유무	3.80				
	평균	3.721				
평가 예산의 합리성	평가예산의 안정적 지원 여부	3.51	0.054	0.188	69.6	보통
	평가예산의 반영 여부	3.45				
	평균	3.480				
총계	3.639	0.288	1.053	73.1	보통	

리성이 3.480로 평균 3.639에 해당하는 점수이다. 평가인력과 평가자료, 그리고 평가조직에 있어서는 양호한 평가를 받았음에도 불구하고, 평가예산 지원 및 적정 규모의 반영 측면에서는 시급한 대책이 필요하다는 것을 알 수 있다. 이는 평가예산의 지원 및 규모에서는 아직도 보완해야 할 점이 많은 것으로 해석할 수 있다.

평가인력의 적정성에서는 평가위원의 전문성 및 경험 유무 3.93, 사업규모 고려 평가위원 구성 여부 3.50, 평가위원 구성의 공정성 여부 4.00, 평가위원의 양적 충분성 여부 3.72, 평가위원 사전 교육실시 여부 3.35 등으로 평균 3.700으로 100점 만점에 74.0점을 얻어 ‘보통’으로 평가되었다. 평가 자료의 충실성에서는 평가 자료의 양적 충분성 여부 3.75, 평가 자료의 질적 우수성 여부 3.66, 평가자료 활용이 용이한지 여부 3.59, 평가 자료의 적기 제공 여부 3.62 등으로 평균 3.656으로 100점 만점에 73.1점을 얻어 ‘보통’으로 평가 되었다. 평가조직의 적합성에서는 평가조직의 평가기능 부합 여부 3.82, 평가조직과 평가결과의 비관련성 여부 3.52, 평가조직의 제도적 독립성 유무 3.74, 평가조직 인력의 전문성 및 경험 유무 3.80 등으로 평균 3.721로 100점 만점에 74.4점을 얻어 ‘보통’으로 평가되었다. 평가예산의 합리성에서는 평가예산의 안정적 지원 여부 3.51, 평가예산의 반영 여부 3.45 등으로 평균 3.480로 100점 만점에 69.6점을 얻어 ‘보통’으로 평가되었다.

5.3.3 평가수행요소

평가수행요소에 대한 메타평가 결과는 <표 11>에서와 같이 100점 만점에 74.8점으로 ‘보통’으로 평가되었다. 평가항목별로는 평가절차의 적정성이 3.680, 평가과정의 공정성이 3.973, 평가방법의 적합성이 3.740, 평가기준의 합리성이 3.589로 평균 3.746에 해당하는 점수이다. 평가절차, 평가과정 및 평가방법 등에 있어서는 양호한 평가를 받았으나, 특히, 평가과정 중의 공정성에 대해서는 매우 높은 평가를 받았다. 그러나 평가기준에 대해서는 비교적 양호한 평가를 받지 못하였다.

평가 절차의 적정성에서는 계획된 평가절차의 일관성 여부 3.96, 평가절차의 적절한 수립 여부 3.84, 피평가기관 의견 반영 유무 3.50, 평가기간 및 시간의 합리성 여부 3.42 등으로 평균 3.680로 100점 만점에 73.6점을 얻어 ‘보통’으로 평가되었다. 평가과정의 공정성에서는 특수이해집단과의 연계성 유무 3.97, 기간 중 내·외부의 압력 행사 유무 4.14, 이해관계자간 의사소통 유무 3.68, 평가주관기관의 공정성 여부 4.10 등으로 평균 3.973로 100점 만점에 비교적 높은 점수인 79.5점을 얻어 ‘보통’으로 평가되었다. 평가방법의 적합성에서는 평가팀 구성의 균형 여부 3.93, 평가분석 방법의 적절한 사용 여부 3.55, 평가방법의 효율적 수행 여부 3.71, 평가방법의 평가목적 부합 여부 3.78 등으로 평균 3.740로 100점 만점에 74.8점을 얻어 ‘보통’으로 평가되었다. 평가기준의 합리성에서는 평가지표의 평가대상사업 특성 반영 여부 3.75, 평가지표의 양적·질적 적절성 여부 3.52, 평가지표의 측정요소별 적절성 여부 3.50 등으로 평균 3.589로 100점 만점에 71.8점을 얻어 ‘보통’으로 평가되었다.

<표 11> 평가수행 요소의 메타평가 결과

평가 항목		평균	가중치	가중치 반영점수	100점 환산점수	평가
평가 절차의 적정성	계획된 평가절차의 일관성 여부	3.96	0.065	0.239	73.6	보통
	평가절차의 적절한 수립 여부	3.84				
	피평가기관 의견 반영 유무	3.50				
	평가기간 및 시간의 합리성 여부	3.42				
	평균	3.680				
평가 과정의 공정성	특수이해집단과의 연계성 유무	3.97	0.060	0.238	79.5	보통
	기간 중 내·외부의 압력 행사 유무	4.14				
	이해관계자 간 의사소통 유무	3.68				
	평가주관기관의 공정성 여부	4.10				
	평균	3.973				
평가 방법의 적합성	평가팀 구성의 균형 여부	3.93	0.069	0.258	74.8	보통
	평가분석 방법의 적절한 사용 여부	3.55				
	평가방법의 효율적 수행 여부	3.71				
	평가방법의 평가목적 부합 여부	3.78				
	평균	3.740				
평가 기준의 합리성	평가지표의 평가대상사업 특성 반영	3.75	0.068	0.244	71.8	보통
	평가지표의 양적·질적 적절성 여부	3.52				
	평가지표의 측정요소별 적절성	3.50				
	평균	3.589				
총 계		3.746	0.262	0.980	74.8	보통

점을 얻어 ‘보통’으로 평가되었다.

5.3.4 평가효과요소

평가효과요소에 대한 메타평가 결과는 <표 12>에서와 같이 100점 만점에 69.0점으로 비교적 낮은 점수의 ‘보통’으로 평가되었다. 평가항목별로는 평가결과의 신뢰성이 3.510, 평가결과의 환류성이 3.411, 평가결과의 활용성

〈표 12〉 평가효과 요소의 메타평가 결과

평가 항목		평균	가중치	가중치 반영점수	100점 환산점수	평가
평가 결과의 신뢰성	피평가기관의 소명 기회 유무	3.57	0.083	0.291	70.2	보통
	평가보고서 작성의 충실성 여부	3.51				
	평가결과의 적절한 공개 여부	3.51				
	평가보고서의 의사결정권자 보고 여부	3.52				
	평가보고서의 적기 제공 여부	3.46				
	평 균	3.510				
평가 결과의 환류성	성과부진 및 문제사업의 개선 여부	3.49	0.075	0.256	68.2	보통
	보완·반영사항의 이행 여부	3.49				
	평가정보시스템 구축 및 접근 가능성	3.33				
	평가결과의 예산배정 연계성 여부	3.34				
	평 균	3.411				
평가 결과의 활용성	평가결과의 차년도 연구계획에 반영 정도	3.51	0.074	0.254	68.6	보통
	평가시스템 개선 시 평가결과 활용 여부	3.45				
	평가결과 활용을 위한 연계장치 구비 여부	3.33				
	평 균	3.429				
총 계		3.450	0.232	0.801	69.0	보통

이 3.429으로 평균 3.45에 해당하는 점수이다. 평가효과 요소의 평가항목들이 다른 평가요소 및 평가항목들에 비해 비교적 낮게 평가되었다. 특히, 평가정보시스템 구축을 통한 정보의 접근 가능성과 평가결과의 예산배정 연계성 여부 그리고 평가결과 활용을 위한 연계장치의 적절한 구비 여부 등이 비교적 낮은 평가를 받았다. 평가결과의 신뢰성에서는 피평가기관의 소명기회 유무 3.57, 평가보고서 작성의 충실성 여부 3.51, 평가결과의 적절한 공개 여부 3.51, 평가보고서의 의사결정권자 보고 여부 3.52, 평가보고서의 적기 제공 여부 3.46 등으로 평균 3.510로

100점 만점에 70.2점을 얻어 ‘보통’으로 평가되었다. 평가 결과의 환류성에서는 성과부진 및 문제사업의 개선 여부 3.49, 보완·반영사항의 이행 여부 3.49, 평가정보시스템 구축 및 접근 가능성 여부 3.33, 평가결과의 예산배정 연계성 여부 3.34 등으로 평균 3.411로 100점 만점에 비교적 낮은 점수인 68.2점을 얻어 ‘보통’으로 평가되었다. 평가결과의 활용성에서는 평가결과의 차년도 연구계획에 반영 정도 3.51, 평가시스템 개선 시평가결과 활용 여부 3.45, 평가결과 활용을 위한 연계장치 구비 여부 3.33 등으로 평균 3.429으로 100점 만점에 비교적 낮은 점수인 68.6점을 얻어 ‘보통’으로 평가되었다.

6. 결론 및 제언

본 연구에서는 국방연구개발사업 자체평가에 대한 메타평가를 실시해 본 결과 100점 만점에 74.3점을 획득하여 ‘보통’의 수준으로 평가되었다. 이러한 결과는 국방연구개발사업 자체평가를 3년째 시행해 왔지만 아직까지 개선해야 할 사항이 많으며, 향후에도 지속적인 평가시스템 개선이 필요한 것으로 사료된다.

국방연구개발사업 자체평가에 대한 메타평가 결과 평가상황요소는 80.2점으로 ‘우수’한 것으로 평가되었다. 평가항목별로는 평가목적의 타당성이 4.073, 평가제도의 명확성이 4.038, 평가대상의 합리성이 3.911로 평균 4.007에 해당하는 비교적 높은 점수이다. 평가투입요소에 대한 메타평가 결과는 73.1점으로 ‘보통’으로 평가되었다. 평가항목별로는 평가인력의 적정성이 3.700, 평가자료의 충실성이 3.656, 평가조직의 적합성이 3.721, 평가예산의 합리성이 3.480로 평균 3.639에 해당하는 점수이다. 평가수행요소에 대한 메타평가 결과는 74.8점으로 ‘보통’으로 평가되었다. 평가항목별로는 평가절차의 적정성이 3.680, 평가 과정의 공정성이 3.973, 평가방법의 적합성이 3.740, 평가기준의 합리성이 3.589로 평균 3.746에 해당하는 점수이다. 평가효과 요소에 대한 메타평가 결과는 69.0점으로 비교적 낮은 점수의 ‘보통’으로 평가되었다. 평가항목별로는 평가결과의 신뢰성이 3.510, 평가결과의 환류성이 3.411, 평가결과의 활용성이 3.429으로 평균 3.450에 해당하는 점수이다.

본 연구는 그 동안 국방연구개발사업 자체평가에 대한 메타평가 논의가 거의 이루어지지 않은 시점에서 국방분야 메타평가 모형을 설계함으로써 메타평가의 새로운 영역을 개척하는데 기여하였다. 그리고 메타평가에 대한 문헌연구 뿐만 아니라 메타평가를 위한 세부 설문지를 작성하여 통계적인 방법을 사용하여 국방연구개발사업 자체평가에 대한 메타평가를 실시하였다는 점에서

는 매우 의의가 크다고 할 수 있다. 특히, 메타평가 항목을 도출하는데 평가요소별로 구성되는 평가지표들에 대해서 요인분석을 실시하여 평가항목을 도출한 점과, 이러한 평가항목에 AHP 기법을 적용하여 평가항목별 가중치를 결정하여 메타평가 설문분석에 이용한 점은 메타평가와 관련한 타 연구들과 비교하여 독창성과 차별성을 지닌다고 할 수 있겠다.

향후 본 연구결과를 토대로 국방연구개발사업 자체평가 시스템 전반에 대한 개선방안이 지속적으로 모색되어야 할 것이며, 평가결과의 활용성 및 환류성 증진 방안도 함께 모색되어야 할 것이다. 아울러 본 연구를 토대로 국방연구개발사업 자체평가 발전방안들이 활발하게 논의되고 평가시스템이 합리적인 방향으로 개선됨으로써, 평가시스템의 조기 정착 및 연구개발사업 성과 극대화에 크게 기여할 것으로 기대된다.

참고문헌

- [1] 과학기술부; “국가연구개발사업 등의 성과평가 및 성과 관리에 관한 법률”, 법률 제7808호, 2005.
- [2] 과학기술혁신본부; “2008년도 국가연구개발사업 자체평가매뉴얼”, 2008.
- [3] 김병철, “정부출연연구기관 평가에 대한 메타평가 모형 설계 및 적용”, 박사학위논문, 서울시립대학교 대학원, 2008.
- [4] 노유진, “우리나라 정부부처에 있어서 정책평가결과의 저활용요인에 관한 연구”, 행정논총, 43(4) : 245-276, 2005.
- [5] 류영수; “기술영향평가의 메타평가 모형개발 및 적용”, 박사학위논문, 한양대학교 대학원, 2007.
- [6] 박중수; “정보화 사업 메타평가를 위한 CIPOU 모형 개발”, 박사학위논문, 충남대학교 대학원, 2003.
- [7] 방위사업청; “방위력개선사업관리규정”, 2006.
- [8] 이학식, 임지훈; “SPSS 14.0 매뉴얼”, 법문사, 2008.
- [9] 이형우; “국가연구개발사업 평가체계에 관한 소고”, 정책분석평가학회보, 15(2) : 191-219, 2005.
- [10] 이찬구; “영국의 범부처간 과학기술정책 평가제도-메타평가와 한국에의 시사점”, 한국행정논문집, 15(4) : 809-833, 2003.
- [11] American Evaluation Association “Guidelines for Evaluator,” by Shadish W. R., D. L. Newman, M. A. Scheirer and C.; Wye, San francisco, Jossey-Bass, 1995.
- [12] Chelimsky, Eleanor; “Program Evaluation: Patterns and Directions,” 1-35, Washington DC, *The American Society for Public Administration*, 1985.
- [13] Cook T D. and Gruder C. L.; “Metaevaluation Research,” *Evaluation Quarterly*, 2(1) : 5-55, 1978.
- [14] Hong H. D.; “Meta-evaluation of National Large-Scale R&D Programs : A Comparison of Evaluation Systems of 6 National R&D Programs,” *University of Manchester*, 2000.
- [15] Joint Committee on Standards for Educational Evaluation; “The Programme Evaluation Standards,” *Thousand Oaks, Sage Publications*, 1994.
- [16] Larson, Richard C and Leni Berliner; “On Evaluating Valuations,” *Policy Science*, 16(2) : 147-163, 1983.
- [17] Leviton, Laura C and Edward F. X. Hughes; “Research on the Utilization of Evaluations : A Review and Synthesis,” *Evaluation Review*, 5(4) : 1981.
- [18] OECD, “Improving Evaluation Practices-Best Practice Guidelines for Evaluation and Background Paper,” *Paris*, 1999.
- [19] Scriven M.; “Evaluation Thesaurus(4th ed.),” Newbury Park, CA, Sage Publications, 1991.
- [20] Stevenson, John F., Richard H. Longabaugh and Dwight N. Mcneil; “Metaevaluation in Human Services,” *The Evaluator and Management*, Beverly Hills, CA, Sage Publications, 2000.
- [21] Stufflebeam L.; “Metaevaluation : Concepts, Standard and Uses,” *Educational Evaluation Methodology*, The JHU Press, 1981.
- [22] Worthen, Blaine R., James R. Sanders and Jody L.; Fitzpatrick, “Program Evaluation: Alternative Approaches and Practical Guidelines,” Second Edition, White Plains, Longman Publishers, 1997.
- [23] Yi, Chan Goo, “The National R&D Evaluation System in the UK : Meta-evaluation and Applicability to the Korea Case,” Thesis for the degree of PhD, PREST, University of Manchester, 2003.