

국방연구개발사업 자체평가시스템의 메타평가 모형분석 및 검증

김순영¹, 하규수^{2*}

¹국방기술품질원, 호서대학교 벤처전문대학원

²호서대학교 벤처전문대학원

A Meta Evaluation for the Self Evaluation System of Defense R&D Programs

Soon-Yeong Kim¹ and Kyu-Soo Ha^{2*}

¹Defense Agency for Technology and Quality, The Graduate School of Venture, Hoseo University

²The Graduate School of Venture, Hoseo University

요약 본 논문은 국방연구개발사업 자체평가시스템에 대한 메타평가 실증분석을 통하여 자체평가시스템 개선 및 제도발전을 도모하고, 국방예산의 비효율성을 예방하며, 연구 성과를 효율적으로 관리하기 위한 것이다. 국방연구개발사업에 대한 메타평가결과 전체적인 총점은 74.3점을 획득하여 '보통' 수준으로 나타났다. 그리고 집단 간 차이분석을 실시한 결과 일부 평가항목에서 유의미한 차이점이 나타났다. 향후에도 본 연구결과를 토대로 국방연구개발사업 자체평가 시스템 전반에 대한 개선방안이 지속적으로 고려되어야 할 것이며, 특히, 평가결과의 활용성 및 환류성 증진 방안도 함께 모색되어야 할 것이다.

Abstract This study is the result of meta evaluation for the self evaluation of defense R&D programs in Korea. The overall meta evaluation result gained 74.3 points out of 100, so it was evaluated as 'Good'. T- test and one-way ANOVA were used to identify the difference between groups. As a result of difference analysis between groups, some evaluation items showed difference within significance level. From this 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 performances.

Key Words : Defense R&D Programs, Performance Evaluation, Meta Evaluation

1. 서론

과학기술을 중심으로 세계 경쟁체제의 강화가 가속화되고 있는 가운데 연구개발이 국가 경쟁력의 핵심요소로 부각되고 있다. 그리고 2009년도 우리나라 국가 연구개발예산은 12조원을 넘어서는 등 연구개발예산이 크게 증가하고 있는 추세다. 그러나 최근에는 이러한 규모의 성장을 넘어, 국가연구개발사업의 효율적이고 효과적인 수행을 위해 성과관리의 중요성이 크게 부각되고 있는 상황이다.

정부에서는 국가 연구개발의 효율성과 생산성을 제고

하기 위하여 2005년 12월에 '국가연구개발사업 등의 성과평가 및 성과관리에 관한 법률'을 제정하였으며, 이에 따라 각 부처는 소관 연구개발사업에 대한 조사·분석 평가를 실시하고 있다. 연구개발사업은 과학기술정책의 틀 속에서 중·장기적으로 대규모의 예산이 투입되기 때문에, 그것의 성공이나 실패여부는 국가경제에 미치는 영향이 매우 크다고 볼 수 있다. 따라서 최근의 과학기술 환경변화에 국방연구개발사업이 능동적으로 잘 대응하고 있는지를 평가하는 것은 매우 중요하다.

국방연구개발사업 자체평가의 주된 목적은 평가의 객관성 및 공정성을 제고하고, 연구자 중심의 투명한 평가

*교신저자 : 하규수(ksh@hoseo.edu)

접수일 09년 06월 04일

수정일 09년 07월 23일

게재확정일 09년 08월 19일

제도를 정착시키며, 엄정한 평가를 통한 효율적인 사업관리 체계를 구축하는데 있다. 이런 점을 고려해볼 때, 본 연구의 필요성은 국방연구개발사업의 평가가 제대로 수행되고 있는지를 확인하고, 시행 3년을 맞는 현 시점에서 국방연구개발사업 자체평가 시스템의 개선사항을 도출하여 반영하여 국방연구개발사업을 보다 효율적으로 관리하고 연구성과에 대한 신뢰성과 객관성을 제고시킬 수 있도록 하는 것이다.

본 연구와 기존의 메타평가 연구와의 차별성으로는, 국방분야 연구개발사업의 자체평가시스템에 대하여 처음으로 객관적인 메타평가 모형을 분석하고 평가지표를 개발하여 실증분석을 실시한 점이다. 또한, 메타평가 항목을 도출하는데 있어서는 요인분석을 실시하여 평가항목 간 신뢰성 및 타당성을 입증하였고, AHP기법을 적용하여 평가항목별 가중치를 객관적으로 산정하여 적용하였다.

따라서 본 연구의 목적은 국방연구개발사업 자체평가에 대한 메타평가 모형을 분석하고 실증 연구를 통하여 자체평가시스템 개선 및 제도발전을 도모하고, 나아가 국방예산의 비효율성을 예방하고 선진 무기체계를 성공적으로 개발하여 국방력을 강화하는 것이다.

2. 이론적 고찰

2.1 국방연구개발사업 특성 및 평가 시스템

2.1.1 국방연구개발사업 특성

국방연구개발은 무기체계를 구성하는 주요 기술이나 구성품 등과 같은 제반 요소들을 개발함으로써 궁극적으로 국내에서 무기체계를 생산하여 획득하고자 하는 일련의 활동을 의미한다[8].

국방연구개발을 사업추진방법으로 구분하면, 방위력 개선사업은 무기체계 및 핵심기술 연구개발사업과 무기체계 구매사업으로 구분한다. 그리고 국방과학기술 경쟁력 제고 및 국가자원의 효율적인 활용을 위하여 정부 관련 부처와 공동으로 민군 겸용기술 사업이나 국책사업을 추진할 수 있으며, 민군 겸용기술 사업의 경우 사업비의 부담 및 추진절차·사업관리 등에 관한 사항은 민군 겸용기술 사업촉진 법령에 따르고, 기타 국책사업에 관한 제반사항은 부처의 합의에 의해 제정된 별도의 법령에 따른다.

국방연구개발은 ‘대내·외적인 위협을 제거하기 위해 국방에 필요한 무기체계, 자동화체계에 관한 과학적 기술적 조사, 연구, 개발 및 시험 등을 행하는 일련의 과정’이

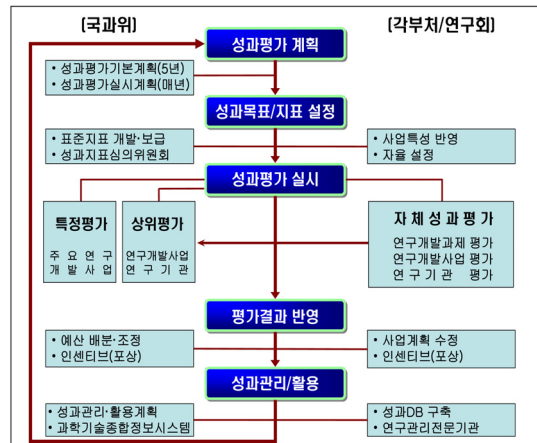
라고 정의 할 수 있는데, 그 범주는 국방력 형성에 관련된 전 분야를 망라한 것이라 하겠다.

2.1.2 국방연구개발사업 평가 시스템

우리나라에서 정부 예산으로 지원되고 있는 국가연구개발 활동에 대한 평가는 국가과학기술위원회(국과위)에서 실시하는 특정평가와 상위평가, 그리고 부처에서 자체적으로 시행하는 자체성과평가로 구성되며, 세부 평가절차는 [그림 1]과 같다.

국방연구개발사업 자체평가는 국가 연구개발사업 평가시스템 틀 속에서 수행되고 있다. 부처 자체평가는 연구개발사업을 수행하는 각 중앙행정기관은 당해 연도 특정평가 대상이 아닌 사업 중 착수 후 1년 이상 경과된 사업에 대해서는 사업특성에 따른 성과중심의 자체평가를 실시한다. 중앙행정기관들은 평가대상 사업별로 전략목표와 성과목표를 설정한 후, 국과위의 표준성과지표를 활용하여 사업특성을 반영한 성과지표를 자율적으로 설정하여 평가를 실시한다.

자체평가 결과는 차년도 국방연구개발사업에 대한 국방예산의 지정 및 배분에 반영하고, 또한, 연구개발 사업을 추진하는 관련 기관에서는 자체평가결과를 반영하여 연구개발사업 등의 추진계획을 수정·보완하고, 과학적인 성과평가 지표개발 및 평가시스템상의 제도개선 사항을 도출하여 반영할 수 있다. 국방연구개발사업의 자체평가 대상은 국과위연구개발사업, 업체주관연구개발사업 및 국책연구개발사업 등으로 관리되고 있다.



[그림 1] 국가연구개발사업 등의 성과평가 및 성과관리 절차

한편, 특정평가는 국과위의 심층적인 성과평가 기능을 강화하기 위하여 주요 사업에 대한 평가를 실시하는 것

이며, 상위평가는 자체평가 성과목표/지표의 적절성 및 자체평가 방법의 객관성 공정성 등을 점검하는 것이다.

2.2 메타평가 선행연구 사례분석

메타평가의 등장은 평가결과의 활용과 직접적인 관련이 있다[23]. 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[17]. 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[2].

메타평가(meta evaluation)는 일종의 평가에 대한 평가로 평가의 효율성을 평가하기 위한 수단이다. 메타평가의 모델을 제안한 외국 학자들[16,21,24]과 주요 평가기관의 가이드라인[12,14,18], 그리고 국내 사례 연구[5,7,11,15]에서 제시하고 있는 메타평가 요소들을 정리하면 [표 1]과 같다.

[표 1] 메타평가에 대한 선행연구

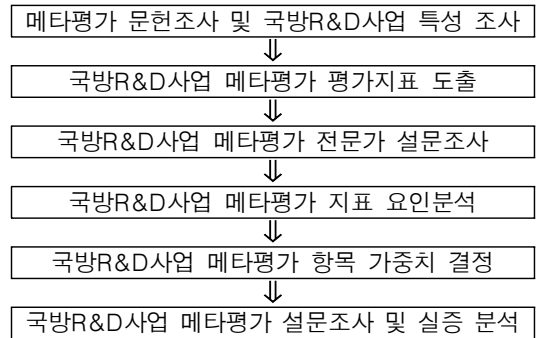
| 구분 | 평가요소 |
|-------------------------|----------------------------------|
| Hong(2000) | 평가상황, 집행배너즘, 평가결과, 결과활용 |
| 이찬구(2003) | 평가기준, 평가자원, 평가수행, 평가활용 |
| 박종수(2003) | 평가상황, 평가투입, 평가수행, 평가결과, 평가활용 |
| 류영수(2007) | 평가기획, 평가투입, 평가수행, 평가결과, 평가활용 |
| Sufflebean(1981) | 적절성 기준, 유용성 기준, 효율성 기준 |
| Larson & Berliner(1983) | 평가투입요소, 평가과정요소, 평가결과요소 |
| Scriven(1991) | 서론, 기만, 하위평가, 결론 및 함의 |
| Joint Committee(1994) | 중요성, 실현가능성, 타당성, 정확성 |
| AEA(1995) | 체계적 조사, 능력, 통합성/정직, 인격존중, 공공책임 |
| OECD(1999) | 평가로부터의 수집, 평가 틀의 조직화, 효과적인 평가 수행 |

2.3 연구 절차 및 가설의 설정

2.3.1 연구 절차

국방연구개발사업의 메타평가 지표를 개발하기 위하여 먼저 메타평가 선행연구 및 브레인스토밍을 통하여

질문항목을 도출하였다. 선정된 질문항목들은 국방연구개발사업 자체평가의 전문 평가위원들의 의견을 반영하여 더욱 정제화 하였다. 그리고 메타평가 요소별 도출된 평가지표를 검증하고 보다 효율적으로 분석하고 평가항목 선정에 위해 요인분석을 실시하였으며, 평가항목별 가중치 산정은 AHP기법을 적용하였다. 특히, 국방연구개발사업 자체평가 이해관계자 모두의 의견이 다양하게 수렴되어 평가지표 개발의 객관성과 신뢰성을 제고하도록 하였다. 본 논문의 연구 절차는 [그림 2]와 같다.



[그림 2] 연구 절차

2.3.2 연구 가설의 설정

1) 평가환경요소

평가에 있어서 합리적인 목적이란 책무성 확보, 프로그램 향상, 지식의 창출 등을 들 수 있는데[1,3,19,20,25], 이러한 평가목적은 평가규정이나 평가실시기관의 평가계획 등에 명료하게 드러나 있음으로서 평가주체들이 인식할 수 있어야 한다. 평가의 경우 제안된 행동노선과 관련된 수단뿐만 아니라 목적 그 자체에 초점을 두는 목적 지향적 성격을 지니기 때문에, 정책결정자가 명시적으로 천명한 행위 외에도 표명한 목표(goal)와 관련된 목적(objectives)까지도 의문을 가지게 된다[3].

평가상황을 뒷받침하기 위해 평가에 대한 법적·제도적 장치가 마련되어 있어야 한다. 아울러 평가에 대한 기본계획을 수립하고 수립된 기본계획을 바탕으로 시행계획이 작성되어 실제 평가를 위한 지침으로 제공되어야 한다. 평가의 주안점은 평가의 목적분류와 그에 상응하는 질문을 구성하는데 있으며, 각 질문에 대한 답에 이용될 방법론을 확인하고 실행 가능한 계획으로 만들어낼 수 있어야 한다[20].

평가대상 프로그램은 일정한 선정기준이 공식적 근거가 된다. 이러한 선정기준은 내·외부적 환경요인, 특히, 이해관계자들의 정치적 의도에 따라 쉽게 변경되거나 용인되지 않아야 한다[9].

가설 1 : 국방연구개발사업 메타평가의 평가환경 요소는 집단 간 차이가 있을 것이다.

- 1-1) 평가목적이 타당하다는 판단에 차이가 있을 것이다.
- 1-2) 평가제도가 명확하다는 판단에 차이가 있을 것이다.
- 1-3) 평가대상이 합리적이라는 판단에 차이가 있을 것이다.

2) 평가투입요소

평가와 관계되는 인력은 다양하나, 그 중 가장 평가에 영향을 미치는 것은 평가에 직접적으로 관여하는 평가전문가라고 할 수 있다. 이들 평가인력은 능률성을 확보할 수 있는 범위 내에서 양적으로 충분하여야 한다. 평가는 개인적 차원의 업무가 아니고, 조직적인 업무분담을 요구하며, 또한 평가인력은 동일하더라도 계층단계가 높은 경우와 낮은 경우 등 조직적 구조에 따라 실무인원에서는 크게 차이가 발생할 수 있다[6].

일반적으로 평가조직은 공식적으로 평가에 대한 권한과 책임이 있는 전담 평가조직을 설치하여 운영하게 된다. 평가의 전문성과 효율성을 높이기 위해 위원회를 구성하여 함께 운영되며[7], 프로그램 또는 정책의 중요성, 규모, 파급효과 등이 클수록 더욱 강조된다.

평가자료는 질적인 측면에서 활용가치가 있어야 한다. 정책연구에서 정보는 단순한 지식의 전달 이상의 의미를 포함하는데[25], 평가자료 역시 이러한 기준을 확보하여야 한다.

메타평가에 소요되는 비용은 평가실시자의 인건비와 평가관련 비용, 평가전문가의 활용비 등이라고 할 수 있는데[22], 이외 소요되는 예산은 적정수준의 것이어야 한다.

가설 2 : 국방연구개발사업 메타평가의 평가투입 요소는 집단 간 차이가 있을 것이다.

- 2-1) 평가인력이 적합하다는 판단에 차이가 있을 것이다.
- 2-2) 평가조직이 적합하다는 판단에 차이가 있을 것이다.
- 2-3) 평가자료가 충실하다는 판단에 차이가 있을 것이다.
- 2-4) 평가예산이 합리적이라는 판단에 차이가 있을 것이다.

3) 평가수행요소

평가에서 중요한 것은 하나의 프로그램이나 정책에 대

한 평가를 누가하는 것이 바람직한가 하는 평가조직의 구성에 관한 문제이다. 일반적으로 외부적 평가가 내부적 평가보다 객관적 의도로서 더 큰 신뢰를 얻는다[25].

평가는 프로그램이나 정책의 목적을 나타내는 측정수단을 식별하는데서 출발한다[4]. 즉, 평가목적에 부합하는 평가기준을 개발하여야 한다. 이러한 평가기준들은 다양하게 제시되고 있으나, 개별적인 프로그램의 특성을 반영하여 작성하는 것이 필요하다.

평가는 그 유형이나 목적에 따라 실제로 평가를 담당하고 있는 사람들이나 평가를 의뢰하는 의뢰인들 간에는 평가절차에 대하여 각기 다른 의견을 가질 수 있고, 이것이 실제 정책을 평가하는 과정에서 상호간에 불신과 불편한 관계를 가져올 수 있는 소지가 될 수 있으며, 평가의 질을 떨어뜨리고 비능률적인 평가작업의 원인이 될 수도 있다[4].

모든 평가는 정치적 요소(political forces)들로부터 영향을 받으며, 이러한 요소들이 의미 있고 공정한 평가를 저해하는지를 살펴볼 필요가 있다[15].

가설 3 : 국방연구개발사업 메타평가의 평가수행 요소는 집단 간 차이가 있을 것이다.

- 3-1) 평가방법이 적합하다는 판단에 차이가 있을 것이다.
- 3-2) 평가기준이 합리적이라는 판단에 차이가 있을 것이다.
- 3-3) 평가절차가 적정하다는 판단에 차이가 있을 것이다.
- 3-4) 평가과정이 공정하다는 판단에 차이가 있을 것이다.

4) 평가효과요소

평가결과는 여기에 참여한 평가집단들의 합의가 필요하다. 즉, 평가결과는 평가집단의 공통된 의견이 종합적으로 정리되어야 하는데, 평가집단의 수가 많은 경우 이외에도, 일정한 위계질서를 지니거나 분야별로 구분되는 경우, 특히, 피평가기관의 의견을 적절하게 반영하여 이러한 합의를 이끄는데 많은 노력이 요구된다. 평가의 활용을 제고하고 평가결과를 능률적으로 환류하기 위해서는 평가보고서의 질적 수준이 일차적으로 요구된다[10].

평가의 활용여부는 정도의 문제로서, 완전한 활용과 완전한 불용이라는 양극단 사이의 연속선상에서 관찰될 수 있는 상대적인 성질의 것이다[1].

국가 연구개발사업의 평가에서 성과 중심의 평가시스템이 제대로 정착되려면 우리나라 공공부문에서 나타나는 문제점을 개선할 필요가 있다[4]. 평가의 주된 관심이

수단적 활용을 극대화하려는데 집중되고 있다는 것을 부인하기 어렵다[3]. 즉, 평가를 수행하는 가장 중요한 이유는 프로그램의 실행 가능성을 검토하고 운용실태와 효과를 파악함으로써 그 정보를 다음 정책을 결정할 때 환류시키는데 있다[10].

가설 4 : 국방연구개발사업 메타평가의 평가효과 요소는 집단 간 차이가 있을 것이다.

- 4-1) 평가결과가 신뢰성 있다는 판단에 차이가 있을 것이다.
- 4-2) 평가결과가 활용되고 있다는 판단에 차이가 있을 것이다.
- 4-3) 평가결과가 환류되고 있다는 판단에 차이가 있을 것이다.

2.4 국방연구개발사업의 메타평가 지표개발

국방연구개발사업의 메타평가 지표를 개발하기 위하여 위에서 살펴본 선행연구에서 활용된 변수를 참조하면서, 본 연구의 주제인 국방연구개발사업의 특성에 공통적으로 적용할 수 있는 측면들을 고려하여 평가의 전 과정에서 필요한 평가지표 구성요소를 도출하였다. 즉, 국방연구개발사업 평가시스템 개선을 위한 메타평가 모형은 [표 2]와 같이 평가상황요소(evaluation context), 평가투입요소(evaluation input), 평가수행요소(evaluation process), 평가효과요소(evaluation outcome) 등 4개의 평가요소와 52개의 평가지표로 구성되었다.

본 논문에서는 국방연구개발사업 자체평가에 대한 메타평가의 모형을 dCIPO(defense Context Input Process Outcome)모델이라고 명명한다.

[표 2] 메타평가 세부지표 구성(안)

| 평가요소 | 세부 평가지표 구성요소 |
|------|--|
| 평가상황 | 평가목적의 합리성, 명확성, 국방사업 가치 부합 여부, 평가대상사업 분류의 적절성, 평가대상사업 분류의 적절성, 평가대상사업 기준의 명확성, 평가대상사업의 국방분야 특수성 고려, 평가과연 이해관계자 및 수요 파악, 법적·제도적 근거의 타당성, 평가규정의 합리성, 평가모형의 적절성 |
| 평가투입 | 평가인력의 양적 충분성, 평가위원의 전문성 및 경험, 평가대상을 고려한 평가위원 수 배정, 평가위원 구성의 공정성, 평가위원 사전 교육의 필요성 여부, 평가조직의 제도적 독립성, 평가조직의 평가결과 비관련성, 평가조직의 평가기능 부합 여부, 평가조직 인력의 전문성, 평가자료의 양적 충분성, 평가자료의 질적 우수성, 평가자료 제공의 적시성, 평가자료 활용의 용이 여부, 평가예산 규모의 적정성, 평가예산의 안정적 지원 여부 |

| | |
|------|--|
| 평가수행 | 평가목적과 평가방법의 부합성, 평가방법의 효율적 수행, 평가팀 구성의 균형 여부, 평가분석방법의 사용 여부, 평가지표의 양적·질적 적절성, 평가지표의 평가대상사업 특성 반영 여부, 평가지표의 측정요소별 적절성, 적절한 평가절차 수립 여부, 평가절차의 일관성 여부, 피평가기관의 의견 반영 여부, 평가기간 및 시간의 합리성, 특수 이해집단과의 연계성 여부, 평가수행중 내·외부의 압력 행사 여부, 평가주관기관의 공정성, 이해관계자간 의사소통 여부 |
| 평가효과 | 평가결과 공개 여부, 피평가기관의 소명 기회 여부, 평가보고서의 의사결정권자 보고 여부, 평가보고서의 충실성, 평가보고서의 적기 제공 여부, 평가결과와 차년도 추진계획 반영 여부, 평가시스템 개선에 평가정보 활용 여부, 평가결과 활용 연계장치 구비 여부, 평가결과와 평가예산 배정과 연계성, 평가결과 보완·권고사항 이행 여부, 성과부진 및 문제사업의 개선 여부, 평가정보의 접근 가능성 및 정보시스템 구축 여부 |

3. 신뢰성, 타당성 분석 및 가중치 결정

3.1 신뢰성 분석

본 연구에서 사용하고 있는 평가지표들에 대해서 신뢰성(reliability) 분석을 실시하였다. 신뢰성 분석에는 다양한 방법이 있지만 본 연구에서는 크론바알파값(Cronbach's α)을 이용한 내적일관성 측정방법을 사용하였다. Cronbach's α 값이 0.6이상이면 문항들 간의 신뢰도가 양호한 것으로 볼 수 있다.

국방연구개발사업 메타평가를 위한 평가요소에 대한 신뢰도 분석결과, 평가상황요소의 Cronbach's α 값은 0.877로 비교적 높게 나타났으며, 평가투입요소 및 평가수행요소의 Cronbach's α 값은 각각 0.755와 0.792로 나타났다. 그리고 평가효과요소의 Cronbach's α 값은 0.906으로 매우 높게 나타났다. 메타평가를 구성하는 전체 평가항목 간의 Cronbach's α 값도 0.6이상으로 나타나 본 연구의 평가지표는 신뢰할 만한 수준이라고 할 수 있다. 신뢰도 분석 결과는 [표 3]과 같다.

[표 3] 신뢰도 분석 결과

| 평가요소 | 평가항목 | 분항수 | Cronbach's α |
|------|-----------|-----|---------------------|
| 평가상황 | 평가목적의 타당성 | 3 | .918 |
| | 평가제도의 명확성 | 4 | .860 |
| | 평가대상의 합리성 | 3 | .824 |
| 평가투입 | 평가인력의 적정성 | 5 | .927 |
| | 평가자료의 충실성 | 4 | .912 |
| | 평가조직의 적합성 | 4 | .854 |
| | 평가예산의 합리성 | 2 | .889 |

| | | | | |
|------|-----------|---|------|------|
| 평가수행 | 평가절차의 적정성 | 4 | .950 | .792 |
| | 평가과정의 공정성 | 4 | .949 | |
| | 평가방법의 적합성 | 4 | .940 | |
| | 평가기준의 합리성 | 3 | .898 | |
| 평가효과 | 평가결과의 신뢰성 | 5 | .933 | .906 |
| | 평가결과의 환류성 | 4 | .762 | |
| | 평가결과의 활용성 | 3 | .822 | |

3.2 타당성 분석

수집한 자료들의 측정상의 오류를 제거하고 평가지표들이 과연 원래의 모집단을 대표하고 있는지의 여부를 살펴보고자 타당성(validity) 검증을 실시하였다. 본 연구에서는 변수들의 상관관계를 이용하여 본래의 변수들이 갖고 있는 의미를 최대한 보존하면서 보다 적은 합성요인으로 요약하는 R-type 요인분석(Factor Analysis)을 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.

요인분석은 통계적인 기법을 활용하여 타당성을 확보하는 것이 적절하므로, 국방연구개발사업을 추진하고 있는 일련의 관련 기관인 방사청, 국과연, 기품원, 산학연 및 군 소속 전문가를 대상으로 설문조사를 실시하였다. 설문지 구성은 선행연구 및 전문가 조사를 통하여 선정된 각 평가요소별 평가지표를 근거로 하였으며, 총 4개 평가요소 및 52개 평가지표로 구성되었다. 평가상황요소에는 10개의 평가지표로, 평가투입요소에는 15개의 평가지표로, 평가수행요소에는 15개의 평가지표로, 평가효과요소에는 12개의 평가지표로 구성되었다. 설문지는 각 문항 7점 척도의 리커트 방식으로 작성되었으며, 설문은 국방연구개발사업 자체평가에 참여하였던 전문가를 대상으로 총 53부의 설문지를 회수하여 결과분석에 적용하였다.

3.2.1 평가상황요소 요인분석

평가상황요소에 해당하는 10가지 평가지표에 대하여 요인분석을 실시해 본 결과, [표 4]에서와 같이 3가지 요인으로 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 요인

이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 81.881%로 나타났다.

[표 4] 평가상황요소의 요인분석 결과

| 요인명 | 평가지표 | Factor1 | Factor2 | Factor3 |
|-----------------|--------------------|---------|---------|---------|
| 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 | .209 | .915 | -.001 |
| | 평가의 법적·제도적 근거 유무 | .113 | .782 | .359 |
| | 평가모형의 적절성 구비 여부 | .510 | .537 | .406 |
| 평가 대상의 합리성 | 평가대상사업 선정기준의 명확성 | .329 | .238 | .843 |
| | 국방 특수성 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 |

3.2.2 평가투입요소 요인분석 결과

평가투입요소에 해당하는 15가지 평가지표에 대하여 요인분석을 실시해 본 결과, [표 5]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 5] 평가투입요소의 요인분석 결과

| 요인명 | 평가지표 | Factor1 | Factor2 | Factor3 | Factor4 |
|------------|-----------------|---------|---------|---------|---------|
| 평가 인력의 적정성 | 평가위원의 전문성 및 경험 | .970 | -.164 | -.019 | .015 |
| | 사업규모 고려 평가위원 구성 | .882 | .181 | .001 | -.109 |
| | 평가위원 구성의 공정성 | .849 | .126 | -.164 | -.300 |
| | 평가위원의 양적 충부성 | .835 | .323 | .185 | .309 |
| | 평가위원 사전 교육실시 | .782 | .385 | -.089 | .039 |
| 평가 자료의 | 평가자료의 양적 충부성 | .234 | .940 | -.153 | -.039 |

| | | | | | |
|-----------------|-----------------|--------|--------|--------|--------|
| 총실성 | 평가자료의 질적 우수성 | .242 | .889 | .016 | -.081 |
| | 평가자료 활용이 용이성 | .075 | .885 | -.365 | -.270 |
| | 평가자료의 적기 제공 | .028 | .758 | -.011 | .500 |
| 평가 조직의 적합성 | 평가조직의 평가기능 부합 | -.197 | .093 | .947 | -.116 |
| | 평가조직과 평가결과비 관련성 | -.097 | -.278 | .799 | -.129 |
| | 평가조직의 제도적 독립성 | .411 | .054 | .754 | -.376 |
| 평가 예산의 합리성 | 평가예산의 안정적 지원 | .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 |

3.2.3 평가수행요소 요인분석 결과

평가수행요소에 해당하는 15가지 평가지표에 대하여 요인분석을 실시해 본 결과, [표 6]에서와 같이 4가지 요인으로 축약되었다. 요인 1에는 4개 평가지표를 묶어 ‘평가절차의 적정성’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 명명하였다. 각 요인의 고유치는 최고 3.847부터 2.778까지 나타났고, 4개의 요인이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 91.755%로 나타났다.

[표 6] 평가수행요소의 요인분석 결과

| 요인명 | 평가지표 | Factor1 | Factor2 | Factor3 | Factor4 |
|-----------------|-----------------|---------|---------|---------|---------|
| 평가 절차의 적정성 | 계획된 평가절차의 일관성 | .950 | .181 | -.215 | -.006 |
| | 평가절차의 적절한 수립 | .938 | .029 | -.103 | -.042 |
| | 피평가기관 의견 반영 | .904 | -.286 | -.047 | .113 |
| | 평가위원회 합리적 운영 | .873 | -.127 | -.165 | .159 |
| 평가 과정의 공정성 | 특수이해집단과의 연계성 | -.161 | .969 | .077 | .096 |
| | 기간중 내·외부의 압력 행사 | -.128 | .923 | .191 | .068 |
| | 이해관계자간 의사소통 | .101 | .812 | .371 | .300 |
| 평가 방법의 적합성 | 평가주관기관의 공정성 | .163 | .749 | .516 | .293 |
| | 평가등급 산정의 적절성 | -.128 | .120 | .976 | -.056 |
| | 평가기간 및 시간의 합리성 | -.188 | .290 | .887 | -.014 |
| | 평가방법의 효율적 수행 | -.276 | .583 | .711 | -.235 |
| 평가 기준의 합리성 | 평가방법의 평가목적 부합 | -.441 | .307 | .626 | -.299 |
| | 평가대상사업 특성 반영 | .283 | -.036 | .193 | .920 |
| | 평가지표의 적절성 | -.010 | .180 | -.172 | .903 |
| 평가지표의 요소별 적절성 | | -.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 |

3.2.4 평가효과요소 요인분석 결과

평가결과요소에 해당하는 12가지 평가지표에 대하여 요인분석을 실시해 본 결과, [표 7]에서와 같이 3가지로 축약되었다. 요인 1에는 5개 평가지표를 묶어 ‘평가결과 의 신뢰성’으로 명명하였으며, 요인 2는 4개의 평가지표를 묶어 ‘평가결과 의 환류성’이라는 요인으로 축약되었다. 요인 3은 3가지 평가지표를 묶어서 ‘평가결과 의 활용성’이라는 요인으로 명명하였다. 각 요인의 고유치는 최고 3.889부터 2.373까지 나타났고, 3개의 요인이 전체평가지표를 설명하는 비율인 설명분산의 누적계수는 74.518%로 나타났다.

[표 7] 평가효과요소의 요인분석 결과

| 요인명 | 평가지표 | Factor1 | Factor2 | Factor3 |
|------------------|---------------------|---------|---------|---------|
| 평가 결과의 신뢰성 | 피평가기관의 소명기회 | .873 | .287 | -.043 |
| | 평가보고서 작성의 충실성 | .863 | .363 | .036 |
| | 평가결과 의 적절한 공개 | .860 | .120 | .229 |
| | 피평가기관의 평가결과 수용 | .826 | .175 | .375 |
| 평가 결과의 환류성 | 평가보고서의 적기 제공 여부 | .770 | .144 | .421 |
| | 성과부진 및 문제 사업의 개선 | .226 | .892 | .038 |
| | 보완 반영사항의 이행 여부 | .277 | .708 | .412 |
| 평가 결과의 활용성 | 평가정보시스템 구축 및 접근 가능성 | .295 | .600 | -.039 |
| | 평가결과 의 예산배정 연계성 | .083 | .538 | .183 |
| | 평가결과 의 차년도 계획 반영 정도 | .191 | -.047 | .889 |
| | 평가시스템 개선시 평가결과 활용 | .330 | .387 | .771 |
| 평가결과 활용을 연계장치 구비 | | .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 |

3.3 가중치 결정

요인분석 결과를 토대로 선정된 평가요소 및 평가항목에 대하여 가중치 결정을 위한 AHP 설문조사를 실시하였다. 설문조사는 국방연구개발사업 자체 평가 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 8]과 같다. 그리고 전체 일관성 지수는 0.00052로 나타났으며, 평가상황요소 간 일관성 지수는 0.00903, 평가투입요소 간 일관성 지수는 0.00046, 평가수행요소 간 일관성 지수는 0.00709, 평가효과요소 간 일관성 지수는 0.00021으로 나타나 분석대상인 22명에 대한 전반적인 일관성은

있는 것으로 분석되었다.

[표 8] 평가항목별 가중치 및 일관성 지수

| 평가요소 (가중치) | 평가항목 (가중치) | 최종 가중치 | 일관성 지수 |
|-----------------|------------------|-----------|---------|
| 평가상황 (0.218) | 평가목적의 타당성(0.380) | 0.083 | 0.00903 |
| | 평가제도의 명확성(0.279) | 0.061 | |
| | 평가대상의 합리성(0.341) | 0.074 | |
| 평가투입 (0.288) | 평가인력의 적정성(0.302) | 0.087 | 0.00046 |
| | 평가자료의 충실성(0.189) | 0.054 | |
| | 평가조직의 적합성(0.324) | 0.093 | |
| 평가수행 (0.262) | 평가예산의 합리성(0.186) | 0.054 | 0.00052 |
| | 평가절차의 적정성(0.249) | 0.065 | |
| | 평가과정의 공정성(0.230) | 0.060 | |
| 평가효과 (0.232) | 평가방법의 적합성(0.263) | 0.069 | 0.00709 |
| | 평가기준의 합리성(0.259) | 0.068 | |
| | 평가결과의 신뢰성(0.357) | 0.083 | |
| 평가효과 (0.232) | 평가결과의 환류성(0.324) | 0.075 | 0.00021 |
| | 평가결과의 활용성(0.319) | 0.074 | |
| | | | |

4. 국방연구개발사업 메타평가 분석

4.1 메타평가 방법

메타평가의 분석방법에 있어서 제1단계는 평가항목에 대한 평가를, 제2단계는 평가항목별 가중치를 반영한 평가요소별 평가를, 제3단계는 평가요소별 점수를 합산한 종합평가를 실시하였다.

평가지표는 ‘전혀 그렇지 않다(1점), 별로 그렇지 않다(2점), 보통이다(3점), 그렇다(4점), 매우 그렇다(5점)’ 등 리커트 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’, 93점미만 80점 이상이면 ‘우수’, 80점미만 60점 이상이면 ‘보통’, 60점미만 40점 이상이면 ‘미흡’ 그리고 40점미만인 경우에는 ‘불량’으로 평가하였다.

설문조사 대상은 방위사업청, 국방기술품질원, 국방과학연구소, 항공우주연구원, 개발업체, 평가위원 등 2007

년부터 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.

4.2 종합평가 결과

국방연구개발사업 자체평가에 대한 메타평가 결과를 종합해 보면, [표 9]와 같이 100점 만점에 74.3점을 획득하여 ‘보통’의 수준으로 평가되었다. 메타평가 4개의 평가요소별로는 평가상황요소가 가장 높은 80.2점으로 ‘우수’하게 평가되었다. 평가상황 다음으로 높은 점수를 받은 요소는 평가수행요소로 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 떨어지고 있다. 특히, 평가결과의 활용성과 환류성에서 낮은 평가를 받았는데, 이러한 결과는 평가결과 활용을 위한 연계장치 구비가 미흡한 실정이며, 평가정보시스템 구축을 통한 접근가능성이 미흡하다고 할 수 있다.

[표 9] 메타평가 평가요소별 총점

| 평가요소 | 점수 (평균) | 가중 치 | 가중치 반영점수 | 100점 환산점수 | 평가 |
|------|------------|---------|-------------|--------------|----|
| 평가상황 | 4.007 | 0.218 | 0.874 | 80.2 | 우수 |
| 평가투입 | 3.639 | 0.288 | 1.053 | 73.1 | 보통 |
| 평가수행 | 3.746 | 0.262 | 0.980 | 74.8 | 보통 |
| 평가효과 | 3.450 | 0.232 | 0.801 | 69.0 | 보통 |
| 총점 | | 1 | | 74.3 | 보통 |

4.3 평가요소별 메타평가 결과

4.3.1 평가상황요소

국방연구개발사업 자체평가의 평가상황 요소에 대한 메타평가 결과는 [표 10]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 분류해야할 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 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. 평가상항요소에서는 평가모형의 적절성 구비 측면이 가장 낮은 점수를 얻었는데, 이는 국방연구개발사업 자체평가의 평가모형이 아직은 제대로 정착되지 못하였음을 의미한다 하겠다. 평가대상의 합리성에서는 평가대상사업 선정기준의 명확성 여부 4.02, 평가대상사업의 국방분야 특수성 고려 선정 여부 3.89, 평가대상 분류의 적절성 여부 3.82 등으로 평균 3.911로 100점 만점에 78.2점을 얻어 '보통'으로 평가되었다.

[표 10] 평가상항요소에 대한 메타평가 결과

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

4.3.2 평가투입요소

국방연구개발사업 자체평가의 평가투입요소에 대한 메타평가 결과는 [표 11]에서와 같이 100점 만점에 73.1점으로 '보통'으로 평가되었다. 평가항목별로는 평가인력의 적절성이 3.700, 평가자료의 충실성이 3.656, 평가조직의 적합성이 3.721, 평가예산의 합리성이 3.480로 평균 3.639에 해당하는 점수이다. 평가인력과 평가자료, 그리고 평가조직에 있어서는 양호한 평가를 받았음에도 불구하고

평가예산지원 및 적정 규모의 반영측면에서는 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. 평가자료의 충실성에서는 평가자료의 양적 충분성 여부 3.75, 평가자료의 질적 우수성 여부 3.66, 평가자료 활용이 용이한지 여부 3.59, 평가자료의 적기 제공 여부 3.62 등으로 평균 3.656으로 100점 만점에 73.1점을 얻어 '보통'으로 평가되었다. 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.

[표 11] 평가투입요소에 대한 메타평가 결과

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

4.3.3 평가수행요소

국방연구개발사업 자체평가의 평가수행요소에 대한 메타평가 결과는 [표 12]에서와 같이 100점 만점에 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. 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 유 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, 기간중 내·외부의 압력 행사 유무 4.14, 이해관계자간 의사소통 유무 3.68, 평가주관기관의 공정성 여부 4.10 등으로 평균 3.973로 100점 만점에 비교적 높은점수인 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 12] 평가수행요소에 대한 메타평가 결과

| 평가 항목 | | 평균 | 가중치 | 가중치 반영점수 | 100점 환산점수 | 평가 |
|------------|--------------|-------|-------|----------|-----------|----|
| 평가 절차의 적정성 | 평가절차의 일관성 | 3.96 | 0.065 | 0.239 | 73.6 | 보통 |
| | 평가절차의 적절한 수립 | 3.84 | | | | |
| | 피평가기관 의견 반영 | 3.50 | | | | |
| | 평가기간의 합리성 | 3.42 | | | | |
| | 평균 | 3.680 | | | | |
| 평가 과정의 공정성 | 이해집단과 연계성 | 3.97 | 0.060 | 0.238 | 79.5 | 보통 |
| | 기간중 내·외부의 압력 | 4.14 | | | | |
| | 이해관계자간 의사소통 | 3.68 | | | | |
| | 평가주관기관의 공정성 | 4.10 | | | | |
| | 평균 | 3.973 | | | | |
| 평가 방법의 | 평가팀 구성의 균형 | 3.93 | 0.069 | 0.258 | 74.8 | 보통 |
| | 평가분석 방법 사용 | 3.55 | | | | |

| | | | | | | |
|------------|--------------------|-------|-------|-------|------|----|
| 적합성 | 평가방법 효율적 수행 | 3.71 | 0.068 | 0.244 | 71.8 | 보통 |
| | 평가방법 평가목적 부합 | 3.78 | | | | |
| | 평균 | 3.740 | | | | |
| 평가 기준의 합리성 | 평가지표의 평가대상사업 특성 반영 | 3.75 | 0.262 | 0.980 | 74.8 | 보통 |
| | 평가지표의 적절성 | 3.52 | | | | |
| | 평가지표의 적절성 | 3.50 | | | | |
| | 평균 | 3.589 | | | | |
| 총계 | | 3.746 | | | | |

4.3.4 평가효과요소

국방연구개발사업 자체평가의 평가효과요소에 대한 메타평가 결과는 [표 13]에서와 같이 100점 만점에 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. 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. 평가결과의 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 시 평가결과 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.

[표 13] 평가효과요소에 대한 메타평가 결과

| 평가 항목 | | 평균 | 가중치 | 가중치 반영점수 | 100점 환산점수 | 평가 |
|------------|-------------|------|-------|----------|-----------|----|
| 평가 결과의 신뢰성 | 피평가기관의 소명기회 | 3.57 | 0.083 | 0.291 | 70.2 | 보통 |
| | 평가보고서 충실성 | 3.51 | | | | |
| | 평가결과의 공개 | 3.51 | | | | |

| | | | | | | |
|------------|---------------------|-------|-------|-------|------|----|
| | 평가보고서의 의사결정권자 보고 | 3.52 | | | | |
| | 평가보고서 적기 제공 | 3.46 | | | | |
| | 평균 | 3.510 | | | | |
| 평가 결과의 환류성 | 성과부진/문제사업 개선 | 3.49 | 0.075 | 0.256 | 68.2 | 보통 |
| | 보완·반영사항의 이행 | 3.49 | | | | |
| | 평가정보시스템 구축 및 접근 가능성 | 3.33 | | | | |
| | 평가결과 예산배정 연계 | 3.34 | | | | |
| | 평균 | 3.411 | | | | |
| 평가 결과의 활용성 | 평가결과의 차년도 연구계획에 반영 | 3.51 | 0.074 | 0.254 | 68.6 | 보통 |
| | 평가시스템 개선 시 | 3.45 | | | | |
| | 평가결과 활용을 위한 연계장치 구비 | 3.33 | | | | |
| | 평균 | 3.429 | | | | |
| | 총계 | 3.450 | | | | |
| | | | 0.232 | 0.801 | 69.0 | 보통 |

4.4 집단 간 차이분석 결과

4.4.1 평가집단과 피평가집단 간의 차이분석 결과

평가집단과 피평가집단 간에 따른 메타평가의 차이를 살펴보기 위해서 t-test를 실시하였다. 평가집단과 피평가집단 간의 차이를 분석한 결과 평가수행요소에서 평가절차의 적정성($P<.01$), 평가과정의 공정성($P<.01$)이, 평가효과요소에서는 평가결과의 신뢰성($P<.001$)에서 집단 간에 유의미한 차이가 나타났다. 세부적으로 살펴보면, 평가기관이 피평가기관보다 평가절차의 적정성, 평가과정의 공정성, 평가결과의 신뢰성을 높게 평가하는 것으로 나타났다. 평가집단과 피평가집단 간의 차이분석 결과는 [표 14]와 같다.

[표 14] 평가집단과 피평가집단 간의 차이분석 결과

| 항목 | | 집단 | 시료수 | 평균 | 표준편차 | t값 |
|------|-----------|-------|-----|------|------|----------|
| 평가수행 | 평가절차의 적정성 | 평가집단 | 57 | 3.88 | .60 | 2.932** |
| | | 피평가집단 | 148 | 3.60 | .61 | |
| | 평가과정의 공정성 | 평가집단 | 57 | 4.15 | .58 | 2.801** |
| | | 피평가집단 | 148 | 3.90 | .53 | |
| 평가효과 | 평가결과의 신뢰성 | 평가집단 | 57 | 3.69 | .38 | 4.101*** |
| | | 피평가집단 | 148 | 3.43 | .47 | |

* $P<.05$ ** $P<.01$ *** $P<.001$

4.4.2 연구개발사업 간의 차이분석 결과

연구개발사업에 따른 메타평가의 차이를 살펴보기 위해서 one-way ANOVA를 실시하였으며, 사후검증으로 Scheffe를 수행하였다. 연구개발사업 간의 차이분석을 살

펴본 결과 평가상황요소에서 평가목적의 타당성($P<.001$), 평가수행요소에서 평가절차의 적정성($P<.05$), 평가과정의 공정성($P<.05$), 평가방법의 적합성($P<.01$)에서 집단 간 유의미한 차이가 나타났다.

세부적으로 살펴보면 평가목적의 타당성은 국과연연구개발, 국책연구개발에서 집단 간에 차이를 보이고 있으며, 국과연연구개발사업에서 평가목적의 타당성을 높게 인식하는 것으로 나타났다. 평가수행에서 평가절차의 적정성은 국과연연구개발, 국책연구개발에서 집단 간에 차이를 보이고 있으며, 국과연연구개발사업에서 평가절차의 적정성이 높은 것으로 나타났다. 평가과정성의 공정성에서는 국과연연구개발, 국책연구개발과 업체연구개발에서 집단 간에 차이를 보이고 있으며 국과연연구개발사업에서 평가과정의 공정성이 높은 것으로 나타났다. 평가방법의 적합성에서는 국과연연구개발, 업체연구개발과 국책연구개발에서 집단 간에 차이를 보이고 있으며,

국과연연구개발사업에서 평가방법의 적합성이 높은 것으로 나타났다. 연구개발사업 간의 차이분석 결과는 [표 15]와 같다.

[표 15] 연구개발사업 간의 차이분석 결과

| 항목 | | 연구개발사업 | 시료수 | 평균 | 표준편차 | Scheffe | F값 |
|-----------|-----------|---------|------|------|------|---------|----------|
| 평가상황 | 평가목적의 타당성 | 국과연연구개발 | 23 | 4.36 | .53 | a | 9.028*** |
| | | 국책연구개발 | 82 | 3.89 | .55 | b | |
| | | 업체연구개발 | 100 | 4.15 | .52 | ab | |
| 평가수행 | 평가절차의 적정성 | 국과연연구개발 | 23 | 3.89 | .64 | a | 4.092* |
| | | 국책연구개발 | 82 | 3.53 | .59 | b | |
| | | 업체연구개발 | 100 | 3.74 | .62 | ab | |
| | 평가과정의 공정성 | 국과연연구개발 | 23 | 4.27 | .48 | a | 4.086* |
| | | 국책연구개발 | 82 | 3.96 | .47 | a | |
| | | 업체연구개발 | 100 | 3.90 | .62 | b | |
| 평가방법의 적합성 | 국과연연구개발 | 23 | 3.86 | .46 | a | 5.181** | |
| | 국책연구개발 | 82 | 3.57 | .61 | b | | |
| | 업체연구개발 | 100 | 3.84 | .58 | a | | |

* $P<.05$ ** $P<.01$ *** $P<.001$

4.4.3 자체평가 참여횟수간의 차이분석 결과

자체평가 참여횟수에 따른 메타평가의 차이를 살펴보기 위해서 one-way ANOVA를 실시하였으며, 사후검증으로 Scheffe를 수행하였다. 자체평가 참여횟수간의 차이분석 결과 평가효과요소에서 평가결과의 신뢰성($P<.05$), 평가결과의 환류성($P<.05$)에서 유의미한 차이가 나타났다. 세부적으로 살펴본 결과 평가결과의 신뢰성과 평가결과의 환류성 모두 2회와 3회에서 횟수별 차이를 나타내고 있으며, 참여횟수가 2회시일 때 평가결과의 신뢰성과 환류성이 높은 것으로 나타났다. 자체평가 참여횟수 간의

차이분석 결과는 [표 16]과 같다.

[표 16] 평가 참여횟수 간의 차이분석 결과

| 항목 | 참여횟수 | 시료수 | 평균 | 표준편차 | Scheffe | F값 | |
|-------|------------|-----|-----|------|---------|----|--------|
| 평가 효과 | 평가 결과의 신뢰성 | 1회 | 118 | 3.54 | .43 | ab | 3.644* |
| | | 2회 | 42 | 3.58 | .46 | a | |
| | | 3회 | 45 | 3.35 | .49 | b | |
| | 평가 결과의 환류성 | 1회 | 118 | 3.41 | .48 | ab | 4.125* |
| | | 2회 | 42 | 3.55 | .43 | a | |
| | | 3회 | 45 | 3.25 | .52 | b | |

*P<.05 **P<.01 ***P<.001

4.4.4 근속년수간의 차이분석 결과

근속년수에 따른 메타평가의 차이를 살펴보기 위해서 one-way ANOVA를 실시하였으며, 사후검증으로 Scheffe를 수행하였다. 근속년수간의 차이분석결과 평가투입요소에서 평가인력의 적정성(P<.01), 평가자료의 충실성(P<.05), 평가조직의 적합성(P<.001)이 평가수행요소에서는 평가절차의 적정성(P<.05), 평가과정의 공정성(P<.01), 평가방법의 적합성(P<.05)에서 유의미한 차이가 나타났으며, 세부결과는 [표17]과 같다.

[표 17] 근속년수 간의 차이분석 결과

| 항목 | 근무기간 | 시료수 | 평균 | 표준편차 | Scheffe | F값 | |
|-------|------------|---------|----|------|---------|----|----------|
| 평가 투입 | 평가 인력의 적정성 | 5년이하 | 50 | 3.69 | .58 | a | 5.198** |
| | | 5년~10년 | 28 | 3.31 | .65 | b | |
| | | 10년~20년 | 76 | 3.74 | .62 | a | |
| | | 20년이상 | 51 | 3.85 | .51 | a | |
| | 평가 자료의 충실성 | 5년이하 | 50 | 3.81 | .59 | a | 3.684* |
| | | 5년~10년 | 28 | 3.41 | .59 | b | |
| | | 10년~20년 | 76 | 3.67 | .47 | ab | |
| | | 20년이상 | 51 | 3.60 | .50 | ab | |
| | 평가 조직의 적합성 | 5년이하 | 50 | 3.73 | .54 | a | 5.429*** |
| | | 5년~10년 | 28 | 3.38 | .63 | b | |
| | | 10년~20년 | 76 | 3.83 | .48 | a | |
| | | 20년이상 | 51 | 3.72 | .41 | a | |
| 평가 수행 | 평가 절차의 적정성 | 5년이하 | 50 | 3.71 | .67 | ab | 3.012* |
| | | 5년~10년 | 28 | 3.36 | .70 | b | |
| | | 10년~20년 | 76 | 3.76 | .53 | a | |
| | | 20년이상 | 51 | 3.69 | .62 | ab | |
| | 평가 과정의 공정성 | 5년이하 | 50 | 3.90 | .61 | ab | 5.062** |
| | | 5년~10년 | 28 | 3.66 | .64 | b | |
| | | 10년~20년 | 76 | 3.98 | .48 | ab | |
| | | 20년이상 | 51 | 4.15 | .48 | a | |
| | 평가 방법의 적합성 | 5년이하 | 50 | 3.77 | .62 | | 2.702* |
| | | 5년~10년 | 28 | 3.50 | .65 | | |
| | | 10년~20년 | 76 | 3.85 | .51 | | |
| | | 20년이상 | 51 | 3.66 | .62 | | |

*P<.05 **P<.01 ***P<.001

세부적으로 살펴보면 평가투입의 평가인력의 적정성에서는 근무기간이 5년~10년인 경우와 5년 이하와 10년 이상인 집단에서 유의미한 차이가 나타났으며, 근무기간이 20년 이상인 경우에 평가인력의 적정성을 가장 높게 인식하는 것으로 나타났다. 평가자료의 충실성에서는 근무기간이 5년 이하와 5년~10년인 경우에 유의미한 차이가 나타났으며, 5년 이하에서 평가자료의 충실성을 높게 인식하는 것으로 나타났다. 평가조직의 적합성에서는 근무기간이 5년~10년인 경우와 5년 이하와 10년 이상인 집단에서 유의미한 차이가 나타났으며, 근무기간이 10년~20년인 경우에 평가조직의 적합성을 가장 높게 인식하는 것으로 나타났다. 평가수행요소에서는 평가절차의 적정성의 경우 근무기간이 5년~10년과 10년~20년에 따라서 집단 간의 유의미한 차이를 보이고 있으며, 근무기간이 10년~20년인 경우에 평가절차의 적정성을 가장 높게 인식하는 것으로 나타났다. 평가과정의 공정성의 경우 근무기간이 5년~10년과 20년 이상에 따라서 집단 간의 유의미한 차이를 보이고 있으며, 근무기간이 20년 이상인 경우에 평가과정의 공정성을 가장 높게 인식하는 것으로 나타났다. 평가방법의 적합성에서는 근무기간에 따른 세부적인 집단 간의 차이는 나타나지 않았으나 근무기간이 10년~20년인 경우에 평가방법의 적합성을 가장 높게 인식하고 있는 것으로 나타났다.

4.4.5 직업별 차이분석 결과

직업에 따른 메타평가의 차이를 살펴보기 위해서 one-way ANOVA를 실시하였으며, 사후검증으로 Scheffe를 수행하였다. 직업별 차이분석결과 평가상황요소에서 평가목적의 타당성(P<.001), 평가투입요소에서 평가인력의 적정성(P<.05), 평가조직의 적합성(P<.05), 평가예산의 합리성(P<.01), 평가수행의 요소에서 평가절차의 적정성(P<.05), 평가과정의 공정성(P<.001), 평가방법의 적합성(P<.01), 평가기준의 합리성(P<.01)에서 유의미한 차이가 나타났다.

세부적으로 살펴보면, 평가상황요소에서 평가목적의 타당성은 교수와 공무원, 군인, 연구원간의 집단 간 차이가 나타났으며, 교수가 평가목적의 타당성을 가장 높게 인식하는 것으로 나타났다. 평가투입요소에서 평가인력의 적정성은 교수와 공무원, 군인, 회사원간의 집단 간 차이가 나타났으며, 교수가 평가인력의 적정성을 가장 높게 인식하는 것으로 나타났다. 평가조직의 적합성은 교수와 공무원, 군인 간의 집단 간 차이가 나타났으며, 교수가 평가조직의 적합성을 가장 높게 인식하는 것으로 나타났다. 평가예산의 합리성은 교수와 공무원, 군인, 연구원간의 집단 간 차이가 나타났으며, 교수가 평가예산의 합리성을

가장 높게 인식하는 것으로 나타났다. 평가수행요소에서 평가절차의 적정성은 교수와 군인, 연구원간의 집단 간 차이가 나타났으며, 교수가 평가절차의 적정성을 가장 높게 인식하는 것으로 나타났다. 평가과정의 공정성은 교수와 공무원, 군인, 연구원, 회사원간의 집단 간 차이가 나타났으며, 교수가 평가과정의 공정성을 가장 높게 인식하는 것으로 나타났다. 평가방법의 적합성과 평가기준의 합리성은 직업별 세부적인 집단 간의 차이는 나타나지 않았으나, 평가방법의 적합성은 교수가 가장 높게 인식하는 것으로 나타났으며, 평가기준의 합리성은 회사원이 가장 높게 인식하는 것으로 나타났다. 직업별 차이분석 결과는 [표 18]과 같다.

[표 18] 직업별 차이분석 결과

| 항목 | | 직업 | 시료수 | 평균 | 표준 편차 | Scheffe | F값 |
|------------|------------|-----|------|------|-------|---------|----------|
| 평가 상황 | 평가 목적의 타당성 | 공무원 | 11 | 3.90 | .68 | b | 8.482*** |
| | | 군인 | 24 | 3.87 | .37 | b | |
| | | 연구원 | 88 | 3.92 | .55 | b | |
| | | 회사원 | 73 | 4.26 | .50 | ab | |
| | | 교수 | 9 | 4.66 | .37 | a | |
| 평가 투입 | 평가 인력의 적정성 | 공무원 | 11 | 3.47 | .68 | b | 2.785* |
| | | 군인 | 24 | 3.65 | .54 | b | |
| | | 연구원 | 88 | 3.70 | .54 | ab | |
| | | 회사원 | 73 | 3.67 | .67 | b | |
| | | 교수 | 9 | 4.31 | .59 | a | |
| | 평가 조직의 적합성 | 공무원 | 11 | 3.52 | .77 | b | 2.790* |
| | | 군인 | 24 | 3.64 | .51 | b | |
| | | 연구원 | 88 | 3.67 | .46 | ab | |
| | | 회사원 | 73 | 3.77 | .51 | ab | |
| | | 교수 | 9 | 4.19 | .70 | a | |
| | 평가 예산의 합리성 | 공무원 | 11 | 3.27 | .87 | b | 4.535** |
| | | 군인 | 24 | 3.27 | .65 | b | |
| | | 연구원 | 88 | 3.36 | .69 | b | |
| | | 회사원 | 73 | 3.62 | .73 | ab | |
| | | 교수 | 9 | 4.22 | .44 | a | |
| 평가 수행 | 평가 절차의 적정성 | 공무원 | 11 | 3.86 | .65 | ab | 3.152* |
| | | 군인 | 24 | 3.64 | .53 | b | |
| | | 연구원 | 88 | 3.57 | .62 | b | |
| | | 회사원 | 73 | 3.71 | .61 | ab | |
| | | 교수 | 9 | 4.27 | .56 | a | |
| | 평가 과정의 공정성 | 공무원 | 11 | 3.68 | .79 | b | 4.891*** |
| | | 군인 | 24 | 4.04 | .35 | b | |
| | | 연구원 | 88 | 3.97 | .47 | b | |
| | | 회사원 | 73 | 3.90 | .62 | b | |
| | | 교수 | 9 | 4.66 | .37 | a | |
| | 평가 방법의 적합성 | 공무원 | 11 | 3.88 | .75 | | 3.443** |
| | | 군인 | 24 | 3.69 | .42 | | |
| | | 연구원 | 88 | 3.59 | .59 | | |
| | | 회사원 | 73 | 3.84 | .59 | | |
| | | 교수 | 9 | 4.19 | .46 | | |
| 평가 기준의 합리성 | 공무원 | 11 | 3.42 | .53 | | 3.653** | |
| | 군인 | 24 | 3.51 | .59 | | | |
| | 연구원 | 88 | 3.45 | .60 | | | |
| | 회사원 | 73 | 3.80 | .63 | | | |
| | 교수 | 9 | 3.55 | .81 | | | |

*P<.05 **P<.01 ***P<.001

4.4.6 학위별 차이분석 결과

학위에 따른 메타평가의 차이를 살펴보기 위해서 one-way ANOVA를 실시하였으며, 사후검증으로 Scheffe를 수행하였다. 학위별 차이분석결과 평가상황요소에서

평가목적의 타당성(P<.05), 평가대상의 합리성(P<.05), 평가투입요소에서 평가조직의 적합성(P<.05), 평가수행요소에서 평가절차의 적정성(P<.001), 평가과정의 공정성(P<.01), 평가방법의 적합성(P<.05), 평가기준의 합리성(P<.01)에서 유의미한 차이가 나타났다. 학위별 차이분석 결과는 [표 19]와 같다.

세부적으로 살펴보면, 평가상황요소에서 평가목적의 타당성은 학사와 석사 간에 집단 간 차이를 나타내고 있으며, 학사가 평가목적의 타당성을 가장 높게 인식하는 것으로 나타났다. 평가대상의 합리성은 학사와 석사 간에 집단 간 차이를 나타내고 있으며, 학사가 평가대상의 합리성을 가장 높게 인식하는 것으로 나타났다. 평가투입요소에서 평가조직의 적합성은 박사가 가장 높게 인식하는 것으로 나타났다.

평가수행요소에서 평가절차의 적정성은 학사와 석사 간에 집단 간 차이를 나타내고 있으며, 학사가 평가절차의 적정성을 가장 높게 인식하는 것으로 나타났다. 평가과정의 공정성은 학사와 박사 간에 집단 간 차이를 나타내고 있으며, 박사가 평가과정의 공정성을 가장 높게 인식하는 것으로 나타났다. 평가방법의 적합성은 학사와 박사 간에 집단 간 차이를 나타내고 있으며, 학사가 평가방법의 적합성을 가장 높게 인식하는 것으로 나타났다. 평가기준의 합리성은 학사와 석사, 박사 간에 집단 간 차이를 나타내고 있으며, 학사가 평가기준의 합리성을 가장 높게 인식하는 것으로 나타났다.

[표 19] 학위별 차이분석 결과

| 항목 | | 최종학위 | 시료 수 | 평균 | 표준 편차 | Scheffe | F값 | |
|------------|------------|------|------|------|-------|---------|----------|---------|
| 평가 상황 | 평가 목적의 타당성 | 학사 | 51 | 4.20 | .51 | a | 4.06* | |
| | | 석사 | 93 | 3.95 | .57 | b | | |
| | 평가 대상의 합리성 | 박사 | 61 | 4.13 | .52 | ab | | |
| | | 학사 | 51 | 4.03 | .48 | a | | |
| 평가 투입 | 평가 조직의 적합성 | 석사 | 93 | 3.81 | .54 | b | 3.515* | |
| | | 박사 | 61 | 3.95 | .46 | ab | | |
| | 평가 과정의 공정성 | 학사 | 51 | 3.68 | .61 | | | 3.428* |
| | | 석사 | 93 | 3.64 | .45 | | | |
| 평가 수행 | 평가 절차의 적정성 | 박사 | 61 | 3.86 | .52 | | 6.893*** | |
| | | 학사 | 51 | 3.89 | .61 | a | | |
| | | 석사 | 93 | 3.51 | .61 | b | | |
| | | 박사 | 61 | 3.74 | .59 | ab | | |
| | 평가 방법의 적합성 | 학사 | 51 | 3.79 | .72 | b | 4.985** | |
| | | 석사 | 93 | 3.97 | .46 | ab | | |
| | | 박사 | 61 | 4.12 | .49 | a | | |
| | | 학사 | 51 | 3.91 | .74 | a | | |
| 평가 기준의 합리성 | 평가 기준의 합리성 | 석사 | 93 | 3.64 | .51 | ab | 3.354* | |
| | | 박사 | 61 | 3.74 | .55 | b | | |
| | 평가 기준의 합리성 | 학사 | 51 | 3.85 | .58 | a | | 6.355** |
| | | 석사 | 93 | 3.51 | .58 | b | | |
| 평가 기준의 합리성 | 박사 | 61 | 3.48 | .68 | b | | | |
| | 학사 | 51 | 3.48 | .68 | b | | | |

*P<.05 **P<.01 ***P<.001

5. 결론

본 연구는 국방연구개발사업 자체평가 제도발전 및 시스템 개선에 도움을 주고자 자체평가에 대한 메타평가 모형을 분석하고 메타평가를 실시하기 위한 평가항목을 개발하여 실증분석을 실시한 결과로써, 처음으로 메타평가 모형을 분석하고 평가지표를 개발하였다는 점에서 의의가 매우 크다고 볼 수 있다. 그리고 메타평가 연구를 위해 요인분석을 실시하여 평가항목 간 타당성 및 신뢰도를 입증하고, AHP기법을 적용하여 평가항목 간 가중치를 객관적으로 산정한 점은 기존 연구와의 차별성이라 할 수 있다.

개발된 메타평가 설문지를 이용하여 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. 평가투입요소에 대한 메타평가 결과는 73.1점으로 ‘보통’으로 평가되었다. 평가수행요소에 대한 메타평가 결과는 74.8점으로 ‘보통’으로 평가되었다. 평가효과 요소에 대한 메타평가 결과는 69.0점으로 비교적 낮은 점수의 ‘보통’으로 평가되었다.

그리고 집단 간 차이분석을 실시하였다. 평가집단과 피평가집단 간의 차이를 분석한 결과 평가수행요소에서 평가절차의 적정성($P<.01$), 평가과정의 공정성($P<.01$)이, 평가효과요소에서는 평가결과의 신뢰성($P<.001$)에서 집단 간에 유의미한 차이가 나타났다. 연구개발사업 간의 차이분석을 살펴본 결과 평가상황요소에서 평가목적의 타당성($P<.001$), 평가수행요소에서 평가절차의 적정성($P<.05$), 평가과정의 공정성($P<.05$), 평가방법의 적합성($P<.01$)에서 집단 간 유의미한 차이가 나타났다. 자체평가 참여횟수 간의 차이분석 결과 평가효과요소에서 평가결과의 신뢰성($P<.05$), 평가결과의 환류성($P<.05$)에서 유의미한 차이가 나타났다. 근속년수간의 차이분석 결과 평가투입요소에서 평가인력의 적정성($P<.01$), 평가자료의 충실성($P<.05$), 평가조직의 적합성($P<.001$)이 평가수행요소에서는 평가절차의 적정성($P<.05$), 평가과정의 공정성($P<.01$), 평가방법의 적합성($P<.05$)에서 유의미한 차이가 나타났다. 직업별 차이분석결과 평가상황요소에서 평가목적의 타당성($P<.001$), 평가투입요소에서 평가인력의 적정성($P<.05$), 평가조직의 적합성($P<.05$), 평가예산의 합리

성($P<.01$), 평가수행의 요소에서 평가절차의 적정성($P<.05$), 평가과정의 공정성($P<.001$), 평가방법의 적합성($P<.01$), 평가기준의 합리성($P<.01$)에서 유의미한 차이가 나타났다. 학위별 차이분석결과 평가상황 영역에서 평가목적의 타당성($P<.05$), 평가대상의 합리성($P<.05$), 평가투입 영역에서 평가조직의 적합성($P<.05$), 평가수행의 영역에서 평가절차의 적정성($P<.001$), 평가과정의 공정성($P<.01$), 평가방법의 적합성($P<.05$), 평가기준의 합리성($P<.01$)에서 유의미한 차이가 나타났다.

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 점에서는 매우 의의가 크다고 할 수 있다.

국방분야의 특성상 국방연구개발은 선진국 기술보호와 이전기피 심화로 필요한 무기체계를 개발하는 과정에서 예산의 효율적 배분과 사용은 아무리 강조해도 지나치지 않는다. 이처럼 주어진 예산 범위 내에서 국방연구개발사업을 성공적으로 수행할 수 있도록 하기 위해서는 국방연구개발사업에 대한 자체평가 업무의 효율성을 증대하여야 한다. 지금까지 국방분야의 연구개발사업에 대한 자체평가를 연구개발 기관에서 수행해 오면서 평가에 대한 신뢰성 및 공정성에 의문을 제기하게 되었다.

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 지속적 개선 및 연구개발사업 성과 극대화에 크게 기여하고, 국방예산의 비효율성을 방지하고 선진 무기체계를 성공적으로 개발하여 국방력을 강화하는데 큰 역할을 할 것으로 기대된다.

참고문헌

- [1] 김명수, “공공정책 평가론”, 박영사, 2003.
- [2] 노유진, “우리나라 정부부처에 있어서 정책평가결과의 저활용요인에 관한 연구”, 행정논총, 제43권, 제4호, pp. 245-276, 2005.
- [3] 노화준, “정책평가론”, 법문사, 2006.

- [4] 노화준, “정책평가론”, 법문사, 2008.
- [5] 류영수, “기술영향평가의 메타평가 모형개발 및 적용”, 한양대학교 대학원, 박사학위논문, 2007.
- [6] 문영세, “정부업무평가기본법에 대한 메타평가”, 정책분석학회보, 제11권, 1호, pp. 19-45, 2001.
- [7] 박중수, “정보화 사업 메타평가를 위한 CIPOU 모형 개발”, 충남대학교 대학원, 박사학위논문, 2003.
- [8] 방위사업청, “방위력개선사업관리규정”, 훈령 제13호, 2006.
- [9] 송위진, 이은경, 송성수, 김병운, “과학기술과 사회의 주요 쟁점분석 연구”, 과학기술정책연구원, 2003.
- [10] 이진주, 서건수, “정책평가를 위한 새로운 모형”, 남출판, 1996.
- [11] 이찬구, “영국의 범부처간 과학기술정책 평가제도- 메타평가와 한국에의 시사점”, 한국사화와 행정연구, 제15권, 제4호, pp. 809-833, 2003.
- [12] American Evaluation Association, "Guidelines for Evaluator", by Shadish W. R., D. L. Newman, M. A. Scheirer and C. Wye, San Francisco, Jossey-Bass, 1995.
- [13] Cook T. D. and Gruder C. L., “Metaevaluation Research”, Evaluation Quarterly, Vol. 2 No. 1, pp. 5-55, 1978.
- [14] Joint Committee on Standards for Educational Evaluation, "The Programme Evaluation Standards", Thousand Oaks, Sage Publications, 1994.
- [15] 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.
- [16] Larson, Richard and Leni Berliner, "On Evaluating Evaluation", Policy Sciences, Vol. 16, No. 2, 1983.
- [17] Leviton, Laura C & Edward F. X. Hughes, “Research on the Utilization of Evaluations: A Review and Synthesis”, Evaluation Review, Vol. 5, No. 4, 1981.
- [18] OECD, “Improving Evaluation Practices-Best Practice Guidelines for Evaluation and Background Paper”, Paris, 1999.
- [19] Owen, John M. and Patricia J. Rogers, "Program Evaluation: Forms and Approaches", International Edition, London and Thousand Oaks and New Delhi: Sage Publication, 1999.
- [20] Rossi, Peter H., Mark W. Lipsey and Howard E. Freeman, “Evaluation: A Systematic Approach”, Seventh Edition, Thousand Oaks and London and New Delhi: Sage Publication, 2004.
- [21] Scriven M., “Evaluation Thesaurus 4th Edition”, Newbury Park, CA, Sage Publications, 1991.
- [22] Shepard, L. A. “A Checklist for Evaluating Large-Scale Assessment Programs 9th Edition”, Occasional Paper Series, Kalamazoo: The Evaluation Center, western Michigan University, 1977.
- [23] Stevenson, John F., Richard H. Longabaugh & Dwight N. Mcneil, “Metaevaluation in Human Services”, The Evaluator and Management, Beverly Hills, CA, Sage Publications, 2000.
- [24] Stufflebeam L., “Metaevaluation: Concepts, Standard and Uses”, Educational Evaluation Methodology, The JHU Press, 1981.
- [25] Vedung, Evert, “Public Policy and Program Evaluation”, New Brinswick and London, Transactions Publishers, 1997.

김 순 영(Soon Yeong Kim)

[정회원]



- 1990년 2월 : 전북대학교 대학원 기계공학과 (기계공학석사)
- 2007년 3월 ~ 현재 : 호서대학교 벤처전문대학원(경영학박사)
- 1984년 3월 ~ 현재 : 국방기술 품질원 선임연구원, 기술거래사

<관심분야>
경영, 연구개발사업평가

하 규 수(Kyu Soo Ha)

[정회원]



- 1995년 2월 : 고려대학교 경영대학원 (경영학석사)
- 1998년 6월 : TOURO 법과전문대학원(J.D.), 미국 뉴욕주변호사
- 2009년 2월 : 한양대학교 경영학과 (경영학박사)
- 2002년 2월 ~ 현재 : 호서대학교 벤처전문대학원 교수

<관심분야>
벤처경영, 창업, 경영전략