

유치원, 어린이집 통합평가체계의 방향과 정책과제: 전문가 델파이 조사결과를 중심으로*

박창현¹⁾

요약

본 연구는 유아교육·보육 평가체계의 통합 방향과 핵심 정책과제를 전문가 델파이 조사를 통해 도출하는 데 목적이 있다. 평가관련 전문가 20명을 대상으로 2차에 걸친 델파이 조사를 실시하였으며, 수집된 자료는 평균, 표준편차, 내용타당도비율(CVR), 변동계수(CV)로 분석하였다. 주요 결과는 다음과 같다. 첫째, 법·제도적 이원 구조가 통합의 근본적 장애요인으로 확인되었으며, 제도 보완과 거버넌스 조정이 핵심 해결방안으로 합의되었다. 둘째, 통합평가의 설계 원칙으로 성찰·대화 기반 자체평가, 개선 중심 평가, 과정의 질, 교직원 성장이 도출되었다. 셋째, 핵심 요소에서는 과정의 질, 교사 전문성, 아동 권리가, 지표 영역에서는 놀이 중심 교육과정과 교직원 전문성이 최상위 합의를 보였다. 넷째, 실행 체계에서는 법·지침 정렬과 데이터 허브 구축이 중요하며, 단기 우선 과제로 거버넌스 합의, 공통지표 정비, 평가자 전문성, 로드맵 설정이 확인되었다. 본 연구는 통합평가체계가 제도적 적합성과 데이터 기반 질 관리를 강화하는 동시에 평가결과의 환류와 현장지원을 체계적으로 설계해야 함을 시사한다.

주제어: 유치원·어린이집 평가, 유보통합, 전문가 델파이 조사

I. 서론

한국의 영유아교육·보육(Early Childhood Education and Care, 이하 ECEC)은 아동 발달과 평생 학습의 기초를 형성하는 공공서비스로서, 그 질은 시설과 규정 준수가 아니라 아동의 일상적 경험, 교직원 전문성, 제도적 지원 구조가 함께 만들어내는 다차원적 결과로

* 본 논문은 2025년 육아정책연구소에서 수행된 「유아교육·보육 평가체계의 통합방향과 미래과제」(연구보고 2025-6)의 일부 내용을 수정·보완하여 재구성한 것임.

1) 육아정책연구소 연구위원

이해된다(박창현·김근진·정다솔, 2025; 박창현·정수지·정다솔, 2025). OECD의 Starting Strong 시리즈는 20여 년에 걸쳐 구조적 질과 과정적 질을 포괄하는 통합적 정책 설계를 강조해 왔다. 교사-아동 간 상호작용의 응답성, 정서적 지지, 놀이의 깊이와 같은 과정적 질은 아동 발달 성과를 직접 매개하는 핵심 변인이며, 구조적 질은 이를 뒷받침하는 필요 조건이되 충분조건은 아니라는 점이 반복적으로 확인되고 있다(OECD, 2015; 2021; 2022; Slot, 2018). 이러한 관점에서 질 관리 정책은 구조적·과정적 요소를 함께 다루어야 하며, 제도 기반과 데이터 시스템 위에서 작동하되 현장 개선과 정책 학습으로 환류되어야 한다(박창현·정수지·정다솔, 2025; OECD, 2022).

국제적 논의는 평가체계가 결과 판단의 기제가 아니라 공공 책임과 질 개선을 결합하는 정책 장치로 설계되어야 함을 강조한다. OECD는 질 모니터링이 서비스와 인력의 질을 높이고 아동 발달로 환류되어야 한다고 보며(OECD, 2015; 2019), UNICEF와 UNESCO는 권리 기반 접근에 따라 영유아 돌봄·교육을 공공 책무로 재구성할 것을 강조한다(UNICEF & UNESCO, 2024). 이러한 논의는 ECEC 평가가 기관 운영의 적합성 점검을 넘어 아동의 발달 조건 보장과 연결되어야 함을 시사한다(National Research Council & Institute of Medicine, 2000).

Slot(2018)은 OECD 워킹페이퍼에서 교사 대 아동 비율, 집단 규모, 교직원 자격 등 구조적 조건이 과정적 질을 뒷받침하되 그 관계가 맥락에 따라 달라진다고 보고, 평가체계가 구조와 과정을 결합한 설계로 접근해야 함을 제안하였다. 유럽연합은 접근성, 인력, 교육과정, 모니터링, 거버넌스를 하나의 질 프레임으로 제시하였으며(Council of the European Union, 2019), 다섯 영역이 결합된 시스템 개혁 과제로 작동할 때 그 효과가 확장된다는 점이 확인된다(European Commission, 2023; Hulpia, Sharmahd, Bergeron-Morin et al., 2024). 호주는 국가품질프레임워크(NQF) 아래 「국가법」, 「국가규정」, 국가품질표준, 평가·등급 체계, 기관별 질 향상계획(QIP)을 순환 구조로 결합하여 운영하고 있으며(ACECQA, 2018), 잉글랜드 Ofsted와 스코틀랜드 HGIOELC는 각각 교사-아동 상호작용 관찰과 자가평가를 제도화하였다(Education Scotland, 2016; Ofsted, 2019).

국내에서는 유치원과 어린이집이 서로 다른 법·제도 및 행정체계 아래 운영되어 왔으며, 이러한 분절은 평가체계에도 구조적으로 반영되었다. 유치원 평가는 「유아교육법」 제19조 및 같은 법 시행령 제20조부터 제22조에 근거하여 교육부(시·도교육청) 주관으로 실시된다(국가법령정보센터, 2025a). 2007년 시범평가를 거쳐 2008년부터 본격 시행된 유치원 평가는 현재 교육과정 및 방과후 과정, 교육환경 및 안전관리, 건강·급식·방역, 교직원, 운영관리 등의 영역으로 구성되며, 3년 주기 자체평가와 서면평가를 중심으로 운영된다(박창

현·정수지·정다솔, 2025).

어린이집 평가는 「영유아보육법」 제30조에 근거하여 시행되며, 2024년 6월 27일 정부 조직 개편을 기점으로 보건복지부에서 교육부로 주관 부처가 이관되었고, 한국보육진흥원에 위탁하여 운영된다. 2006년 평가인증제로 출발한 뒤 2019년 의무 평가제로 전환되었으며(국가법령정보센터, 2025b; 교육부·한국보육진흥원, 2024), 2024년 7월 개정 평가제는 영유아 중심 일과 운영, 영유아가 주도하는 배움 지원 등 6개 영역 15개 항목으로 구성되어 등급제 대신 영역별 서술형 결과서를 제공한다.

두 체계는 법적 근거, 평가 방법, 지표 구성 논리, 결과 활용 방식 등에서 구조적 차이를 보인다(박창현·정수지·정다솔, 2025). 유치원 평가는 교육과정과 교원 전문성에, 어린이집 평가는 보육서비스 적정성과 운영 관리에 초점을 두어, 동일 연령대 영유아의 질을 다루면서도 관점과 지표 구조에서 공통점과 차이점이 동시에 나타난다.

국내 연구는 이러한 분절이 평가체계 통합을 어렵게 하는 구조적 조건임을 반복적으로 지적해 왔다(정미영·김남희, 2025). 임명희(2014)는 평가 목적과 기준, 운영 방식이 분절된 상황에서는 기관 간 비교 가능성과 질 개선 환류가 제약된다고 보고, 통합 설계가 지표의 단순 결합이 아니라 제도적 틀과 운영 체계의 정합성 확보를 포함해야 함을 강조하였다. 최윤경·박창현·하연섭 외(2016)와 정미영·김남희(2025)도 평가를 포함한 질 관리 체계의 일원화가 행정적 통합을 넘어 실질적 질 보장으로 이어지기 위한 핵심 과제임을 제시하였다. 염지숙·고영미(2014)는 유아교육·보육 서비스 질 향상을 위한 평가제도 통합 방안을 분석하면서, 공통 평가 기준 마련과 함께 기관 유형별 특성을 반영하는 유연한 설계가 필요하다고 제안하였다.

현장 수준에서도 평가가 질 개선의 학습 장치로 작동하기보다 행정 부담을 가중시키는 방식으로 경험될 수 있다는 우려가 축적되어 왔다. 문서 중심의 증빙과 규정 준수 점검이 강화될수록 기관은 서류 대응과 형식 충족에 자원을 투입하게 되고, 상호작용이나 놀이와 같은 과정적 질은 정책 목표로 선언되더라도 평가 실천에서는 주변화될 위험이 있다(박창현·정수지·정다솔, 2025). 국제적으로도 측정이 목적을 대체하고 관리가 배움을 압도하는 현상에 대한 경계가 이어져 왔으며(Power, 1999), 유아교육 분야에서도 지표와 인증의 과잉이 교사 자율성과 상호작용의 질을 잠식할 수 있다는 문제의식이 제기되어 왔다(Dahlberg, Moss, & Pence, 2013). 이러한 흐름 속에서 평가의 목적이 통제가 아닌 학습과 개선에 있어야 한다는 논의가 국제적으로 확산되고 있다(European Commission, 2023; OECD, 2015; 2022).

최근 국내 유보통합 논의가 확장되면서 평가체계 통합은 교육·보육 정책의 일관된 목표

달성과 질 개선을 위한 핵심 기반으로 다시 부상하고 있다(정미영·김남희, 2025). 교사 자격·양성과정, 근로 조건, 운영 지원 체계 등 인력·전문성 구조가 통합 정책의 성패를 좌우한다는 점이 강조되며(최윤경·박창현·하연섭 외, 2016), 이는 평가체계가 교직원 전문성 개발과 지원 구조를 구조적으로 포함해야 함을 시사한다. 그러나 유보통합이 관리체계 일원화와 지원 확대에 머물 경우, 통합의 실질적 성과가 영유아가 경험하는 교육·보육의 질로 충분히 이어지지 못할 위험이 있다. 행정 효율 중심의 접근이 질의 하향평준화와 서비스 동질화로 귀결될 수 있다는 우려가 제기되어 왔으며(박창현, 2025), OECD 역시 ECEC 체계의 질은 단일 지표나 단일 제도 변경만으로 확보될 수 없고, 법·재정·인력·교육과정·모니터링이 일관된 방향으로 정합성을 갖추도록 조정될 때 비로소 체제 수준의 질이 형성된다는 점을 강조해 왔다(OECD, 2015; 2022; 2025b).

이와 같은 국내·국제 맥락은 유치원·어린이집 평가체계 통합이 단일 주관체계 선택이나 평가 주기 조정과 같은 단편적 설계로 해결될 수 없음을 보여준다. 평가체계 통합은 제도적 정합성 확보, 과정 중심의 질 측정, 데이터 기반 질 관리, 현장 지원과 환류 구조의 결합이라는 통합적 접근을 요구한다. 이러한 정책 과제의 복잡성과 불확실성을 고려할 때, 다양한 분야의 전문가 판단을 체계적으로 수렴하는 델파이 기법은 합의 기반의 설계 원칙과 우선과제를 도출하는 데 적합한 방법이다(Hasson, Keeney, & McKenna, 2000). 이에 본 연구는 한국 ECEC 평가체계의 질적 통합 방향과 실행 과제를 전문가 합의를 통해 체계적으로 탐색하고, 이를 바탕으로 정책적·실천적 설계 지침을 제시하는 데 목적이 있다. 이에 따른 연구문제는 다음과 같다.

연구문제 1. 우리나라 ECEC 평가체계 통합의 장애요인과 해결 전략은 무엇인가?

연구문제 2. 통합평가체계의 설계 원칙과 핵심 요소는 무엇인가?

연구문제 3. 통합평가체계의 실행 체계와 단기 우선 과제는 무엇인가?

II. 연구방법

1. 조사설계

본 연구는 유아교육·보육 평가체계의 통합 방향과 과제를 체계적으로 도출하기 위해 수정 델파이 기법에 기반한 구조화된 전문가 합의 연구로 설계되었다. 수정 델파이 기법은 고전적 델파이와 달리 1차 조사에서부터 문헌 분석과 정책 검토를 기반으로 구조화된 설문

을 제시하여 전문가 의견을 효율적으로 수집하고, 통제된 피드백을 통해 합의 수렴을 촉진하는 방법이다(Hasson, Keeney, & McKenna, 2000). 본 연구에서는 국내외 선행연구, 정책 문헌, 한국 ECEC 평가체계의 구조적 특성 분석을 토대로 설문 문항을 사전 구성한 뒤, 전문가 패널을 대상으로 2차에 걸친 조사를 실시하였다.

2. 연구대상

패널 선정은 의도적 표집방법을 적용하였으며, 다음 세 가지 기준을 설정하였다. 첫째, 유아교육과 보육·아동복지 두 분야의 전문성이 균형 있게 대표될 것, 둘째, 학계의 이론적 관점과 정책·행정 현장의 실행 경험이 함께 포함될 것, 셋째, 평가체계 설계·운영·개선에 직접 관여한 경험을 보유할 것이었다. 이 기준에 따라 총 20명의 전문가를 선정하였다. 직능별로는 학계(대학 교수 및 연구기관 연구자) 13명(65.0%), 행정·정책 실무(교육부·보건복지부 관련 부서, 시·도교육청, 한국보육진흥원 등 평가 운영·관리 실무) 7명(35.0%)으로 학술적 관점과 정책·행정 관점이 균형 있게 반영되었다. 전공 분야는 교육·유아교육 10명(50.0%), 보육 8명(40.0%), 특수·장애영유아 보육 2명(10.0%)으로 나타났다. 유아교육 전공 전문가 중 상당수는 보육과정, 돌봄정책, 영유아 권리 등 보육·복지 영역을 겸비하고 있어, 실질적으로 양 분야의 관점이 고르게 반영되었다. 경력은 6-10년 3명(15.0%), 11-15년 3명(15.0%), 16-20년 5명(25.0%), 21-25년 3명(15.0%), 26년 이상 6명(30.0%)이며, 패널의 85.0%가 11년 이상의 경력을 보유하고 있다. 성별은 남성 2명(10.0%), 여성 18명(90.0%)이다.

〈표 1〉 패널 인구통계학적 배경

			단위 : 명, %
구분	범주	빈도	백분율
직능	학계	13	65.0
	행정	7	35.0
전문분야	교육·유아교육	10	50.0
	보육	8	40.0
	특수·장애영유아 보육	2	10.0
총 경력	6~10년	3	15.0
	11~15년	3	15.0
	16~20년	5	25.0
	21~25년	3	15.0
	26년 이상	6	30.0
성별	남성	2	10.0
	여성	18	90.0
	합계	20	100.0

3. 연구도구

설문 문항 개발은 다음의 절차에 따라 수행되었다. 첫째, 문항 개발의 기초 단계로서 국내외 선행연구 및 정책 문헌을 분석하여 핵심 쟁점을 도출하였다. OECD(2015; 2022)의 ECEC 질 보장 프레임워크, EU Council(2019)의 질 권고안, 호주 국가품질프레임워크(ACECQA, 2018), 그리고 국내 선행연구(박창현·정수지·정다솔, 2025; 염지숙·고영미, 2014; 이윤아·곽승주·강유선, 2018; 임명희, 2014) 등을 종합적으로 검토하여, 통합 평가 체계의 장애요인, 설계 원칙, 핵심 요소, 지표 구성, 환류 구조, 거버넌스, 우선과제 등 핵심 영역을 도출하였다. 둘째, 도출된 핵심 영역을 바탕으로 예비 문항을 구성하였다. 예비 문항은 리커트 5점 척도(1=전혀 중요하지 않음, 5=매우 중요함)를 기반으로 전문가의 판단을 정량적으로 수집하되, 필요시 서술형 의견을 병행 수집할 수 있도록 구성하였다. 셋째, 예비 문항에 대해 ECEC 정책·평가 분야 전문가 3인의 사전 검토를 실시하였다. 사전 검토에서는 내용타당도, 문항 표현의 적절성, 영역 간 중복 여부, 응답 용이성 등을 점검하였으며, 검토 결과를 반영하여 문항의 표현을 수정하고 중복 문항을 통합하였다. 최종 설문지는 8개 영역, 총 73개 문항으로 확정되었다. 8개 영역은 ①통합의 장애요인(4문항), ②통합의 장애요인 해결방안(4문항), ③통합평가의 원칙(21문항), ④미래형 평가체계 핵심 요소(11문항), ⑤통합평가 지표 핵심 영역(8문항), ⑥차별화·연령구분·환류 구조(3문항), ⑦실행 체계(거버넌스·운영)(9문항), ⑧단기 우선 과제(13문항)으로 구성되었다.

〈표 2〉 설문 구성 영역 및 문항 수

단위: 문항 수(개)

영역	설명	문항 수
통합의 장애요인	유치원·어린이집 평가체제 통합의 어려움 요인 탐색	4
통합의 장애요인 해결방안	장애요인 극복을 위한 실질적 전략	4
통합평가의 원칙	평가체계 설계 시 준수해야 할 기본 가치 및 원칙	21
미래형 평가체계 핵심 요소	향후 평가체계가 반드시 포함해야 할 핵심 요소	11
통합평가 지표 핵심 영역	통합평가 지표의 필수 영역 및 구성 요소	8
차별화·연령구분·환류구조	연령별, 환류체계 관련 평가 구조적 요건	3
실행 체계(거버넌스·운영)	통합평가 운영·법제·관리·데이터 체계	9
단기 우선 과제	통합 평가 실행을 위한 우선 과제	13
총계		73

4. 조사절차

본 연구의 델파이 조사는 2차에 걸쳐 진행되었다. 1차 조사는 2025년 9월 14일부터 20일까지, 2차 조사는 9월 26일까지 구조화된 설문지를 배포·회수하는 방식으로 실시하였다. 1차 조사에서는 전문가 패널 20명에게 8개 영역 73개 문항에 대한 리커트 5점 척도 기반의 중요도 평가를 요청하였으며, 동시에 각 영역마다 서술형 의견란을 배치하여 정량 응답의 맥락과 판단 근거를 정성적으로 수집하였다. 1차 조사가 완료된 후 문항별 평균, 표준편차, 내용타당도비율(Content Validity Ratio, CVR), 변동계수(Coefficient of Variation, CV)를 산출하여 합의 수준과 수렴 정도를 분석하였고, 서술형 응답은 내용분석을 통해 범주화하여 2차 설문 문항의 표현 수정, 범위 조정, 정합성 보완에 반영하였다.

2차 조사에서는 1차 분석 결과를 바탕으로 통제된 피드백을 제공하였다. 각 문항에 1차 전체 응답자의 평균과 표준편차를 사전 제시하여, 응답자가 자신의 판단을 집단 경향과 비교·확인하면서 재평가할 수 있도록 하였다. 이는 델파이 기법의 핵심 원리로, 전문가가 집단 의견의 분포를 인지한 상태에서 자신의 판단을 재확인하거나 수정할 기회를 제공하며, 극단적 의견을 조정하고 합의 수렴을 촉진하는 기능을 한다(Hasson, Keeney, & McKenna, 2000). 2차 조사 결과 모든 문항의 CV 값이 0.50 미만으로 나타나 전문가 의견이 충분히 수렴된 것으로 확인되었으며(English & Kernan, 1976), 추가 조사 없이 합의를 확정하였다.

5. 자료처리와 분석방법

정량 자료는 SPSS 28.0과 Microsoft 365를 활용하여 문항별 평균과 표준편차를 산출하였다. 전문가 의견의 합의 수준은 다음의 두 지표로 평가하였다. 첫째, 내용타당도비율(CVR)은 Lawshe(1975)의 공식 $CVR = (ne - N/2) / (N/2)$ 에 따라 산출하였다. 여기서 ne 는 해당 문항에 대해 4점 또는 5점으로 응답한 전문가 수이며, N 은 전체 패널 수이다. 패널 20명을 기준으로 $CVR \geq 0.42$ 를 내용타당도 확보의 임계값으로 적용하였다(Lawshe, 1975). 둘째, 변동계수(CV)는 표준편차를 평균으로 나눈 상대적 분산 지표로, $CV < 0.50$ 일 경우 의견이 충분히 수렴된 것으로 해석하였다(English & Kernan, 1976). 최종 결과에서는 CVR 충족 여부와 CV 수렴도를 함께 고려하여 합의 여부를 판정하였다.

III. 연구결과

1. 통합의 장애요인과 해결전략

가. 1차 분석 결과

1) 통합의 장애요인

통합의 장애요인 가운데 법·제도 요인이 평균 4.80($SD=0.52$), CVR 0.90으로 가장 높은 중요도와 합의를 보였다. 운영 요인과 재정·인력 요인은 각각 평균 4.15에 CVR 0.50과 0.60으로 기준을 충족하였다. 현장 요인은 평균 4.20으로 비교적 높았으나 CVR 0.40으로 기준(0.42)에 미달하여 2차에서 재확인 대상으로 분류되었다.

〈표 3〉 통합의 장애요인 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 법·제도 요인(교육부·복지부 이원 구조, 상위법 미비 등)	4.80	0.52	0.90	0.11
2) 운영 요인 (지표·절차 불일치, 행정부담 등)	4.15	1.04	0.50	0.25
3) 현장 요인 (기관 수용성 부족, 특성화 기관 반발 등)	4.20	0.89	0.40	0.21
4) 재정·인력 요인 (안정적 지원 부족 등)	4.15	1.14	0.60	0.27

2) 통합의 장애요인 해결방안

통합의 장애요인 해결방안에서는 제도 보완(평균 4.85, CVR 0.90), 단계적 이행(평균 4.50, CVR 0.80), 현장 지원(평균 4.50, CVR 0.70), 기관 다양성 존중(평균 4.35, CVR 0.60) 순으로 4개 항목 모두 CVR 기준을 충족하였다.

〈표 4〉 통합의 장애요인 해결방안 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 제도 보완 (법·지침 정비, 거버넌스 조정 등)	4.85	0.49	0.90	0.10
2) 단계적 이행 (시범사업→점진적 확대 등)	4.50	0.95	0.80	0.21
3) 현장 지원 (재정·인력 확충, 연수·컨설팅 제공 등)	4.50	0.89	0.70	0.20
4) 기관 다양성 존중 (공통지표+선택지표 구조 설계 등)	4.35	1.04	0.60	0.24

나. 2차 분석 결과

1) 통합의 장애요인

통합의 장애요인에서 법·제도 요인은 평균 4.80에 CVR 1.00으로 전원 합의에 도달하여 1차보다 합의가 강화되었다. 재정·인력 요인(평균 4.30, CVR 0.80), 현장 요인(평균 4.20, CVR 0.60), 운영 요인(평균 4.10, CVR 0.70) 순으로 4개 항목 모두 CVR 기준을 충족하였다. 특히 1차에서 기준에 미달하였던 현장 요인이 2차에서 CVR 0.40에서 0.60으로 상승하여 합의 기준을 충족하였다는 점이 주목된다.

〈표 5〉 통합의 장애요인 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 법·제도 요인 (교육부-복지부 이원 구조, 상위법 미비 등)	4.80	0.40	1.00	0.08
2) 운영 요인 (지표·절차 불일치, 행정부담 등)	4.10	0.89	0.70	0.22
3) 현장 요인 (기관 수용성 부족, 특성화 기관 반발 등)	4.20	0.75	0.60	0.18
4) 재정·인력 요인 (안정적 지원 부족 등)	4.30	0.64	0.80	0.15

2) 통합의 장애요인 해결방안

통합의 장애요인 해결방안은 제도 보완(평균 4.80, CVR 1.00), 현장 지원(평균 4.45, CVR 0.90), 단계적 이행(평균 4.35, CVR 0.80), 기관 다양성 존중(평균 4.15, CVR 0.80) 순으로 합의가 확보되었다. 1차와 비교하면 제도 보완은 CVR 0.90에서 1.00으로 합의가 강화되었고, 현장 지원은 평균 4.50에서 4.45로 유사한 수준을 유지하면서 CVR이 0.70에서 0.90으로 상승하여 합의가 분명해졌다.

〈표 6〉 통합의 장애요인 해결방안 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 제도 보완 (법·지침 정비, 거버넌스 조정 등)	4.80	0.40	1.00	0.08
2) 단계적 이행 (시범사업 → 점진적 확대 등)	4.35	0.91	0.80	0.21
3) 현장 지원 (재정·인력 확충, 연수·컨설팅 제공 등)	4.45	0.74	0.90	0.17
4) 기관 다양성 존중 (공통지표+선택지표 구조 설계 등)	4.15	0.73	0.80	0.18

2. 통합평가의 설계 원칙과 핵심 요소

가. 1차 분석 결과

1) 통합평가의 원칙

“자체평가는 단순 점검이 아니라 성찰·대화 기반으로 운영한다”는 원칙이 전원 만점(평균 5.00, $SD=0.00$)으로 가장 높은 합의를 보였다. “과정의 질 우선”과 “교직원 성장 촉진”이 각각 평균 4.95, “개선 중심·서열화 지양”과 “아동 개인정보 보호·윤리”가 각각 평균 4.90으로 매우 강한 합의를 보였다. 반면 “위험 기반·비례성 원칙 적용”은 평균 3.90, CVR 0.20으로 합의가 형성되지 않았다. 위험 기반·비례성 원칙을 제외한 나머지 20개 문항은 모두 CVR 기준(0.42)을 충족하여, 1차 단계에서부터 통합평가 원칙에 대한 광범위한 합의가 확인되었다.

〈표 7〉 통합평가의 원칙 1차 분석 결과

문항	평균	표준 편차	CVR	CV
철학·가치 기반 원칙	4.74	0.54	0.93	0.11
1) 평가는 아동권·참여·관계성·돌봄을 핵심 가치로 삼는다.	4.65	0.67	0.80	0.14
2) 평가는 상호작용·놀이·정서적 안녕 등 과정의 질을 우선한다.	4.95	0.22	1.00	0.05
3) 평가는 형평성·포용성(장애·이주·저소득 등)을 별도 영역으로 반영한다.	4.45	0.94	0.90	0.21
4) 평가는 개선 중심이어야 하며, 서열화·낙인을 지양한다.	4.90	0.31	1.00	0.06
참여·전문성 원칙	4.85	0.11	1.00	0.07
5) 교육·보육 공동체는 평가의 설계·실행·환류 전 과정에 참여한다.	4.75	0.44	1.00	0.09
6) 평가는 교직원의 성장과 전문성 향상을 촉진한다	4.95	0.22	1.00	0.05
7) 평가는 자율성과 책무성의 균형을 유지하며, 기관의 규모·지역·철학을 존중한다.	4.85	0.37	1.00	0.08
연령·다양성 반영 원칙	4.48	0.83	0.80	0.19
8) 평가는 0-2세와 3-5세의 발달 차이를 지표와 방법에 반영한다.	4.55	0.83	0.80	0.18
9) 기관 특성(숲·생태, 발도르프, 레지오, 다문화·다언어 등)은 선택 지표로 반영한다.	4.25	1.16	0.60	0.27
10) 평가는 증빙을 최소화하고 행정부담을 줄이는 것을 원칙으로 한다.	4.65	0.49	1.00	0.11
윤리·데이터·공개 원칙	4.88	0.04	1.00	0.07
11) 평가는 데이터 처리 전 과정에서 아동 개인정보 보호와 윤리를 준수한다.	4.90	0.31	1.00	0.06
12) 평가결과공개는 요약·사례중심으로, 서열화·낙인을 최소화한다.	4.85	0.37	1.00	0.08
보완·추가 원칙	4.66	0.31	0.84	0.13

문항	평균	표준 편차	CVR	CV
13) 평가는 교육·보육의 연속성을 반영한다. (예: 0-2세 → 3-5세 → 초등 전이의 일관된 품질 보장)	4.75	0.72	0.90	0.15
14) 평가는 모든 아동의 접근성을 보장한다. (예: 장애, 다문화·다언어, 저소득 등 배경과 무관하게 동등한 참여)	4.80	0.52	0.90	0.11
15) 평가는 보편적 학습설계(UDL)와 합리적 조정을 반영한다.(예: 발달 수준·언어에 맞는 자료 제공, 장애 아동 환경 조정 등)	4.75	0.55	0.90	0.12
16) 평가는 위험 기반·비례성 원칙을 적용한다. (예: 사고·문제 이력이 많은 기관은 더 자주·엄격하게, 안정적 이고 규모가 작은 기관은 덜 자주·간단하게 평가)	3.90	1.17	0.20	0.30
17) 평가는 지속가능한 교육 요소(환경·기후·생태)를 포함한다. (예: 친환경 자원 사용, 기후위기 교육, 지속가능한 기관 운영)	4.50	0.76	0.90	0.17
18) 평가는 안전·보건·위기 대응 요소를 포함한다. (예: 아동 안전 관리, 위생·건강 관리, 화재·감염병 등 비상 상황 대처 능력)	4.75	0.44	1.00	0.09
19) 평가는 평가자의 전문성과 공정성을 제도적으로 관리한다.	4.85	0.49	0.90	0.10
20) 자체평가는 단순 점검이 아니라 성찰·대화 기반으로 운영한다.	5.00	0.00	1.00	0.00
21) 평가는 특성화 기관(발도르프·숲·레지오·다문화 등)의 자율성을 보장한다.	4.65	0.59	0.90	0.13

2) 미래형 평가체계 핵심 요소

과정의 질(상호작용, 놀이, 정서적 안녕)이 평균 4.95, CVR 1.00으로 가장 높은 합의를 보였다. 교사 전문성 성장 연계(평균 4.90, CVR 1.00), 아동 권리·참여 보장(평균 4.80, CVR 0.90), 안전·보건·위기 대응(평균 4.70, CVR 0.90) 순으로 합의가 확인되었다. 디지털 기반은 평균 4.15, CVR 0.70으로 기준을 충족하였으나 다른 요소에 비해 상대적 중요도는 낮게 평가되었다.

〈표 8〉 미래형 평가체계 핵심 요소 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 아동 권리·참여 보장	4.80	0.52	0.90	0.11
2) 과정의 질(상호작용, 놀이, 정서적 안녕)	4.95	0.22	1.00	0.05
3) 기관 맞춤형 평가(자율성, 맥락 존중)	4.60	0.60	0.90	0.13
4) 교사 전문성 성장 연계(연수·컨설팅 지원)	4.90	0.31	1.00	0.06
5) 가족·지역사회 협력	4.55	0.60	0.90	0.13
6) 공공성·형평성(장애, 다문화, 저소득 등)	4.65	0.49	1.00	0.11
7) 디지털 기반(데이터 허브, 자동화 기록)	4.15	0.81	0.70	0.20

문항	평균	표준 편차	CVR	CV
8) 지속가능성·기후·생태 요소	4.40	0.75	0.90	0.17
9) 안전·보건·위기 대응	4.70	0.57	0.90	0.12
10) 교육·보육 연속성과 초등 전이	4.40	0.75	0.70	0.17
11) 특성화기관 존중(숲·발도르프·레지오·가정어린이집 등)	4.35	0.49	1.00	0.11

3) 통합평가 지표 핵심 영역

놀이 중심 교육과정·방과후 과정 운영과 교직원 역량 및 전문성이 각각 평균 4.90으로 가장 높은 합의를 보였다. 여기서 놀이 중심 교육과정은 기관 수준의 교육과정 편성·운영 체계를 의미하며, 영유아 주도적 배움·상호작용 지원(평균 4.85, CVR 0.90)은 교사-아동 간 일상적 상호작용과 아동 주도성이라는 과정적 질 차원을 의미하는 것으로 구분된다. 이어 건강·안전·위기 대응(평균 4.70, CVR 0.80), 철학·공동체성·자율성(평균 4.45, CVR 0.80) 순으로 합의가 확인되었다.

〈표 9〉 통합평가 지표 핵심 영역 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 놀이중심 교육과정, 방과후과정 운영	4.90	0.45	0.90	0.09
2) 영유아 주도적 배움·상호작용 지원	4.85	0.49	0.90	0.10
3) 교육환경·운영관리	4.65	0.59	0.90	0.13
4) 건강·안전·위기 대응	4.70	0.66	0.80	0.14
5) 교직원 역량 및 전문성	4.90	0.31	1.00	0.06
6) 원장의 리더십·민주적 의사결정	4.50	0.69	0.80	0.15
7) 가정·지역사회 협력	4.65	0.75	0.90	0.16
8) 철학·공동체성·자율성 반영	4.45	0.69	0.80	0.15

4) 차별화·연령구분·환류 구조

“평가 결과가 개선계획 수립 → 맞춤형 컨설팅·연수 → 재평가·환류로 이어져야 한다”는 환류 구조 항목이 평균 4.50, CVR 0.90으로 가장 높은 합의를 보였다. “공통지표 외 선택 지표 반영”은 평균 4.10, CVR 0.50으로 기준을 충족하였으나 표준편차가 1.17로 의견 분산이 컸다. “0-2세와 3-5세 연령별 지표 구분”은 평균 4.00에 머물고 CVR 0.40으로 합의 기준에 미달하였다.

〈표 10〉 차별화·연령구분·환류 구조 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) [차별화] 공통지표 외에 별도의 선택지표(특색지표)를 반영하는 것이 필요하다.	4.10	1.17	0.50	0.28
2) [연령 구분] 0-2세와 3-5세 아동의 발달 차이를 고려하여, 연령별로 구분된 지표군을 별도로 두는 것이 필요하다.	4.00	1.38	0.40	0.34
3) [환류 구조] 통합평가 결과는 개선계획 수립 → 맞춤형 건설링·연수 지원 → 재평가·환류로 이어져야 한다.	4.50	0.76	0.90	0.17

나. 2차 분석결과

1) 통합평가의 원칙

“자체평가는 성찰·대화 기반 운영”이 평균 5.00, CVR 1.00으로 전원 최고점을 유지하였다. “개선 중심·서열화 지양”(평균 4.95, CVR 1.00), “교직원 성장 촉진”(평균 4.95, CVR 1.00), “과정의 질 우선”(평균 4.90, CVR 1.00), “평가자 전문성·공정성 제도화”(평균 4.90, CVR 1.00)가 강한 합의를 형성하였다. 반면 “위험 기반·비례성 원칙 적용”은 평균 3.60, CVR 0.10으로 합의 기준에 미치지 못하였다. 이는 전문가 집단이 평가의 교육적·발전적 기능—개선, 성장, 과정의 질, 평가자 전문성—에는 매우 강한 합의를 형성한 반면, 기관별 위험 수준에 따라 평가 빈도와 강도를 차등 적용하는 규제적 접근에 대해서는 현장 수용성과 형평성 우려로 인해 유보적 판단을 내린 것으로 해석된다.

〈표 11〉 통합평가의 원칙 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 평가는 아동권·참여·관계성·돌봄을 핵심 가치로 삼는다.	4.65	0.48	1.00	0.10
2) 평가는 상호작용·놀이·정서적 안녕 등 과정의 질을 우선한다.	4.90	0.30	1.00	0.06
3) 평가는 형평성·포용성(장애·이주·저소득 등)을 별도 영역으로 반영한다.	4.35	0.73	0.90	0.17
4) 평가는 개선 중심이어야 하며, 서열화·낙인을 지양한다.	4.95	0.22	1.00	0.04
5) 교육·보육 공동체는 평가의 설계·실행·환류 전 과정에 참여한다.	4.65	0.48	1.00	0.10
6) 평가는 교직원의 성장과 전문성 향상을 촉진한다	4.95	0.22	1.00	0.04
7) 평가는 자율성과 책무성의 균형을 유지하며, 기관의 규모·지역·철학을 존중한다.	4.70	0.46	1.00	0.10
8) 평가는 0-2세와 3-5세의 발달차이를 지표와 방법에 반영한다.	4.45	0.74	0.90	0.17
9) 기관 특성(숲·생태, 발도르프, 레지오, 다문화·다언어 등)은 선택지표로 반영한다.	4.00	0.95	0.60	0.24

16 육아정책연구(제20권 제1호)

문항	평균	표준 편차	CVR	CV
10) 평가는 증빙을 최소화하고 행정부담을 줄이는 것을 원칙으로 한다.	4.65	0.48	1.00	0.10
11) 평가는 데이터 처리 전 과정에서 아동 개인정보 보호와 윤리를 준수한다.	4.85	0.36	1.00	0.07
12) 평가결과공개는 요약·사례중심으로, 서열화·낙인을 최소화한다.	4.80	0.40	1.00	0.08
13) 평가는 교육·보육의 연속성을 반영한다.(예: 0-2세 → 3-5세 → 초등 전이의 일관된 품질 보장)	4.70	0.56	0.90	0.12
14) 평가는 모든 아동의 접근성을 보장한다.(예: 장애, 다문화·다언어, 저소득 등 배경과 무관하게 동등한 참여)	4.80	0.40	1.00	0.08
15) 평가는 보편적 학습설계(UDL)와 합리적 조정을 반영한다.(예: 발달 수준·언어에 맞는 자료 제공, 장애 아동 환경 조정 등)	4.65	0.48	1.00	0.10
16) 평가는 위험 기반·비례성 원칙을 적용한다. (예: 사고·문제 이력이 많은 기관은 더 자주·엄격하게, 안정적이고 규모가 작은 기관은 덜 자주·간단하게 평가)	3.60	0.73	0.10	0.20
17) 평가는 지속가능한 교육 요소(환경·기후·생태)를 포함한다. (예: 친환경 자원 사용, 기후위기 교육, 지속가능한 기관 운영)	4.40	0.58	0.90	0.13
18) 평가는 안전·보건·위기 대응 요소를 포함한다.(예: 아동 안전 관리, 위생·건강 관리, 화재·감염병 등 비상 상황 대처 능력)	4.60	0.49	1.00	0.11
19) 평가는 평가자의 전문성과 공정성을 제도적으로 관리한다.	4.90	0.30	1.00	0.06
20) 자체평가는 단순 점검이 아니라 성찰·대화 기반으로 운영한다.	5.00	0.00	1.00	0.00
21) 평가는 특성화 기관(발도르프·숲·레지오·다문화 등)의 자율성을 보장한다.	4.45	0.59	0.90	0.13

2) 미래형 평가체계 핵심 요소

과정의 질이 평균 4.90, CVR 1.00으로 가장 높았으며, 교사 전문성 성장 연계(평균 4.80, CVR 1.00), 아동 권리·참여(평균 4.75, CVR 1.00), 안전·보건·위기 대응(평균 4.65, CVR 0.90), 공공성·형평성(평균 4.60, CVR 1.00)이 뒤를 이었다. 디지털 기반(평균 4.10, CVR 0.80)은 기준을 충족하였으나 다른 요소에 비해 상대적 중요도는 낮게 평가되었다.

〈표 12〉 미래형 평가체계 핵심 요소 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 아동 권리·참여 보장	4.75	0.43	1.00	0.09
2) 과정의 질(상호작용, 놀이, 정서적 안녕)	4.90	0.30	1.00	0.06
3) 기관 맞춤형 평가(자율성, 맥락 존중)	4.45	0.50	1.00	0.11
4) 교사 전문성 성장 연계(연수·컨설팅 지원)	4.80	0.40	1.00	0.08

문항	평균	표준 편차	CVR	CV
5) 가족·지역사회 협력	4.45	0.50	1.00	0.11
6) 공공성·형평성(장애, 다문화, 저소득 등)	4.60	0.49	1.00	0.11
7) 디지털 기반(데이터 허브, 자동화 기록)	4.10	0.70	0.80	0.17
8) 지속가능성·기후·생태 요소	4.40	0.58	0.90	0.13
9) 안전·보건·위기 대응	4.65	0.57	0.90	0.12
10) 교육·보육 연속성과 초등 전이	4.50	0.50	1.00	0.11
11) 특성화 기관 존중(숲·발도르프·레지오·가정어린이집 등)	4.15	0.48	0.90	0.11

3) 통합평가 지표 핵심 영역

8개 핵심 영역이 모두 CVR 기준을 충족하여 강한 합의가 확인되었다. 놀이 중심 교육 과정과 교직원 역량·전문성이 각각 평균 4.95, CVR 1.00으로 가장 높았으며, 영유아 주도적 배움·상호작용 지원(평균 4.90), 건강·안전·위기 대응(평균 4.75), 교육환경·운영관리(평균 4.65), 가정·지역사회 협력 및 원장 리더십(각 평균 4.50), 철학·공동체성·자율성(평균 4.25) 순으로 나타났다.

〈표 13〉 통합평가 지표 핵심 영역 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 놀이중심 교육과정, 방과후과정 운영	4.95	0.22	1.00	0.04
2) 영유아 주도적 배움·상호작용 지원	4.90	0.30	1.00	0.06
3) 교육환경·운영관리	4.65	0.48	1.00	0.10
4) 건강·안전·위기 대응	4.75	0.54	0.90	0.11
5) 교직원 역량 및 전문성	4.95	0.22	1.00	0.04
6) 원장의 리더십·민주적 의사결정	4.50	0.50	1.00	0.11
7) 가정·지역사회 협력	4.50	0.50	1.00	0.11
8) 철학·공동체성·자율성 반영	4.25	0.54	0.90	0.13

4) 차별화·연령구분·환류 구조

환류 구조(평가-개선-재평가 연결 강화)가 평균 4.55, CVR 1.00으로 가장 높은 합의를 보였다. 연령 구분(0-2세/3-5세)은 평균이 4.00으로 1차와 동일하였으나 CVR이 0.40에서 0.50으로 상승하고 표준편차가 1.38에서 0.95로 감소하여 합의 기준을 충족하게 되었다. 반면 차별화(공통지표 외 선택지표)는 평균 4.10에서 3.90으로 하락하고 CVR이 0.50에서 0.40으로 떨어져, 1차에서 충족하였던 합의 기준이 2차에서 미달로 전환되었다.

〈표 14〉 차별화, 연령구분, 환류 구조 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) [차별화] 공통지표 외에 별도의 선택지표(특색지표)를 반영하는 것이 필요하다.	3.90	0.83	0.40	0.21
2) [연령 구분] 0-2세와 3-5세 아동의 발달 차이를 고려하여, 연령별로 구분된 지표군을 별도로 두는 것이 필요하다.	4.00	0.95	0.50	0.24
3) [환류 영역] 통합평가 결과는 개선계획 수립 → 맞춤형 컨설팅·연수 지원 → 재평가·환류로 이어져야 한다.	4.55	0.50	1.00	0.11

3. 실행 체계와 단기 우선 과제

가. 1차 분석결과

1) 실행 체계(거버넌스·운영)

“상위법은 원칙을, 하위 고시·지침은 세부 운영 기준을 담는 방식”으로 법·제도를 정렬해야 한다는 항목이 평균 4.70, CVR 1.00으로 가장 높은 합의를 보였다. 에듀파인 연계 데이터 허브 구축(평균 4.60, CVR 1.00), 사고·학대 위험기관 컨설팅·점검 강화(평균 4.55, CVR 0.80) 순으로 합의가 확인되었다. 반면 교육청 단위 자율평가 이행(평균 3.95, CVR 0.30), 3년 1회 정기평가와 1.5년 중간 점검 병행(평균 3.95, CVR 0.30), 가정어린이집 간소화 기준(평균 3.75, CVR 0.40)은 합의 기준에 미달하여 추가 검토가 필요한 쟁점으로 확인되었다.

〈표 15〉 실행 체계(거버넌스·운영) 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) [주관 체계] 중장기적으로 통합평가는 교육청(교육감) 단위 자율평가 체계로 일원화	3.95	1.32	0.30	0.33
2) [법·지침 정렬] 통합평가에서는 상위법은 원칙을, 하위 고시·지침은 세부 운영 기준을 담는 방식으로 법·제도를 정렬	4.70	0.47	1.00	0.10
3) [평가 주기] 통합평가에서는 3년 1회 정기평가 + 1.5년 중간 점검, 위험 기반 차등 적용 방식 적용	3.95	0.94	0.30	0.24
4) [디지털 허브] 통합평가에서는 에듀파인과 연계된 질관리 데이터 허브를 구축해야 한다.	4.60	0.50	1.00	0.11
5) [인센티브 연계] 통합평가에서는 평가 결과를 재정 지원·교사 연수·가점 등 인센티브와 연계해야 한다.	3.90	1.07	0.50	0.27

문항	평균	표준 편차	CVR	CV
6) [위험군 컨설팅] 사고·학대 이력 등 위험군 기관은 컨설팅 중심 지원과 현장 점검 강화가 필요하다.	4.55	0.69	0.80	0.15
7) [공시 체계] 부모 공시는 단일 템플릿(핵심항목)으로 제공하고, 아동학대·중대사건은 반드시 공시해야 한다.	3.90	1.21	0.50	0.31
8) [유형별 특례] 가정어린이집·소규모·다양한 운영 철학을 가진 기관은 간소화·맞춤형 기준을 적용해야 한다.	3.75	1.07	0.40	0.29
9) [로드맵] 통합평가는 단기 병행 운영 → 중기 교육청 이관 → 장기 교육청 단일 자율평가 단계로 이행해야 한다.	3.95	1.10	0.50	0.28

2) 단기 우선 과제

거버넌스 합의(교육청·한국보육진흥원·유아교육진흥원 역할 조정)가 평균 4.80, CVR 1.00으로 가장 높았다. 로드맵 설정(평균 4.75, CVR 1.00), 공통지표 합의(평균 4.70, CVR 0.80) 순으로 합의 기준을 충족하였다. 반면 공개 원칙(부모 공시)은 평균 3.85, CVR 0.30으로, 지역 격차 완화는 평균 3.95, CVR 0.40으로 합의 기준에 미달하였다.

〈표 16〉 단기 우선 과제 1차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 공통지표 합의(유치원·어린이집 공통 영역 설정)	4.70	0.80	0.80	0.17
2) 0-2세 특화(영아 발달 차이에 맞는 지표 설계)	4.30	1.13	0.60	0.26
3) 형평성 패키지(장애·다문화·저소득 지원 반영)	4.20	0.89	0.60	0.21
4) 데이터 허브 구축(에듀파인 연계 질관리 시스템)	4.25	0.85	0.70	0.20
5) 평가자 전문성(자격, 윤리)	4.45	0.89	0.70	0.20
6) 행정부담 경감(증빙 최소화, 자동화)	4.35	0.88	0.70	0.20
7) 컨설팅 연계(평가·개선·재평가 환류 고리 강화)	4.45	0.60	0.90	0.14
8) 공개 원칙(부모 공시 템플릿, 낙인 최소화)	3.85	0.99	0.30	0.26
9) 거버넌스 합의(교육청-한국보육진흥원-유아교육진흥원)	4.80	0.41	1.00	0.09
10) 선택지표 반영(숲·발도르프·레지오·가정어린이집 등)	3.95	0.94	0.60	0.24
11) 지역 격차 완화(도시-농촌, 수도권-비수도권 차이 해소)	3.95	1.05	0.40	0.27
12) 인센티브 제도(재정, 연수, 가점 등 동기부여)	4.25	0.72	0.70	0.17
13) 로드맵 설정(단기-중기-장기 단계 이행 계획)	4.75	0.44	1.00	0.09

나. 2차 분석 결과

1) 실행 체계(거버넌스·운영)

상위법-하위 지침 정렬(평균 4.60, CVR 1.00), 사고·학대 위험기관 점검 강화(평균 4.45, CVR 0.80), 데이터 허브 구축(평균 4.35, CVR 1.00)이 합의 기준을 충족하였다. 반면 단계적 로드맵 이행안(평균 3.75), 교육청 단위 자율평가 일원화(평균 3.70), 정기평가와 중간 점검 병행(평균 3.60), 인센티브 연계(평균 3.55), 부모 공시(평균 3.50), 가정어린이집 간소화 기준(평균 3.50)은 합의가 형성되지 않았다. 특히 가정어린이집 간소화 기준은 평균 3.50, CVR 0.00으로 4·5점 응답자가 정확히 절반에 그쳐 전문가 의견이 거의 양분된 양상을 보였는데, 이는 정책 설계 시 별도의 합의 형성 절차가 요구되는 쟁점임을 시사한다.

〈표 17〉 실행 체계(거버넌스·운영) 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 장기적으로 통합평가는 교육청(교육감) 단위 자율평가 체계로 일원화	3.70	1.23	0.30	0.33
2) 통합평가에서는 상위법은 원칙을, 하위 고시·지침은 세부 운영 기준을 담는 방식으로 법·제도를 정렬	4.60	0.49	1.00	0.11
3) 통합평가에서는 3년 1회 정기평가 + 1.5년 중간 점검, 위험 기반 차등 적용 방식이 적절하다.	3.60	0.73	0.10	0.20
4) 통합평가에서는 에듀파인과 연계된 질관리 데이터 허브를 구축해야 한다.	4.35	0.48	1.00	0.11
5) 통합평가에서는 평가 결과를 재정 지원·교사 연수·가점 등 인센티브와 연계해야 한다.	3.55	0.97	0.10	0.27
6) 사고·학대 이력 등 위험군 기관은 컨설팅 중심 지원과 현장 점검 강화가 필요하다.	4.45	0.67	0.80	0.15
7) 부모 공시는 단일 템플릿(핵심항목)으로 제공하고, 아동 학대·중대사건은 반드시 공시해야 한다.	3.50	1.02	0.20	0.29
8) 가정어린이집·소규모·다양한 운영 철학을 가진 기관은 간소화·맞춤형 기준을 적용해야 한다.	3.50	0.81	0.00	0.23
9) 통합평가는 단기 병행 운영 → 중기 교육청 이관 → 장기 교육청 단일 자율평가 단계로 이행해야 한다.	3.75	0.99	0.40	0.26

2) 단기 우선 과제

거버넌스 합의가 평균 4.70, CVR 1.00으로 가장 높았으며, 공통지표 합의(평균 4.65, CVR 0.70), 평가자 전문성(평균 4.55, CVR 0.90), 로드맵 설정(평균 4.50, CVR 1.00)

순으로 합의가 확인되었다. 반면 지역 격차 완화(평균 3.75, CVR 0.30), 선택지표 반영(평균 3.70, CVR 0.40), 공개 원칙(부모 공시)(평균 3.70, CVR 0.30)은 1차에 이어 2차에서도 합의 기준에 도달하지 못하였다.

〈표 18〉 단기 우선 과제 2차 분석 결과

문항	평균	표준 편차	CVR	CV
1) 공통지표 합의(유치원·어린이집 공통 영역 설정)	4.65	0.73	0.70	0.16
2) 0-2세 특화(영아 발달 차이에 맞는 지표 설계)	4.20	0.93	0.50	0.22
3) 형평성 패키지(장애·다문화·저소득 지원 반영)	4.10	0.70	0.60	0.17
4) 데이터 허브 구축(에듀파인 연계 질관리 시스템)	4.15	0.48	0.90	0.11
5) 평가자 전문성(자격, 윤리)	4.55	0.59	0.90	0.13
6) 행정부담 경감(증빙 최소화, 자동화)	4.25	0.54	0.90	0.13
7) 컨설팅 연계(평가-개선-재평가 환류 고리 강화)	4.40	0.49	1.00	0.11
8) 공개 원칙(부모 공시 템플릿, 낙인 최소화)	3.70	0.56	0.30	0.15
9) 거버넌스 합의(교육청-한국보육진흥원-유아교육진흥원)	4.70	0.46	1.00	0.10
10) 선택지표 반영(숲·발도르프·레지오·가정어린이집 등)	3.70	0.84	0.40	0.23
11) 지역 격차 완화(도시-농촌, 수도권-비수도권 차이 해소)	3.75	0.77	0.30	0.20
12) 인센티브 제도(재정, 연수, 가점 등 동기부여)	4.00	0.63	0.60	0.16
13) 로드맵 설정(단기-중기-장기 단계 이행 계획)	4.50	0.50	1.00	0.11

IV. 논의 및 정책적 시사점

1. 통합의 장애요인과 해결 전략(연구문제 1)

법·제도 요인이 두 차례 조사 모두에서 가장 높은 합의를 보인 것은, 유아교육법과 영유아보육법으로 분리된 이원적 법 체계가 평가체계 통합의 근본적 제약으로 작용함을 확인한다. 이는 임명희(2014)가 평가 목적·기준·방식의 분절이 기관 간 비교 가능성과 환류를 제약한다고 분석한 것, 그리고 정미영·김남희(2025)가 유보통합 대비 평가체계 통합의 구조적 어려움을 지적한 것과 일치한다. OECD(2022) 역시 질 보장이 거버넌스, 기준, 자원, 인력, 데이터의 유기적 결합을 통해 달성되어야 하며, 질을 단순 규제 준수로 환원해서는 안 된다고 강조한 바 있다.

해결방안으로 제도 보완이 가장 높은 합의를 얻고, 2차 조사에서 현장 지원, 단계적 이행, 기관 다양성 존중 순으로 나타난 것은 통합이 법·제도 정비를 선행 조건으로 하되, 현

장의 수용성을 확보하면서 점진적으로 추진되어야 함을 의미한다. 이는 염지숙·고영미(2014)가 공통 평가 기준 마련과 함께 기관 특성을 반영하는 유연한 설계의 필요성을 강조한 것, 그리고 호주 NQF가 국가 표준 아래 기관별 질 향상계획(QIP)을 통해 자율적 개선을 유도하는 접근(ACECQA, 2018)과 맥을 같이한다.

2. 통합평가의 설계 원칙과 핵심 요소(연구문제 2)

통합평가의 원칙에서 성찰·대화 기반 자체평가, 개선 중심, 과정의 질 우선, 교직원 성장 촉진, 평가자 전문성·공정성 제도화가 최상위 합의를 보인 것은, 전문가 집단이 평가의 본질을 통제나 서열화가 아닌 교육적 성장과 질 개선에 두고 있음을 보여준다. 이는 정선아·안소영·박보영(2023)이 어린이집 평가제에서 구조 중심에서 과정 중심으로의 패러다임 전환이 필요하다고 제안한 것과 일치하며, Dahlberg, Moss, & Pence(2013)가 질 평가에서 측정 가능성을 넘어 의미와 맥락에 주목해야 한다고 강조한 논의와도 연결된다.

위험 기반·비례성 원칙이 유일하게 CVR 기준에 미달한 것은 주목할 만하다. 서술형 응답을 종합하면, 전문가들은 기관의 위험 수준에 따라 평가 강도를 차등 적용하는 접근이 현장에서 낙인이나 불공정으로 인식될 수 있고, 기준 설정의 어려움과 소규모·취약 기관에 대한 역차별 우려가 반영된 것으로 해석된다. 이는 부모 공시 항목 및가정어린이집 간소화 기준 항목(CVR 0.00, 의견 양분)의 낮은 합의와도 맥락을 공유하는데, 평가 결과의 공개와 차등 적용이 서열화·낙인으로 이어질 수 있다는 우려가 전문가 집단에 넓게 공유되고 있음을 시사한다. OECD(2022)도 평가 결과 공개가 공적 책임성과 질 개선 정보를 균형 있게 제공해야 하며, 과도한 단일 지표 공개는 질 향상 목표를 저해할 수 있다고 지적한 바 있다.

핵심 요소에서 과정의 질, 교사 전문성, 아동 권리가 최상위 합의를 보인 것은 통합평가가 아동의 실제 경험과 교사의 전문적 실천을 중심에 두어야 함을 의미한다. Eadie, Page, Levickis et al.(2024)이 ECEC의 질 지표가 아동의 배움과 발달을 직접 반영해야 한다고 강조한 것, OECD(2021)가 의미 있는 상호작용을 과정적 질의 핵심으로 조명한 것과 부합한다. 연령 구분(0-2세/3-5세) 필요성에 대한 합의가 2차에서 강화된 것은 OECD(2020)가 0-2세 시기의 질 관리에 별도의 세심한 접근이 필요함을 강조한 논의와도 일치한다.

3. 실행 체계와 우선 과제(연구문제 3)

실행 체계에서 법·지침 정렬과 데이터 허브 구축이 높은 합의를 보인 반면, 평가 주기·방법, 인센티브 연계, 부모 공시, 기관 간소화 기준 등은 합의에 이르지 못하였다. 이는 평

가의 제도적 기반에는 전문가 간 인식이 수렴되지만, 구체적 운영 방식에는 현장 조건과 정책 환경에 따라 판단이 갈릴 수 있음을 보여준다. 특히 자율평가 일원화(평균 3.70)와 정기평가·중간점검 병행(평균 3.60) 두 방식 모두 합의 기준에 도달하지 못한 것은, 교사의 전문성·자율성을 평가체계 안에서 어떻게 보장하면서 동시에 공적 책임성을 담보할 것인가라는 근본 과제가 향후 평가 방법 설계 단계에서 재논의되어야 함을 시사한다.

단기 우선 과제에서 거버넌스 합의, 공통지표 정비, 평가자 전문성, 로드맵 설정이 상위 합의를 보인 것은 통합 추진의 순서에 대한 전문가 판단을 반영한다. 즉 통합의 출발점은 관련 기관 간 역할·책임 조정(거버넌스)이며, 이를 기반으로 공통 평가 기준을 마련하고, 이를 실행할 평가 인력의 전문성을 확보하며, 전체 과정을 단계적 로드맵으로 관리하는 것이 합리적 실행 경로임을 의미한다. 이러한 순서는 Klinkhammer, Schäfer, Harring et al.(2017)이 유럽 국가들의 질 모니터링 체계를 분석하면서 거버넌스 정합성이 모니터링 효과의 전제 조건임을 확인한 결과와 부합하며, OECD(2022)가 질 보장 체계에서 제도적 명료성의 중요성을 강조한 논의와도 일치한다

4. 정책적 제언 및 연구의 한계점

이상의 논의에 기반하여 다음과 같은 정책적 제언을 제시한다. 첫째, 거버넌스 합의는 통합평가의 전제조건이다. 교육부·보건복지부·지방자치단체·교육청 간 역할과 책임을 명확히 정렬하는 것이 평가체계 통합의 출발점이며, 이를 위해 상위법에 통합평가의 원칙과 거버넌스 구조를 명시하고 하위 고시·지침으로 세부 운영 기준을 구체화하는 법·제도 설계가 선행되어야 한다. 둘째, 데이터 기반 질 관리 체계 구축은 평가 결과를 지속적 질 개선으로 연결하는 핵심 인프라이다. 평가 결과가 단순 공시를 넘어 정책 설계와 현장 개선의 근거로 활용되어야 하며, 이를 위해 에듀파인 등 기존 시스템과 연계된 질 관리 데이터 허브의 구축이 필요하다(OECD, 2022). 셋째, 현장 지원 중심의 환류 구조가 평가의 교육적 의미를 결정한다. 위험기관에 대한 컨설팅 중심 지원과 교직원 연수·전문성 개발을 평가체계의 환류 고리에 포함시킴으로써, 평가가 처벌이나 서열화가 아닌 성장과 개선의 장치로 작동하도록 설계되어야 한다(OECD, 2019; 2025a). 넷째, 평가 결과의 공시는 서열화·낙인 방지를 전제로 신중하게 설계되어야 한다. 본 연구에서 부모 공시 항목의 합의 수준이 낮게 나타난 것은 평가 결과의 단순 공개가 현장 부담과 역기능을 초래할 수 있다는 전문가 인식을 반영한다. 공시는 요약·해설 중심 정보 제공과 개선 맥락 설명을 포함하는 방식으로 설계될 필요가 있다.

종합하면, 본 연구는 평가체계 통합이 지표 정비의 기술적 과제가 아니라 영유아 교육·

돌봄에 대한 공적 책임과 질 보장에 관한 정책적·철학적 선택의 문제임을 보여준다. 전문가들이 과정의 질, 성찰 기반 자체평가, 교직원 성장, 평가자 전문성에 강한 합의를 보인 것은 평가의 본질이 통제와 서열이 아닌 아동 경험과 교사 실천을 향해야 함을 시사한다. 거버넌스 정합성, 데이터 기반 환류, 현장 지원 설계는 평가가 질 개선의 순환 체계로 작동하기 위한 유기적 조건이며, 이러한 설계가 실현될 때 평가는 기관 판정 기제에서 영유아의 배움과 성장을 지원하는 공공 장치로 전환될 수 있다.

유보통합 전환기의 평가체계 통합은 유치원과 어린이집이 쌓아온 경험과 전문성을 존중하면서 모든 영유아가 어떤 기관에서든 질 높은 교육·돌봄을 경험하도록 하는 사회적 약속의 제도화이다. 본 연구의 전문가 합의는 그 약속을 구체화하는 첫걸음이며, 평가체계의 궁극적 지향은 제도의 정합성을 넘어 영유아 일상 경험의 질 보장이다. 그러나 본 연구는 다음 한계를 지닌다. 20명 전문가 대상 조사로 현장 교사·운영자·학부모의 경험이 충분히 반영되지 않았으며, 의견의 맥락적 배경을 심층 탐색하기 어려웠다. 후속 연구에서는 다층적 이해관계자 참여 설계, 질적 인터뷰·사례 연구를 병행한 혼합연구, 양분된 항목(가정어린이집 간소화 등)에 대한 구체 설계 후 재검증, 정책 시나리오 분석과 국제 비교를 통한 확장 가능성 검증이 요구된다.

참고문헌

- 교육부·한국보육진흥원(2024). 2024 개정 어린이집 평가 매뉴얼. 한국보육진흥원.
- 국가법령정보센터(2025a). 유아교육법, 유아교육법 시행령, 유아교육법 시행규칙(인출일: 2025년 12월 1일).
- 국가법령정보센터(2025b). 영유아보육법, 영유아보육법 시행규칙(인출일: 2025년 12월 1일).
- 박창현(2025). 새 정부 국정과제와 유아교육·보육의 방향: 유보통합, 질적 전환과 새로운 상상. 2025년도 제2차 KICCE 정책토론회 자료집. 육아정책연구소.
- 박창현·김근진·정다솔(2025). 영유아기 최상의 교육·보육을 위한 미래 과제: 유보통합의 질적 전환과 재개념화를 중심으로. 육아정책연구소.
- 박창현·정수지·정다솔(2025). 유아교육·보육 평가체제의 통합방향과 미래과제. 육아정책연구소.
- 염지숙·고영미(2014). 유아교육·보육 서비스 질 향상을 위한 평가제도의 통합 방안. 유아교육연구, 34(3), 455-479.
- 이윤아·곽승주·강유선(2018). 유아교육기관 평가에 관한 연구동향 분석: 2007-2017년 국내

- 학위논문 및 학술지를 중심으로. 한국유아교육연구, 20(3), 109-132.
- 임명희(2014). 유아교육기관 평가체계의 통합화 방안 연구. 한국보육학회지, 14(2), 83-102.
- 정미영·김남희(2025). 유보통합에 대비한 유아교육기관 평가체제 통합에 대한 유아교사 인식 비교. 생태유아교육연구, 24(3), 71-100.
- 정선아·안소영·박보영(2023). 보육서비스의 질 제고를 위한 어린이집 평가제의 방향 탐색: 구조에서 과정중심으로의 패러다임의 전환. 영유아교육과정연구, 13(3), 27-51.
- 최윤경·박창현·하연섭·김희수(2016). 유보통합 정착을 위한 실행 방안 연구. 육아정책연구소. ACECQA. (2018). *Guide to the National Quality Framework*. Sydney: Australian Children's Education and Care Quality Authority.
- Council of the European Union. (2019). Council Recommendation of 22 May 2019 on High-Quality Early Childhood Education and Care Systems (2019/C 189/02). *Official Journal of the European Union, C 189*, 4-14.
- Dahlberg, G., Moss, P., & Pence, A. (2013). *Beyond quality in early childhood education and care: Languages of evaluation* (3rd ed.). Routledge.
- Eadie, P., Page, J., Levickis, P., Elek, C., Murray, L., Wang, L., & Lloyd-Johnsen, C. (2024). Domains of quality in early childhood education and care: A scoping review of the extent and consistency of the literature. *Educational Review, 76*(4), 1057-1086. <https://doi.org/10.1080/00131911.2022.2077704>
- Education Scotland. (2016). *How good is our early learning and childcare? (HGIOELC)*. Education Scotland.
- English, J. M., & Kernan, G. L. (1976). The prediction of air travel and aircraft technology to the year 2000 using the Delphi method. *Transportation Research, 10*(1), 1-8.
- European Commission. (2023). *Monitoring and evaluation of ECEC quality: Towards inclusive and improvement-oriented approaches*. Publications Office of the European Union.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015.
- Hulpia, H., Sharmahd, N., Bergeron-Morin, L., De Pré, L., Crêteur, S., & Dunajeva, J. (2024). *Quality in Early Childhood Education and Care (ECEC): Analytical Report*. Publications Office of the European Union.
- Klinkhammer, N., Schäfer, B., Harring, D., & Gwinner, A. (Eds.). (2017). *Monitoring quality in early childhood education and care: Approaches and experiences from selected countries*. Munich: German Youth Institute.

- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- National Research Council & Institute of Medicine. (2000). *From neurons to neighborhoods: The science of early childhood development* (Shonkoff, J. P., & Phillips, D. A., Eds.). National Academy Press.
- OECD. (2015). *Starting Strong IV: Monitoring quality in early childhood education and care*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264233515-en>
- OECD. (2019). *Providing quality early childhood education and care: Results from the Starting Strong Survey 2018*. Paris: OECD Publishing. <https://doi.org/10.1787/301005d1-en>
- OECD. (2020). *Quality early childhood education and care for children under age 3: Results from the Starting Strong Survey 2018*. Paris: OECD Publishing. <https://doi.org/10.1787/99f8bc95-en>
- OECD. (2021). *Starting Strong VI: Supporting meaningful interactions in early childhood education and care*. Paris: OECD Publishing. <https://doi.org/10.1787/f47a06ae-en>
- OECD. (2022). *Quality assurance and improvement in the early education and care sector* (OECD Education Policy Perspectives No. 55). Paris: OECD Publishing. <https://doi.org/10.1787/774688bf-en>
- OECD. (2025a). *Reducing inequalities by investing in early childhood education and care*. Paris: OECD Publishing. <https://doi.org/10.1787/b78f8b25-en>
- OECD. (2025b). *Education policy outlook 2025: Nurturing engaged and resilient lifelong learners in a world of digital transformation*. Paris: OECD Publishing. <https://doi.org/10.1787/c3f402ba-en>
- Ofsted. (2019). *Education inspection framework: Overview of research*. Ofsted.
- Power, M. (1999). *The audit society: Rituals of verification*. Oxford University Press.
- Slot, P. L. (2018). *Structural characteristics and process quality in early childhood education and care: A literature review* (OECD Education Working Papers No. 176). Paris: OECD Publishing. <https://doi.org/10.1787/edaf3793-en>
- UNICEF & UNESCO. (2024). *A world ready to learn: Global report on early childhood development*. UNICEF.

- 논문접수: 2월 9일 / 수정본접수: 4월 6일 / 게재 승인: 4월 14일
- 교신저자: 박창현, 육아정책연구소 연구위원, pch0407@kicce.re.kr

Abstract

Directions and Policy Priorities for an Integrated Evaluation System of Kindergartens and Childcare Centers: Findings from an Expert Delphi Survey

Changhyun, Park

This study identified directions and policy priorities for integrating early childhood education and care (ECEC) evaluation systems through a two-round modified Delphi survey with 20 experts, with data analyzed using means, standard deviations, Content Validity Ratios (CVR), and Coefficients of Variation (CV). The dual administrative and legal structure emerged as the primary barrier, while institutional improvement and governance coordination proved essential strategies. Reflective self-evaluation, improvement-oriented evaluation, process quality, and teacher professional growth were confirmed as core design principles. Priority tasks included governance agreements, common indicator standardization, evaluator professionalism, data hub establishment, and a phased roadmap. An integrated ECEC evaluation system should enhance institutional coherence and data-driven quality management while systematically designing feedback mechanisms and on-site support linked to evaluation results.

Keywords: Kindergarten and Childcare Evaluation, ECEC Integration, Expert Delphi Survey