

초등학교 교사의 형성적 피드백 실행 척도 개발 및 타당화

염 시 창(전남대학교 교수)*

강 대 중(전남대학교 강사)**

<요 약>

이 연구의 목적은 초등학교 교사의 형성적 피드백 실행 척도를 개발하고, 척도의 신뢰도와 타당도를 확인하는 데 있다. 이를 위해, 반구조화된 설문 조사와 면담 및 토의를 통해 예비문항(50 문항)을 개발하였고, 초등학교 교사 470명의 응답 자료를 수합하였으며, 최종적으로 409명의 자료에 대해 Rasch 평정척도 분석, 확인적 요인분석, 상관분석을 실시하였다.

이 연구에서는 형성적 피드백 실행의 6개 하위요인(계획, 과정, 질문, 향상, 격려, 자기규제)에 대해 Rasch 모형을 적용한 후 확인적 요인분석을 실시하였는데, 그 결과 1차 5요인 모형이 적합한 것으로 나타났다. 초등학교 교사의 형성적 피드백 실행 척도는 5개의 하위요인(총 22개 문항), 즉 비계의 계획요인(4개 문항), 비계의 과정요인(6개 문항), 모니터링의 향상요인(4개 문항), 모니터링의 격려요인(4개 문항), 그리고 모니터링의 자기규제요인(4개 문항)으로 구성되었다. 또한 하위요인별 신뢰도가 높았으며, 하위요인 간 상관도 유의하였다. 초등학교 교사의 형성적 피드백 실행 척도의 하위요인은 기존의 학습을 위한 평가척도(단축형) 및 평가인식 척도와 유의한 정적 상관을 보였다. 이 연구에서 개발한 초등학교 교사의 형성적 피드백 실행 척도가 초등학교 교사들의 형성적 피드백 실재를 평가하는 데 유용하게 활용될 수 있을 것으로 기대된다.

주제어 : 형성적 피드백, 초등학교 교사, 척도 개발

I. 서론

교수-학습의 진행 과정에서 평가는 크게 진단평가, 형성평가, 총합평가로 구분된다. 진단평가(diagnostic assessment)는 학습자의 인지적·정의적 출발점 행동을 평가하는데 초점을 둔 반면, 총합평가(summative assessment)는 학기 중간 혹은 학기말에 학습자의 도달점 행동을

* 제1저자: sichang@chonnam.ac.kr

** 교신저자: kdjung92@hanmail.net

판단하는 성취도 평가에 해당된다. 위의 평가 유형과는 달리 형성평가(formative assessment)는 수업 중에 학습자의 학습 향상도 및 교사의 교수개선을 위한 피드백 제공을 목적으로 실행하는 평가이다. 형성평가는 프로그램 평가와 관련하여 Scriven(1967)이 처음으로 소개하였고, Bloom, Hastings와 Madus(1971), Bloom(1977)이 교수-학습 관련 용어로 사용하기 시작하면서 널리 활용된 평가 유형이다. 이러한 형성평가는 교육력 향상과 질 높은 교실평가와 관련된 논의가 활성화되면서 그 중요성이 크게 강조되어 왔다.

형성평가는 그 설계와 실천의 우선순위가 학생의 학습 증진이기 때문에 책무성이나 등급 혹은 자격증 교부를 목적으로 하는 평가와는 다르다. 또한, 이 평가 활동은 교수-학습 활동의 개선을 위해 교사와 학생들이 서로 피드백 정보를 제공하여 학습에 도움을 주는데 주안점을 둔다. 최근 들어 형성평가는 학습을 위한 평가(assessment for learning)로 외연을 확장함으로써 학습요구에 부합하도록 증거를 활용하여 교수활동을 적응시켜 나아가는 평가로 자리매김되고 있다(Black et al., 2004; Wiliam, 2011). 그러나 학습을 위한 평가 혹은 학습지향적 평가를 형성평가로 규정하고 학습결과의 평가를 총괄평가로 구분하는 것이 개념의 혼란만 가중시킨다는 지적도 있다(Bennett, 2011). 이 연구에서는 학습지향적 평가, 그리고 학습자가 자신의 학습에 주인의식을 갖고 자기 점검과 자기 수정을 이행하는 활동을 강조하는 개념인 학습으로서의 평가(assessment as learning)를 포함하여 형성평가로 정의한다. 즉, 형성평가는 학습을 극대화시키기 위해 교사가 수업에서 학생의 자료를 다각적으로 수집하고 피드백 하는 일련의 계획된 과정이다(김성숙 외, 2015). 또한, Shute(2008)는 형성적 피드백(formative feedback)을 학습향상을 위해 자신의 생각이나 행동을 수정하고자 노력하는 학습자에게 전달되는 정보로 정의하였는데, 이는 교수-학습 활동 개선을 위해 피드백을 제공하는 형성평가와 유사하지만 피드백의 핵심적 특성을 반영하고 있기 때문에, 이 연구에서는 형성평가라는 용어 대신에 형성적 피드백이라는 용어를 사용하고자 한다.

질적으로 양호한 교실평가를 위한 핵심 요소는 명확한 목적과 목표 설정, 건전한 설계, 효과적인 의사소통 및 학생참여이다(Stiggins, 2010). 그렇다면 교실평가에서 형성적 피드백의 하위요소와 핵심전략은 무엇인가? Wiliam(2011)은 학습목표의 이해, 학습증거 수집, 학습자 향상을 위한 피드백 제공, 동료의 학습 자원화, 학습자 자신의 주인의식 함양이라는 다섯 가지 핵심전략을 제시하였다. 또한, Florez & Sammons(2013)는 형성적 피드백의 전략으로 질문, 피드백, 동료평가와 자기평가, 총합평가 결과의 형성적 활용을 들었다. Nicol & Macfarlane-Dick(2006)은 학습자의 내적 처리 과정에 영향을 미치는 교사 및 동료 학생의 피드백에 주목하면서 학습자의 자기규제를 지원 및 개발하기 위한 7가지 피드백 전략을 제시하였다. 앞에서 언급한 전략 및 유형과는 별도로 Shute(2008)는 교사가 형성적 피드백을 실행할 때 반드시 해야 할 일, 피해야 할 일, 학습자의 특성을 고려한 피드백 실행 방법 등 구체적인 피드백 지침을 제안한 바 있다.

위와 같이 형성적 피드백 실행을 위한 전략이나 지침은 다양하게 제안되고 있지만, 정작 교사의 형성적 피드백 실행을 측정하기 위한 노력은 활발하게 전개되지 못했다. 국외에서 교사의 형성적 피드백 실행을 측정한 연구로 Yan & Cheng(2015)은 계획행동이론(theory of planned behavior)을 적용한 구조모형 검정을 위해 단지 두 문항의 형성평가 실행 문항을 사용하였다. 한편, 국내에서 지은림(2009)은 피드백 수행 변인을 평가목적, 평가태도, 평가유형별 피드백 수행, 피드백 수행방법으로 구분하여 17개 문항을 개발하였으며, 이후 피드백 유형, 피드백 효과를 포함한 22개 문항으로 확장한 평가 문항을 구성하였다(지은림, 2010). 그러나 이들 척도는 형성적 피드백의 영역을 형성적 피드백의 전략이나 지침에 근거하여 설정하기보다는 형성적 피드백의 정의에 기초하여 구분했고, 더 나아가 현장 교사들의 형성적 피드백 실행에 관한 의견을 충실하게 반영하지 못했다. 박민애, 손원숙(2016, 2018), 손원숙, 박정, 이빛나(2019)는 국내의 초·중·고등학생, 중등 예비교사 등을 연구대상으로 형성적 피드백에 관한 연구를 수행하면서 Pat-EI 외(2013)가 개발한 학습을 위한 평가척도(AFL-Q)를 변안하여 연구에 사용하였는데, 이는 국내 교실에서 실제로 교사들이 실행하고 있는 형성적 피드백을 반영하지 못한 한계를 안고 있다.

초등학교에서는 교사 한 명이 한 반을 하루 5~6시간 책임지고(이빛나, 2017), 여러 교과와 교수-학습 활동 과정에서 피드백이라고 할 수 있는 교사와 학생 간, 학생 상호 간의 상호작용의 기회가 중·고등학교에 비해 많이 형성되고 있다. 또한, 초등학교 교사보다 중·고등학교의 교사가 제공하는 피드백이 학생에게 미치는 영향이 줄어든다(노현중, 손원숙, 2015)는 특징도 고려하여 이 연구에서는 초등학교 교사를 중심으로 연구를 수행하고자 한다.

따라서 이 연구는 선행연구에서 논의된 형성평가의 정의와 형성적 피드백 실행 전략 및 지침, 그리고 학교현장 교사의 의견을 충실하게 반영한 초등학교 교사의 형성적 피드백 실행 척도를 개발하고 타당화 하는데 목적이 있다. 이 연구결과가 초등학교 교사의 형성적 피드백이 교수-학습 활동에 미치는 영향력을 검정하는 연구는 물론이고 초등학교 교사의 형성적 피드백에 영향을 미치는 학교 및 교사 변인을 밝히는 연구를 수행하는데 중요한 측정도구로 활용될 수 있을 것으로 기대한다.

II. 이론적 배경

1. 형성적 피드백의 실행 전략

그동안 형성적 피드백 혹은 학습지향적 평가와 관련하여 다양한 전략이 제안되어 왔는데, 이 연구에서는 형성적 피드백 실행 척도 구성과 연관성이 높은 전략을 중심으로 논의하고자

한다. 먼저 Wiliam(2011)은 형성적 피드백의 다섯 가지 핵심전략을 제시하였다. 첫째, 교사가 학습목표를 명시하고 이해시키며 이를 학생들과 공유하고, 둘째 학습의 증거를 수집하기 위해 질문, 교실 토론, 과제 및 활동을 실시하며, 셋째 학습자의 현재 성취수준과 학습목표 사이의 격차를 줄이기 위해 피드백을 제공하고, 넷째 동료평가(peer assessment) 활용을 포함하여 동료 학생들이 상호간의 학습자원이 되도록 하며, 다섯째 학습자 스스로 자기평가(self-assessment)를 통해 학습의 주체가 되도록 하는 전략이다. <표 II-1>은 3가지 핵심과정(학습자의 도달목표, 학습자의 현 위치, 목표달성 방법)과 평가주체(교사, 동료, 학습자)가 행해야 할 핵심전략을 제시하고 있다.

<표 II-1> 형성적 피드백의 핵심전략

평가주체	학습자의 도달목표	학습자의 현 위치	목표달성(격차 줄이기) 방법
교사	학습목표를 명시 · 공유함(1)	학습의 증거를 보여주는 효과적인 토론, 활동, 과제 시행(2)	학습자의 향상을 위한 피드백 제공(3)
동료	학습목표를 명시 · 공유함(1)	학습자들이 상호간의 학습자원으로 활동(4)	
학습자	학습목표를 명시 · 공유함(1)	학습자 스스로 학습의 주인으로 활동(5)	

다음으로 Florez & Sammons(2013)는 형성적 피드백의 실천적 전략으로 질문, 피드백, 동료평가와 자기평가 및 총합평가 결과의 형성적 활용을 들었다. 이들은 먼저 실수나 오개념으로부터 자유롭게 답할 수 있는 비위협적 대화 분위기를 조성하여 질문에 집중할 수 있게 하는 전략을 강조한다. 또한, 피드백 제공 전략으로 개인 참조 방식의 서술적 피드백, 학습목표와 연계된 교정 지침과 효과적인 실행을 담은 피드백을 사용해야 한다고 주장한다. 유사한 수준의 언어 사용을 통한 접근가능성, 이해도 및 수용력이 높은 동료평가의 적극적인 활용과 학습자 자신의 학습에 대한 자율성과 메타인지를 높이기 위한 자기평가도 중요한 실천전략이다. 더 나아가 총괄평가 결과를 형성적 피드백의 일부로 보고 학습개선 정보로 활용하는 전략이 필요하다고 하였다.

수업에서 주요한 기능을 수행하는 교사의 피드백 유형은 크게 수행 피드백(performance feedback), 동기 피드백(motivational feedback), 귀인 피드백(attributional feedback) 및 전략 피드백(strategy feedback)으로 구분된다(신중호 역, 2013). 여기에서 수행 피드백은 정답에 대한 교정적 피드백을 포함한 과제의 정확성에 대한 피드백 제공을 의미하고, 동기 피드백은 학생의 향상도와 능력 관련 정보 제공을 말한다. 귀인 피드백은 학습자 수행의 원인에 대한 귀인 정보 제공을 의미하고, 전략 피드백은 학생이 어떤 학습전략을 사용할 때 학습이 어떻게 향상되는지를 알려주는 피드백이다. 이 같은 교사의 피드백 유형을 보면, 형성적 피드백의 핵심 요인이 학습자의 교정을 위한 발판을 제공함과 동시에 학습향상에 초점을 맞춘 지원 활동을 알 수 있다.

그동안 형성적 피드백의 중요성은 꾸준히 강조되어 왔지만, 형성적 피드백 실행 척도를 개발하려는 시도는 활발하게 이루어지지 못했다. 서론에서도 언급한 바와 같이, Yan & Cheng (2015)은 ‘초등학교 교사의 형성평가에 대한 태도, 의도, 실제’라는 연구에서 형성평가의 실행을 단 두 문항으로 질문하였다. 첫 번째 문항은 “지난 6개월 동안 형성평가를 얼마나 자주 실시했습니까?”이며, 두 번째 문항은 “지난 6개월 동안 선생님이 수업 중에 실시한 수행평가 횟수가 어느 정도라고 생각하십니까?” 였다. 또한, 지은림(2009)은 ‘교사의 피드백 수행을 위한 요인구성 및 특성에 관한 연구’에서 피드백 수행 변인을 평가목적, 평가태도, 평가유형별 피드백 수행, 피드백 수행방법으로 구분하고, 5점 척도의 17개 문항을 구성하였다. 이어 지은림(2010)은 ‘교사의 피드백 수행 척도 개발 및 타당화-Rasch 모형의 적용’ 연구에서 형성평가를 위한 피드백 수행 변인을 평가목적, 평가유형별 평가수행, 평가유형별 피드백 수행, 피드백 유형, 피드백 효과로 구분하여 5점 척도의 22개 문항을 구성한 바 있다.

2. 형성적 피드백의 구성 요인

수준 높은 형성적 피드백은 교사의 교실 행동과 관련이 있고(Martin & Dowson, 2009), 학습 내용에서 학생들의 지식, 전략, 이해를 높이며(Shute, 2008), 학습향상을 도모하기 위한 중요한 도구이다(Wiliam, 2011). 이러한 형성적 피드백은 학생의 현 수준과 목표 수준의 비교를 통해 그 차이를 줄일 수 있는 구체적인 정보를 제공하고, 이후의 수행에서 발전적 학습을 할 수 있도록 하는 교육적 비계의 역할을 수행한다(Pat-El et al., 2013). 즉, 형성적 피드백은 학습목표와 평가기준에 대한 이해를 토대로 학습향상을 위해 의사소통하는 전반적인 과정을 포괄한다고 볼 수 있다(박민애, 2019).

교사의 형성적 피드백의 요소를 측정하기 위한 형성적 피드백의 구성요인은 크게 비계(scaffolding)와 모니터링(monitoring)으로 구분된다(박민애, 손원숙, 2016; Pat-El et al., 2013). Pat-El 외(2013)는 비계를 학생들이 나아가야 할(개선해야 할) 영역을 인식하도록 돕는 활동으로, 학습목표와 준거의 명료화와 수업 진행 과정의 주를 이루는 교실 내 질문행동(questioning)을 포함한다고 정의하였다. 또한, 모니터링은 학습목표에 도달하기 위해 학생의 현재 상태에 대한 피드백을 제공하고, 학생의 학습 진행 상황을 추적하는 것이라고 하였다. 즉, 모니터링이란 공유된 학습목표를 위해 학습을 최적화하기 위한 교사의 피드백 지원과 학습자의 자기 모니터링을 포함한다.

한편, Brown, Harris & Harnett(2012)는 “누가, 어떻게/언제, 무엇을, 왜 피드백 하는가?”에 관한 피드백 선행연구 검토를 통해 총 9가지 교사 피드백 인식(teachers' conceptions of feedback) 요인을 구분한 바 있다. 즉, ① 무관련 인식(irrelevance, 학생의 무시), ② 향상 인식(improvement, 학생들의 활용), ③ 책무성 인식(accountability, 기대), ④ 격려 인식(encouragement)

nt, 칭찬), ⑤ 과제수행(task, 조력), ⑥ 과정 실제(process), ⑦ 자기규제(self-regulation), ⑧ 동료/자기 피드백(peer & self), ⑨ 피드백 제공 시기(적시성, timeliness)가 그것이다. 이들 요인 중 교사가 교실 수업에서 실행하는 형성적 피드백 실제와 관련이 적은 요인인 무관련 인식과 책무성 인식을 제외하고 각 요인별 문항을 예시하면 <표 II-2>와 같다.

<표 II-2> 교사의 형성적 피드백 인식 요인과 예시 문항

요인	예시 문항
과제	<ul style="list-style-type: none"> 교사의 피드백은 학생들에게 자신의 활동을 변경해야 할 것을 알려준다. 교사의 의견은 학생들이 교사가 기대하는 일을 하는데 도움이 된다.
과정	<ul style="list-style-type: none"> 교사는 학생들에게 교사의 피드백에 반응할 기회를 준다. 피드백은 학생들과 교사와의 쌍방향 과정이다.
시기	<ul style="list-style-type: none"> 교사는 학생들이 활동을 마치면 바로 피드백을 준다. 양질의 피드백은 학생들이 학습하는 동안 교실에서 대화로 즉시 진행된다.
항상	<ul style="list-style-type: none"> 교사는 학생들에게 피드백을 제공한 후 학생 활동에서 진행 과정을 확인한다. 교사가 제공한 피드백 의견을 학생들이 활용한다.
격려	<ul style="list-style-type: none"> 교사들은 학생 활동에 대한 피드백에서 항상 칭찬을 포함해야 한다. 피드백 제공의 목적은 학생의 자존심을 향상시키는 것이다.
자기규제	<ul style="list-style-type: none"> 교사는 학생들이 자신의 활동을 스스로 평가해 보게 한다. 피드백은 학생들은 자신들의 활동을 평가하는데 도움이 된다.
동료/자기	<ul style="list-style-type: none"> 학생들은 서로 정확하고 유용한 피드백을 제공할 수 있다. 동료 학생은 최고의 피드백 자원이다.

비계는 학습자가 자신의 수준보다 높은 사고와 문제해결을 하는데 도움을 준다. Bransford, Brown & Cocking(2000)에 따르면, 비계를 활용한 목표지향적 피드백이란 학습자의 흥미를 유발하고, 과제 해결이 가능하도록 단순화시키며, 학습자들이 목표달성에 초점을 맞추도록 지침을 제공할 뿐만 아니라 학습자의 현 수준과 목표 간의 차이를 알려주고, 좌절감과 위험요인을 감소시키며, 수행해야 할 활동의 목표를 명세화하는 기능을 수행한다. 일반적으로 간단한 지침이나 단서만 제공하는 촉진적 피드백(facilitative feedback)이 구체적인 교정적 정보를 제공하는 지시적 피드백(directive feedback)보다 학습향상에 도움이 덜 될 것으로 생각하지만, 실제로 교육적 장면에서 학습자가 특정 주제나 교과 내용을 학습할 때 지시적 피드백이 더 큰 도움을 줄 수 있다(Shute, 2008). 비계가 학습 과정 중에 학습자에게 제공하는 명시적 지지가기 때문에, 비계 피드백에는 모형, 구체적인 단서와 프롬프트, 힌트, 부분적인 해결책, 직접 교수 등이 포함된다. 이 같은 지식적 피드백은 학습 초기 단계에 도움을 주기 위한 것으로 학습이 진행되면서 점차 제거된다.

한편, Pat-El 외(2013)는 학습지향적 평가 질문지(Assessment for Learning Questionnaire, AFL-Q)를 개발하였다. 이 질문지는 교사용(TAFL-Q)과 학생용(SAFL-Q)이 있고, 하위요인으로는 학생의 향상도를 추적하는 모니터링 활동(16개 문항)과 학생들이 수정이 필요한 영역을 인지하도록 도움을 주는 비계 활동(12개 문항)이 있다. 국내에서는 박민애, 손원숙(2016)이

초등학교 5학년과 6학년 학생들을 대상으로 Pat-EI 외(2013)의 SAFL-Q를 번안하여 타당화를 시도하였다. 이 척도는 원척도와 마찬가지로 2개의 하위척도 즉, 모니터링(16개 문항)과 비계(10개 문항)로 구성되어 있다. 여기에서 모니터링 척도는 얼마나 자주, 어떤 행태의 피드백이 제공되었는지, 그리고 학생의 향상 과정에 대한 모니터링이 어떻게 시도되었는지를 질문한다. 비계는 학습목표나 평가기준의 명료화 및 교사의 질문 등과 같은 수업 관련 내용으로 이루어져 있다. 박민애, 손원숙(2018)은 박민애, 손원숙(2016)의 학습을 위한 평가척도 26개 문항에 대해 Rasch 평정척도 모형을 사용하여 단축형 검사를 개발하였는데 여기에는 모니터링 7개 문항과 비계 5개 문항이 포함되어 있다.

지금까지 살펴본 형성적 피드백의 구성요인을 종합하면, 형성적 피드백의 하위요인은 크게 비계와 모니터링으로 구분된다. 비계 요인은 평가기준, 과제, 설명/발표/질문 등을 포함한 과정을 묻는 문항, 그리고 모니터링 요인은 시험결과 검토, 격려, 장단점 및 향상도 점검, 학생의 요구 등을 묻는 문항을 포함한다. 각 요인별 문항 평가 요소 및 근거를 요약하여 제시하면 <표 II-3>과 같다.

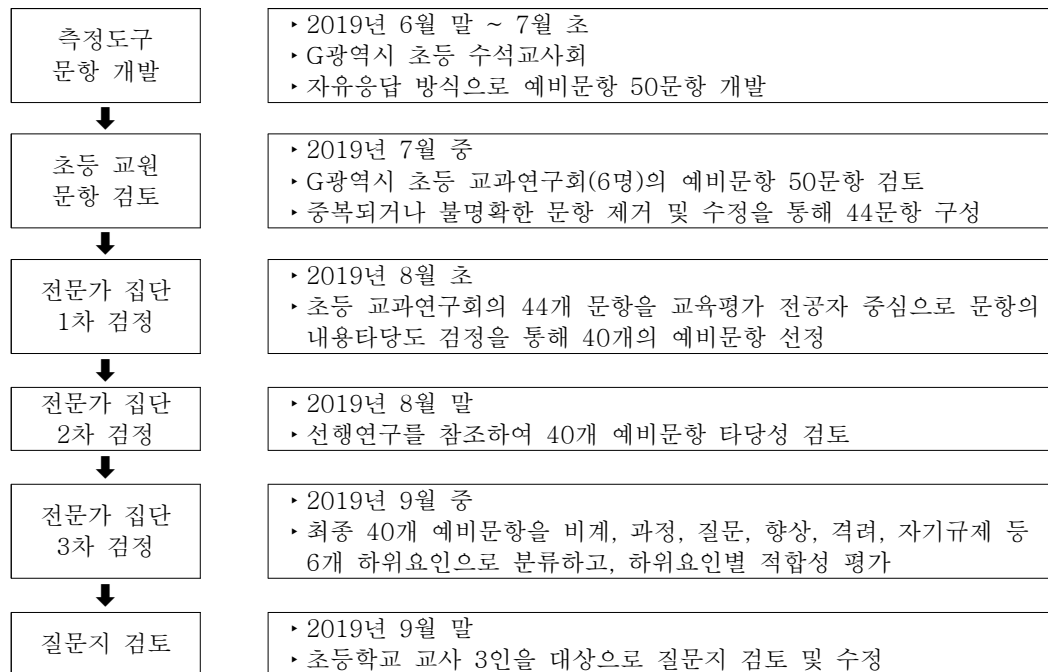
<표 II-3> 형성적 피드백의 하위 요인, 평가 요소, 그리고 근거

하위요인	평가 요소			근거
비계	◦ 과제 유형	◦ 과정 유형	◦ 시기	Brown et al.(2012)
	◦ 평가기준 제시 ◦ 과제 내용 확인	◦ 추가 설명 ◦ 발표/질문, 참여기회 부여		Pat-EI et al.(2013) 박민애, 손원숙(2016, 2018)
	◦ 과제 단순화	◦ 목표 성취 지침 제공		Bransford et al.(2000)
모니터링	• 향상 • 격려/칭찬	• 자기규제 • 동료/자기평가		Brown et al.(2012)
	• 격려 및 강점 강조 • 조언 • 학생의 요구 파악 및	• 향상도 파악 및 시험결과 검토 • 부족한 점 보충 강약점 확인		Pat-EI et al.(2013) 박민애, 손원숙(2016, 2018)
	• 좌절감 줄이기 • 흥미 유발	• 목표와 현 상태의 격차 점검		Bransford, et al.(2000)

III. 연구방법

1. 예비연구

이 연구는 2019년 6월 말부터 시작하여 9월 말까지 [그림 III-1]과 같은 예비연구 절차를 걸쳐 형성적 피드백 실행 척도 예비문항을 개발하였다.



[그림 III-1] 예비연구 절차

먼저, G광역시 초등 수석교사회에서 이 연구의 목적을 설명하고, 형성적 피드백 핵심전략과 피드백 유형 및 지침 등의 자료를 제공하였다. 이를 토대로 수석교사들에게 성공적인 수업과 질 높은 교실평가를 위해 그동안 사용하였거나 사용해야 한다고 보는 형성적 피드백에 대해 사전에 제공한 반구조화 설문지에 자유응답 방식으로 기술하도록 하였고, 이 내용을 정리하여 50개의 예비문항을 구성하였다. 이 예비문항을 G광역시 초등 교과연구회 교사 6명에게 검토를 의뢰하여 중복되거나 의미가 불명료한 문항을 수정·보완·제거하여 44개 문항으로 정리하였다. 다음으로 교육평가 교수 및 박사 등 전문가를 중심으로 세 차례의 검정 작업을 거쳤다. 1차 검정에서는 문항의 내용타당도 점검을 통해 4개 문항을 삭제하여 40개의 예비문항을 선정하였고, 2차 검정에서는 형성적 피드백과 관련된 선행연구를 참조하면서 40개 예비문항의 타당성을 검토하였다. 3차 검정에서는 최종 예비문항(40개 문항)을 6개의 하위요인 즉, 비계의 계획요인, 과정요인, 질문요인, 그리고 모니터링의 향상요인, 격려요인, 자기규제요인으로 명명 및 분류하였고, 하위요인별 적합성을 평가하였다. 마지막으로 초등교원 3명이 질문지에 응답하면서 질문지의 형식과 문항의 내용 등을 검토한 후 일부 문항을 수정하였다. 예비연구의 통해 구성한 최종 예비문항(40개 문항)은 <표 III-1>과 같다. 이 예비척도의 특징은 초등 교사들이 비계와 관련하여 계획과 질문이 중요하다는 점에 합의하여 하위요인 및 문항에 이를 포함시켰다는 점이다.

<표 III-1> 최종 예비문항(40문항)

구분	하위요인	예비문항
비계 (21개 문항)	계획요인(6문항)	1. 학생들이 자기들의 활동을 수정할 수 있도록 구체적인 정보를 준비한다. 4. 학생들이 과제를 기대하는 수준으로 완성하는데 도움을 주기 위한 피드백을 준비한다. 7. 학생들의 수준을 고려하여 모둠을 구성한다. 10. 학생들의 활동결과를 확인할 수 있는 다양한 평가방법을 계획한다. 14. 수업의 핵심개념과 관련된 내용을 확인할 수 있도록 수업을 설계한다. 18. 학생들의 개별 활동을 평가하고 피드백할 시간을 미리 준비해둔다.
	과정요인(9문항)	2. 학생들에게 교사의 피드백에 반응할 기회를 준다. 5. 학생들의 수준에 적합한 피드백을 주고받는다. 8. 학생들이 공부할 문제를 읽고 해결방법을 스스로 찾도록 안내한다. 11. 문제해결을 위한 학습 활동 순서를 학생들이 이해할 수 있도록 안내한다. 12. 문제해결 과정에서 학생들의 사고를 촉진하기 위해 단계적이고 명료하게 시범을 보여준다. 15. 문제해결 방법을 학생들에게 직접 제시하기보다 단서(실마리)를 먼저 제시한다. 16. 학습의 과정이 드러나도록 판서를 구조화시킨다. 19. 동료 학생은 최고의 피드백 자원이다. 20. 친구들과 모둠 활동을 통해 상호협력하면서 문제해결을 하도록 안내한다.
	질문요인(6문항)	3. 공부할 문제를 해결하기 위해 학생들의 배경지식을 활성화시키는 질문을 한다. 6. 공부할 문제가 어떤 역량과 관련이 있는지 확인하는 질문을 한다. 9. 공부할 문제가 학생들의 실생활과 관련이 있는지 확인하는 질문을 한다. 13. 학습활동 간의 유기적 관계가 문제해결에 도움이 되는지 확인하는 질문을 한다. 17. 학생들의 사고력을 촉진하기 위해 학생들의 경험과 관련된 질문을 한다. 21. 학생들의 수준에 적합한 구체적이고 명료한 질문을 한다.
	향상요인(8문항)	22. 피드백은 학생들의 학습에 큰 도움을 주기 때문에 중요하다. 23. 학생들의 활동을 수정하기 위해 피드백을 제공한다. 26. 학생들에게 피드백을 제공한 후 학습 활동의 진행과정을 점검한다. 29. 학생들이 공부할 문제를 보고 자신이 해결할 수 있는 수준의 문제인지 확인하게 한다. 30. 공부할 문제와 관련하여 경험해 본 적이 있는 방법이 있는지 확인한다. 33. 학생들이 문제를 해결하는 도중에 어려움에 직면했을 때 조인자로서 적절히 개입한다. 36. 학생들의 학습개선을 위해 적합한 수업매체를 활용했는지 확인한다. 39. 수업에 필요한 자료를 적절한 시기에 제시하고 활용한다.
모니터링 (19개 문항)	격려요인(5문항)	24. 학생들에게 칭찬하고 격려하는 긍정적인 피드백을 제공한다. 27. 학생들이 자신들의 활동에 쏟은 노력에 적극적으로 피드백한다. 31. 학생들의 학습동기와 자존감을 높이는 피드백을 제공한다. 34. 학생들의 질문을 경청하고 의견을 존중하는 편이다. 37. 학생들과의 친밀감과 유대감을 기반으로 허용적인 분위기를 조성한다.
	자기규제요인(6문항)	25. 학생들이 자신의 활동을 스스로 평가해 보도록 한다. 28. 학생들이 스스로 학습향상을 위한 아이디어를 내보도록 한다. 32. 피드백이 학생들 스스로 자신들의 활동을 평가하는데 도움이 된다. 35. 학생들이 교사의 지시가 없어도 스스로 자신들의 활동을 분석한다. 38. 학생들이 상호 간에 유용한 피드백을 제공할 수 있다. 40. 학생들이 스스로 자신들의 실수를 찾아보게 한다.

비계와 모니터링의 하위요인별 의미를 살펴보면, 먼저 비계의 계획요인은 학생의 학습 효과의 변화를 촉진할 수 있는 형성적 피드백에 대한 사전 계획 및 설계의 실행 정도를 의미하며, 과정요인은 수업이 진행되는 과정에서 실행되는 형성적 피드백이고, 질문요인은 학생의 사고력을 자극 및 촉진할 수 있는 교사의 발문에 대한 형성적 피드백이다. 다음으로 모니터링의 향상요인은 학생의 학습능력을 향상시키기 위한 교사의 교수 개선과 관련된 형성적 피드백이며, 격려요인은 교실에서 학생이 자신을 긍정적으로 느낄 수 있도록 도와주는 교사의 지지와 격려를 가리킨다. 자기규제요인은 교사의 형성적 피드백이 학생이 자신의 일을 스스로 평가하는데 도움을 줄 수 있는 피드백이다.

2. 연구대상

이 연구에서는 본 연구를 위해 G광역시 초등학교에 재직하고 있는 교사 470명의 자료를 임의표집 방식으로 수집하였다. 470명 중에서 불성실한 응답, 무응답 등을 제외한 결과 교사 443명의 자료가 분석에 포함되었는데, 이 중에서 다변량 정규성을 벗어난 사례 34명을 제외하고 총 409명의 교사 응답을 분석하였다(<표 III-2> 참조). 이 연구에서 초등학교 교사를 연구대상으로 선정한 이유는 초등학교 교사들이 범교과적인 교수-학습 활동에서 학생들의 학습향상을 위해 형성적 피드백을 더 적극적이고 능동적이므로 실행한다고(안희상, 손원숙, 2017; Brown et al., 2012) 보았기 때문이다.

<표 III-2> 조사대상의 일반적 특성

구분		사례수	%
성별	남교사	35	20.8
	여교사	324	79.2
교육경력	10년 이하	127	31.1
	20년 이하	209	51.1
	20년 초과	73	17.8
교육평가 연수 경험	없음	205	50.1
	있음	204	49.9
전체		409	100.0

조사대상자의 일반적 특성을 보면, 여교사의 비율이 79.2%로 남교사보다 높았고, 교육경력이 10년을 초과하고 20년 이하인 교사의 비율이 51.1%로 가장 높았다. 교육평가와 관련한 연수를 15시간(1학점) 이상 경험한 교사의 비율과 그렇지 않은 교사의 비율은 거의 유사한 수준을 보였다.

3. 측정도구

이 연구의 측정도구는 예비연구에서 개발한 초등학교 교사용 예비척도와 이 예비척도의 공인타당도 검정을 위해 사용한 기존검사로 구분된다. 기존검사로선 먼저 형성적 피드백 경험을 측정한 박민애, 손원숙(2018)의 평가 척도를 활용하였다. 이 척도는 비계 5문항과 모니터링 7문항의 하위요인을 포함한 단축형 학습을 위한 평가척도(AFL-Q)이며, α 계수는 비계 0.79, 모니터링 0.89, 그리고 단축형 검사 전체의 α 계수는 0.91이었다. 이 연구에서 산출한 α 계수는 비계 0.80, 모니터링 0.83, 그리고 검사 전체의 α 계수는 0.87이었다.

예비척도의 공인타당도 검정을 위해 활용한 두 번째 검사는 Brown(2011)의 평가인식 척도

(Students' Conceptions of Assessment: SCoA-VI)이다. 이 척도는 평가인식 측정을 위한 25 문항으로 구성되어 있는데, 이 연구에서는 항상 영역에 속한 교수개선 5문항과 학습향상 5문항을 사용하였다. Brown(2011)이 보고한 α 계수는 0.78이었고, 이 연구에서 산출한 α 계수는 교수개선 0.86, 학습향상 0.86, 검사 전체의 α 계수는 0.91이었다. 위에서 언급한 기존검사 관련 내용을 표로 정리하면 <표 III-3>과 같다.

<표 III-3> 교사용 형성적 피드백 기존 검사의 내용과 신뢰도

영역	문항 예시	문항 수	α 계수
형성적 피드백	비계 <ul style="list-style-type: none"> · 학생이 수업에 활발하게 참여할 수 있는 허용적인 분위기를 만든다. · 학생에게 과제의 평가기준을 알려준다. 	5	0.80
	모니터링 <ul style="list-style-type: none"> · 학생의 공부를 도와주는 조언을 한다 · 학생의 부족한 부분을 보충할 수 있는 방법을 학생과 함께 고민한다. 	7	0.83
평가 인식	교수개선 <ul style="list-style-type: none"> · 다음에 더 잘하기 위해 평가 결과에 주의를 기울인다. · 평가 결과를 활용하여 이후 가르칠 내용을 파악한다. 	5	0.86
	학습향상 <ul style="list-style-type: none"> · 평가가 학생들의 학습을 향상시키는데 도움이 된다. · 평가 결과를 보고 주후 공부해야 할 내용을 검토한다. 	5	0.86

4. 자료 분석

이 연구의 자료 분석은 크게 두 가지 유형으로 구분된다. 첫째, 형성적 피드백 실행에 대한 선행연구와 초등학교 수석교사 및 교과연구회, 교육평가 전문가 집단에서 제안한 내용을 근거로 형성적 피드백 실행의 하위요인을 구분하고 예비척도 40개 문항에 대해 각 하위요인별로 평정척도 분석을 실시하였다(신현숙, 염시창, 2017). 이 분석을 위해 Meyer(2014, 2018)의 j METRIK 4.1.1을 사용하였으며, 먼저 Rasch 모형의 중요 가정인 일차원성 가정을 검토하였고 이어 각 요인별로 표준화 잔차를 이용한 주성분분석을 실시하였다(Wolfe & Smith, 2007). Linacre(2017)의 기준에 따라 첫 번째 대비의 고유값이 3미만이면 일차원성 가정이 충족되는 것으로 해석하였다. 또한, 데이터가 Rasch 모형의 일차원성 기댓값으로부터 어느 정도 벗어났는가를 검토하기 위해 외적합도(outfit)와 내적합도(infit)를 확인하였다(설현수, 2007; Bond & Fox, 2015). 여기에서 적합도 평균제곱(mean-square)의 수용가능 범위는 Wright와 Linacre(1994)가 평정척도에 적합한 범위로 제안한 0.6~1.4를 적용하였다. 다음으로 점이연 상관계수(point-measure correlation)의 0.4이상 기준을 적용하였다. 또한, 잔차부하량이 0.6보다 높아 구인타당도를 저해하거나 문항제거 시 α 계수가 문항 전체의 α 계수보다 높은 문항을 제거 대상 문항으로 분류하였다.

둘째, 하위요인별 평정척도 분석으로 도출된 문항에 대해 AMOS 22.0(Arbuckle, 2013)을 활용하여 확인적 요인분석을 실시한 후 문항을 확정하였다. 확인적 요인분석의 모형 적합도 판정을 위해서는 RMSEA 0.80 이하(Browne & Cudeck, 1993), CFI 0.90 이상 및 SRMR 0.80

이하 기준(Hu & Bentler, 1999)을 적용하였다. 또한, 구인(하위요인)과 측정변수 간의 일치성 정도를 확인하기 위해 구인신뢰도(construct reliability, CR), 평균분산추출(average variance extracted, AVE)을 산출하였으며, 각각 0.7이상, 0.5이상이면 적절한 수준으로 판단하였다(Hair et al., 2006). 마지막으로 하위요인 내의 하위변수 간 상관계수와 기술통계량을 산출하였으며, 기존검사와의 상관계수를 통해 공인타당도를 검토하였다.

IV. 연구결과

1. 요인별 Rasch 평정척도 분석

비계의 계획요인 문항은 6개 문항으로 α 계수는 0.77이었고, Rasch 측정모형의 피험자 신뢰도(person reliability)는 0.75였으며, 피험자 분리지수(separation index)는 1.75였다. Rasch 모형의 일차원성 확인을 위해 표준화 잔차를 활용한 주성분분석을 실시한 결과, 첫 번째 요인의 아이겐 값이 1.62로 3.0보다 작아 일차원성을 지지하였다. 또한, 계획요인 문항적합도와 잔차부하량 등을 표로 제시하면 <표 IV-1>과 같다.

<표 IV-1> 비계의 계획요인 문항적합도와 잔차부하량

원척도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
1	0.11	0.74	0.77	0.62	0.56	0.719
4	-0.43	0.73	0.73	0.62	0.41	0.723
7	0.39	1.68	1.68	0.42	-0.14	0.775
10	0.32	0.82	0.85	0.63	-0.52	0.712
14	-0.80	0.81	0.85	0.59	0.66	0.729
18	0.41	1.07	1.11	0.56	-0.65	0.727

비계의 계획요인 문항에 대해 내적합도와 외적합도 판정준거 0.6~1.4를 적용한 결과 7번 문항이 제거 대상 문항인 것으로 나타났다. 또한, 점이연상관은 모두 0.40보다 높아 기준에 부합하였다. 한편, 14번 문항과 18번 문항의 잔차부하량이 0.6을 초과하여 제거 대상 문항이었지만 다른 측정값 상에서 큰 문제가 없어서 일단 제거하지 않는 것으로 판정하였다. 7번 문항의 경우 문항 제거 시 α 계수도 0.77보다 높아 제거 대상 문항(음영처리)으로 확정하였다.

둘째, 비계의 과정요인 문항은 9개 문항으로 α 계수는 0.81이었고, Rasch 측정모형의 피험자 신뢰도는 0.78이었으며, 피험자 분리지수는 1.90이었다. 일차원성 확인을 위한 주성분분석 결과를 보면, 첫 번째 요인의 아이겐 값은 1.49로 3.0보다 작아 일차원성을 지지하였다. 비계

의 과정요인 문항적합도와 잔차부하량 등은 <표 IV-2>와 같다.

<표 IV-2> 비계의 과정요인 문항적합도와 잔차부하량

원척도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
2	-0.24	0.73	0.75	0.62	0.57	0.787
5	-0.19	0.75	0.76	0.61	0.59	0.788
8	1.00	1.12	1.12	0.54	-0.23	0.791
11	0.40	1.12	1.11	0.61	-0.41	0.784
12	-0.44	0.84	0.86	0.57	-0.07	0.791
15	0.02	0.98	0.99	0.55	0.05	0.791
16	0.55	1.28	1.34	0.50	-0.60	0.798
19	-0.38	1.15	1.15	0.51	0.02	0.796
20	-0.72	0.90	0.95	0.58	0.48	0.790

비계의 과정요인 문항에 대해 내적합도와 외적합도를 판정한 결과 준거를 벗어난 문항은 없었고, 점이연상관 또한 매우 높아 적절한 것으로 나타났다. 다만, 16번 문항의 경우 잔차부하량이 0.60으로 판정 경계선인 0.60상에 놓여 있지만, 문항제거 시 α 계수가 전체 문항의 0.81 보다는 낮아 제거 대상 문항에서 제외하였다.

셋째, 비계의 질문요인 문항은 6개 문항으로 α 계수는 0.76이었고, Rasch 측정모형의 피험자 신뢰도는 0.74였으며, 피험자 분리지수는 1.67이었다. 주성분분석 결과, 첫 번째 요인의 아이겐 값은 1.62로 일차원성을 지지하였다. 비계의 질문요인 문항적합도와 잔차부하량 등은 <표 IV-3>과 같다.

<표 IV-3> 비계의 질문요인 문항적합도와 잔차부하량

원척도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
3	-0.43	0.80	0.85	0.62	0.24	0.719
6	1.46	1.31	1.38	0.46	-0.81	0.756
9	-0.01	0.98	0.98	0.61	-0.23	0.711
13	0.51	0.87	0.90	0.61	-0.13	0.712
17	-1.12	1.09	1.09	0.53	0.71	0.735
21	-0.40	0.86	0.88	0.59	0.56	0.723

비계의 질문요인 문항에 대해 내적합도와 외적합도 판정준거를 적용한 결과 범위를 넘어선 문항은 없었고, 점이연상관은 모두 0.40보다 높았다. 6번과 17번 문항의 경우 잔차부하량이 0.60을 초과하였지만, 17번 문항은 문항제거 시 α 계수 전체 문항의 α 계수인 0.76보다 현저하게 낮은 0.735로 나타나 제거하지 않은 반면, 6번 문항은 α 계수가 0.756으로 제거 대상 문항(음영처리)으로 판정하였다.

넷째, 모니터링의 향상요인 문항은 8개 문항으로 α 계수는 0.82였고, Rasch 측정모형의 피험자 신뢰도는 0.80이었으며, 피험자 분리지수는 2.00이었다. 주성분분석 결과, 첫 번째 요인의 아이겐 값은 1.84로 일차원성을 지지하였고 모니터링의 향상요인 문항적합도와 잔차부하량 등은 <표 IV-4>와 같다.

<표 IV-4> 모니터링의 향상요인 문항적합도와 잔차부하량

원칙도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
22	-0.97	0.96	0.91	0.61	0.65	0.800
23	-0.55	0.79	0.83	0.67	0.60	0.793
26	0.05	0.85	0.82	0.67	0.11	0.790
29	1.87	1.46	1.48	0.46	-0.80	0.820
30	0.38	1.08	1.06	0.61	-0.41	0.797
33	-0.77	0.80	0.90	0.64	0.26	0.798
36	0.48	1.05	1.07	0.59	-0.26	0.799
39	-0.50	0.90	0.89	0.66	0.34	0.792

모니터링의 향상요인 문항의 점이연상관은 모두 0.40을 넘어서서 양호한 수준을 보였다. 다만, 내적합도와 외적합도 판정준거를 적용한 결과 29번 문항의 적합도가 기준을 넘어서서 제거 대상 문항임을 확인할 수 있었다. 또한, 22번 문항과 29번 문항의 경우 잔차부하량이 0.60을 넘어섰고, 29번 문항의 문항제거 시 α 계수 또한 전체 문항의 α 계수인 0.82와 동일한 값을 보여 22번과 29번 문항을 제거 대상 문항(음영처리)으로 판정하였다.

다섯째, 모니터링의 격려요인 문항은 5개 문항으로 α 계수는 0.81이었고, Rasch 측정모형의 피험자 신뢰도는 0.78이었으며, 피험자 분리지수는 1.86이었다. 주성분분석 결과, 첫 번째 요인의 아이겐 값은 1.53으로 일차원성을 지지하였으며, 모니터링의 향상요인 문항적합도와 잔차부하량 등은 <표 IV-5>와 같다.

<표 IV-5> 모니터링의 격려요인 문항적합도와 잔차부하량

원칙도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
24	-0.33	0.97	0.92	0.69	0.52	0.764
27	0.21	1.06	1.07	0.61	0.69	0.783
31	0.34	0.75	0.74	0.74	0.18	0.749
34	-0.65	0.88	0.87	0.70	-0.44	0.767
37	-0.41	1.05	1.05	0.61	-0.71	0.787

모니터링의 격려요인에서는 모든 문항의 내·외적합도가 판정준거의 범위를 벗어나지 않았고, 문항의 점이연상관은 모두 0.40을 넘어서서 양호한 수준을 보였다. 문항제거 시 α 계수 또

한 문제가 되는 수준을 보인 문항이 없었다. 그러나 27번과 37번 문항의 경우 잔차부하량이 0.60을 초과하여 제거 대상 문항이었지만, 격려요인의 문항 수가 다른 요인에 비해 적어 일단 37번 문항만 제거하고(음영처리) 27번 문항은 유지시키는 것으로 판정하였다.

여섯째, 모니터링의 자기규제요인 문항은 6개 문항으로 α 계수는 0.79였고, Rasch 측정모형의 피험자 신뢰도는 0.77이었으며, 피험자 분리지수는 1.82였다. 주성분분석 결과, 첫 번째 요인의 아이겐 값은 1.46으로 일차원성을 지지하였다. 모니터링의 향상요인 문항적합도와 잔차부하량 등은 <표 IV-6>과 같다.

<표 IV-6> 모니터링의 자기규제요인 문항적합도와 잔차부하량

원척도 문항번호	측정치	내적합도	외적합도	점이연상관	잔차부하량	문항제거 시 α 계수
25	-0.43	0.99	1.01	0.55	0.57	0.769
28	0.15	1.03	1.03	0.62	0.01	0.751
32	-0.88	0.78	0.78	0.65	0.42	0.753
35	1.59	1.15	1.20	0.56	-0.78	0.765
38	-0.19	0.99	0.97	0.64	-0.43	0.746
40	-0.10	0.98	0.95	0.58	0.41	0.761

모니터링의 자기규제요인에서는 모든 문항의 내적합도와 외적합도가 판정준거의 범위를 벗어나지 않았고, 문항의 점이연상관도 모두 0.40을 넘어서서 양호한 수준을 보였다. 문항제거 시 α 계수 또한 문제가 되는 수준을 보이지 않았다. 그러나 35번 문항은 잔차부하량이 0.60을 초과하여 제거 대상 문항(음영처리)으로 판정하였다.

2. 확인적 요인분석 및 상관분석

가. 확인적 요인분석

이 연구에서는 비계의 3요인 모형과 모니터링 3요인 모형에 대한 확인적 요인분석을 실행하였다. 그 결과, 모니터링 3요인 모형에서는 문제가 없었지만, 비계의 경우 수렴이 이루어지지 못한 문제가 발생하였고, 이 문제를 해결하기 위해 기존의 과정요인과 질문요인을 동일한 구인으로 보고 비계 2요인 모형을 설정하였다. 실제로 기존의 과정요인과 질문요인을 새로운 과정요인으로 설정하여 Rasch 평정척도 모형의 주성분분석을 실시한 결과, 첫 번째 요인의 아이겐 값이 1.95로 일차원성을 지지하였다.

비계 2모형과 모니터링 3요인 모형에 대한 확인적 요인분석을 실시한 결과, 비계 모형의 적합도는 $\chi^2(34)=72.026$, $p < .001$, RMSEA=0.052, CFI=0.973, SRMR=0.033으로 매우 양호한 수준을 보였고, 모니터링 모형의 적합도 또한 $\chi^2(51)=183.604$, $p < .001$, RMSEA=0.080, CFI

=0.931, SRMR=0.043으로 매우 양호하였다. 또한, 모든 비표준화계수가 .001수준에서 유의한 것으로 나타났다. 다음으로 이 연구에서는 각 모형별로 기여도가 낮은 문항을 제거하기 위해 표준화계수 0.60 기준을 적용하였다. 그 결과, <표 IV-7>과 같이 하위요인별로 제거 대상(음영처리) 문항을 확인할 수 있었다.

<표 IV-7> 하위요인별 문항 표준화계수

계획요인		과정요인		향상요인		격려요인		자기규제요인	
문항	계수	문항	계수	문항	계수	문항	계수	문항	계수
1	0.66	2	0.63	23	0.65	24	0.70	25	0.62
4	0.66	5	0.65	26	0.71	27	0.68	28	0.63
10	0.62	8	0.56	30	0.57	31	0.76	32	0.69
14	0.69	11	0.59	33	0.63	34	0.63	38	0.63
18	0.58	12	0.57	36	0.56			40	0.58
		15	0.56	39	0.67				
		16	0.52						
		19	0.49						
		20	0.61						
		3	0.59						
		9	0.58						
		13	0.62						
		17	0.63						
		21	0.65						

먼저, 비계의 계획요인에서는 18번 문항이 제거되었고, 과정요인에서는 총 8개 문항이 제거 대상 문항으로 선정되었다. 또한, 모니터링의 향상요인의 2개 문항과 자기규제요인의 1개 문항의 표준화계수가 0.60에 미달하여 제거 대상 문항으로 판정하였다. 그 결과, 비계의 계획요인 4개 문항, 과정요인 6개 문항, 그리고 모니터링의 향상요인 4개 문항, 격려요인 4개 문항, 자기규제요인 4개 문항이 최종 문항으로 확정되었다.

다음으로 이 연구에서는 두 가지 모형을 설정하여 확인적 요인분석을 시도하였다. 첫 번째 모형은 계획요인, 과정요인, 향상요인, 격려요인, 자기규제요인을 1차 요인으로 설정한 1차 5요인 모형이다. 두 번째 모형은 1차 모형에 비계와 모니터링을 2차 요인으로 설정한 2차 2요인 1차 5요인 모형이다. 이들 모형 분석에서 산출된 모형의 적합도 지수는 <표 IV-8>과 같다. 두 모형의 적합도 지수를 비교해 보면, 적합도에 큰 차이가 없다는 점을 알 수 있다. 따라서 이 연구에서는 해석의 편의성을 고려하여 1차 5요인 모형의 분석 결과를 해석하였다.

<표 IV-8> 모형의 적합도 비교

모형	χ^2	df	<i>p</i>	RMSEA	CFI	SPMR	AIC	BIC
1차 5요인	510.341	199	.000	0.062	0.920	0.042	618.340	835.081
2차 2요인 1차 5요인	518.100	204	.000	0.061	0.919	0.042	616.100	812.772

또한, <표 IV-9>처럼, 5개 요인의 구인신뢰도(construct reliability, CR)는 0.73~0.81로 0.70보다 높았고, 평균분산추출(average variance extracted, AVE)은 각각 0.40~0.48로 0.50보다 약간 낮은 수준을 보였다. α 계수는 0.72~0.79 수준이었고, 모든 비표준화계수가 유의하였으며, 표준화계수도 2개 문항을 제외하고 0.60을 넘어선 것으로 나타났다.

<표 IV-9> 확인적 요인분석 결과

하위요인	원척도 문항 번호	비표준화계수	표준화계수	표준오차	<i>t</i>	타당도 및 신뢰도
계획 (4개 문항)	1	—	0.68	—	—	CR=0.76
	4	0.90	0.66	0.07	12.11***	AVE=0.44
	10	0.97	0.59	0.09	10.96***	α =0.75
	14	0.96	0.71	0.07	12.96***	
비계 과정 (6개 문항)	2	—	0.63	—	—	CR=0.81
	5	1.07	0.67	0.09	11.56***	AVE=0.41
	20	1.09	0.65	0.10	11.38***	α =0.80
	13	1.07	0.58	0.10	10.27***	
	17	1.05	0.63	0.10	11.02***	
	21	1.11	0.69	0.09	11.91***	
항상 (4개 문항)	23	—	0.65	—	—	CR=0.77
	26	1.21	0.72	0.10	12.60***	AVE=0.45
	33	0.91	0.63	0.08	11.21***	α =0.76
	39	1.11	0.68	0.09	11.96***	
모니터링 격력 (4개 문항)	24	—	0.70	—	—	CR=0.79
	27	1.01	0.68	0.08	12.69***	AVE=0.48
	31	1.07	0.76	0.08	13.96***	α =0.79
	34	0.83	0.63	0.07	11.85***	
자기규제 (4개 문항)	25	—	0.63	—	—	CR=0.73
	28	1.15	0.61	0.11	10.51***	AVE=0.40
	32	1.03	0.69	0.09	11.65***	α =0.72
	38	1.11	0.60	0.11	10.44***	

*** $p < .001$.

나. 상관분석

확인적 요인분석에서 확정된 문항에 대해 각 요인별로 하위요인 합성점수를 측정변수로 설정하여 상관계수와 기술통계량을 계산하였는데, 그 결과는 <표 IV-10>과 같다. 먼저 비계의 계획요인과 과정요인 간에는 0.77로 매우 높은 상관이 있었고 모니터링의 향상요인, 격려요인, 자기규제요인 간에도 0.70~0.72의 상관이 있었다. 그러나 비계의 과정요인과 모니터링의 향상요인 및 격려요인이 상관도 각각 0.74, 0.71로 비교적 높은 수준을 보였다. 한편, 하위변수의 왜도는 0.07~0.17로 2.0보다 낮았고, 첨도 또한 0.25~0.72로 8.0보다 크게 낮아 정규성 분포에서 이탈하지 않는 것으로 해석되었다.

<표 IV-10> 형성적 피드백 실행 척도의 하위요인 간 상관계수와 기술통계량(N=409)

구분	하위요인(문항 수)	①	②	③	④	⑤
비계	①계획요인(4개)	-				
	②과정요인(6개)	0.77***	-			
모니터링	③향상요인(4개)	0.71***	0.74***	-		
	④격려요인(4개)	0.65***	0.71***	0.72***	-	
	⑤자기규제요인(4개)	0.64***	0.69***	0.70***	0.71***	-
기술통계량	최소값	10.00	13.00	12.00	11.00	10.00
	최대값	20.00	30.00	20.00	20.00	20.00
	평균	16.07	24.97	16.91	16.86	16.15
	표준편차	2.11	2.83	1.97	2.12	2.10
	왜도	-0.07	-0.16	-0.10	-0.17	-0.10
	첨도	-0.25	0.43	-0.72	-0.72	-0.27

*** $p < .001$.

다음으로 공인타당도 검정을 위해 <표 IV-11>과 같이 두 가지 척도와의 상관계수를 산출하였다. 먼저, 박민애, 손원숙(2018)의 학습을 위한 평가척도(단축형) 하위요인과의 상관을 보면, 비계의 하위요인들과 준거변수의 비계와의 상관이 모니터링의 하위요인들과의 상관에 비해 상대적으로 낮았고, 모니터링의 자기규제요인은 박민애, 손원숙(2018)의 모니터링과 높은 상관을 보였지만 향상과 격려요인은 오히려 준거변수의 비계와 높은 상관을 보였다. 또한 Brown(2011)의 교수개선 및 학습향상과의 상관에서는 비계와 모니터링의 모든 하위요인이 0.59~0.67의 높은 상관을 보였다. 이렇게 상관이 높다는 것은 이 연구에서 개발된 초등학교 교사의 형성적 피드백 실행 척도가 형성적 피드백의 기본 요소인 교수개선 및 학습향상과의 공인타당도가 양호하다는 점을 의미한다.

<표 IV-11> 준거변수와의 상관

준거변수		비계		모니터링		
		계획요인	과정요인	향상요인	격려요인	자기규제요인
박민애, 손원숙 (2018)	비계	0.57***	0.64***	0.64***	0.63***	0.55***
	모니터링	0.61***	0.64***	0.60***	0.59***	0.69***
Brown (2011)	교수개선	0.64***	0.67***	0.65***	0.60***	0.64***
	학습향상	0.63***	0.64***	0.66***	0.60***	0.59***

*** $p < .001$.

예비문항에 대한 하위요인별 평정척도 분석과 확인적 요인분석의 결과를 반영한 초등학교 교사를 위한 형성적 피드백 실행 척도는 <표 IV-12>와 같이 최종 22문항으로 확정되었다. 각 문항의 진술문은 ‘매우 그렇지 않다(1점)’, ‘그렇지 않은 편이다(2점)’, ‘보통이다(3점)’, ‘그런 편이다(4점)’, ‘매우 그렇다(5점)’의 5점 Likert 척도로 구성되었다.

<표 IV-12> 초등학교 교사를 위한 형성적 피드백 실행 척도

구분	하위요인	문항
비계 (10개 문항)	계획요인(4개 문항)	1. 학생들이 자기들의 활동을 수정할 수 있도록 구체적인 정보를 준비한다.(1) 2. 학생들이 과제를 기대하는 수준으로 완성하는데 도움을 주기 위한 피드백을 준비한다.(4) 3. 학생들의 활동결과를 확인할 수 있는 다양한 평가방법을 계획한다.(10) 4. 수업의 핵심개념을 확인할 수 있도록 수업을 설계한다.(14)
	과정요인(6개 문항)	5. 학생들에게 교사의 피드백에 반응할 기회를 준다.(2) 6. 학생들의 수준에 적합한 피드백을 주고받는다.(5) 7. 친구들과 모둠 활동을 통해 상호협력하면서 문제해결을 하도록 안내한다.(20) 8. 학습 활동 간의 유기적 관계가 문제해결에 도움이 되는지 확인하는 질문을 한다.(13) 9. 학생들의 사고력을 촉진하기 위해 학생들의 경험과 관련된 질문을 한다.(17) 10. 학생들의 수준에 적합한 구체적이고 명료한 질문을 한다.(21)
	향상요인(4개 문항)	11. 학생들의 활동을 수정하기 위해 피드백을 제공한다.(23) 12. 학생들에게 피드백을 제공한 후 학습활동의 진행과정을 점검한다.(26) 13. 학생들이 문제해결 과정에서 어려움에 직면했을 때 조언자로서 적절히 개입한다.(33) 14. 수업에 필요한 자료를 적절한 시기에 제시하여 활용한다.(39)
	모니터링 (12개 문항)	15. 학생들을 칭찬하고 격려하는 긍정적인 피드백을 제공한다.(24) 16. 학생들이 자신들의 활동에 쏟은 노력에 적극적으로 피드백을 한다.(27) 17. 학생들의 학습동기와 자존감을 높이는 피드백을 제공한다.(31) 18. 학생들의 질문을 경청하고 의견을 존중하는 편이다.(34)
모니터링 (12개 문항)	격려요인(4개 문항)	19. 학생들이 자신의 활동을 스스로 평가해 보도록 안내한다.(25) 20. 학생들이 스스로 학습향상을 위한 아이디어를 내보도록 한다.(28) 21. 피드백이 학생들 스스로 자기활동을 평가하는데 도움이 된다.(32) 22. 학생들이 상호간에 유용한 피드백을 제공한다.(38)

주) 문항의 괄호 안의 번호는 예비검사 문항 번호임.

V. 논의 및 결론

이 연구는 초등학교 교사의 형성적 피드백 실행 척도를 개발하여 교실 내에서 수업 중에 이루어지는 형성평가의 실재를 강화하는데 도움을 주기 위해 수행되었다. 이 연구의 분석 결과, 초등학교 교사의 형성적 피드백 실행 척도는 비계의 계획요인, 과정요인, 그리고 모니터링의 향상요인, 격려요인, 자기규제요인 등 5개의 하위요인으로 구성되었다. 또한 하위요인별로 높은 신뢰도를 보였고, 기존 검사와의 공인타당도 분석 결과도 양호한 것으로 나타났다. 이 연구의 절차를 요약하면 다음과 같다.

먼저, 예비연구에서는 G광역시 수석교사회 구성원들에게 연구의 목적을 설명하고 형성적 피드백 관련 자료를 제공한 다음, 그동안 수석교사들이 수업 중에 사용하였거나 사용해야 할 형성적 피드백을 제공된 반구조화 설문지에 자율응답 방식으로 기술하도록 하여 50개 문항을 구성하였다. 이 문항을 G광역시 초등 교과연구회 교사들의 검토 과정을 거쳐 중복되거나 불명료한 문항을 제거 또는 수정하여 44개 문항으로 정리하였고, 교육평가 전문가들의 예비척도 내용타당도 검토를 통해 6개 하위요인의 40개 문항을 구성하였다. 이 예비척도를 초등교원 3인을 대상으로 질문지를 검토하여 일부 문항내용을 수정하여 최종적으로 40개 문항을 예비척도 문항으로 확정하였다.

이 연구의 본 연구에서는 G광역시 초등학교 교사 470명으로부터 자료를 임의 표집방식으로 수집하였으며, 성의가 없거나 이상점에 해당되는 자료로 확인된 61명을 제외하고 총 409명의 자료를 분석에 포함시켰다. 다음으로 하위요인별 Rasch 평정척도 분석과 확인적 요인분석을 실시하였고, 기존검사들과의 상관분석을 통해 공인타당도를 확인하였다. 연구의 결과를 중심으로 논의를 전개하면 다음과 같다.

첫째, 6개 하위요인별 Rasch 평정척도 분석 결과를 살펴본 결과, 비계의 계획요인 6개 문항 중에서 모둠 활동에 관한 문항이 제거되었는데, 사실 이와 관련된 문항은 비계의 과정요인에 포함되어 있다. 비계의 과정요인 9개 문항은 여러 준거에서 양호한 결과를 보여 그대로 유지되었다. 또한, 비계의 질문요인은 6개 문항에서 1개 문항이 제거 대상 문항으로 선정되었다. 다음으로 모니터링의 향상요인은 8개 문항 중 2개의 문항이 제거 대상 문항이 되어 6개 문항으로 수정되었고, 격려요인과 자기규제요인에서는 각각 1개의 문항이 제거 대상 문항이 되어 5개 문항씩으로 구성되었다.

둘째, 비계와 모니터링을 각각 3요인 모형으로 확인적 요인분석을 실시한 결과, 모니터링은 향상, 격려, 자기규제 3요인 모형이 유지된 반면, 비계의 과정요인과 질문요인이 수렴하지 못

하여 두 요인을 동일한 구인, 즉 과정요인으로 설정하였다. 다음으로 이 연구에서는 비계 2모형과 모니터링 3요인 모형에 대한 확인적 요인분석을 실시하였고, 기여도가 낮은 문항을 제거하기 위해 표준화계수 0.60 기준을 적용하였다. 그 결과, 비계의 계획요인에서는 1개 문항, 비계의 과정요인에서는 8개 문항이 제거 대상 문항으로 선정되었다. 또한, 모니터링의 향상요인에서는 2개 문항, 자기규제요인에서는 1개 문항이 제거 대상 문항으로 판정되었다. 따라서 비계의 계획요인 4개 문항과 과정요인의 6개 문항, 그리고 모니터링의 향상요인 4개 문항, 격려요인 4개 문항, 자기규제요인 4개 문항이 선정되었다.

셋째, 1차 5요인 모형과 2차 2요인 모형을 설정하여 분석하였는데, 최종적으로 1차 5요인 모형이 선정되어 비계의 계획요인 4개 문항, 과정요인 6개 문항, 그리고 모니터링의 향상요인 4개 문항, 격려요인 4개 문항, 자기규제요인 4개 문항이 최종 확정되었다. 여기에서 비계의 과정요인과 질문요인을 하나의 과정요인으로 재설정 한 이유는 질문 행동이 수업 진행 과정에서 주가 되는 활동이고, Pat-El 외(2013)도 비계를 정의하면서 이와 유사한 내용을 주장한 바가 있기 때문이다. 하위요인별 문항 구성과 관련하여 비계의 계획요인은 초등학교 교사들이 제안한 비계의 하위요인으로 Brown, Harris와 Harnett(2012)의 과제 요인을 준비하는 활동, Bransford, Brown와 Cocking(2000)의 과제 단순화 및 목표성취 지침 준비, 그리고 Pat-El 외(2013)의 평가기준 제시와 관련이 있는 것으로 볼 수 있는데, 여기에는 수업설계, 평가방법 계획, 정보 및 피드백 준비 문항이 포함된다. 비계의 과정요인에는 피드백 주고받기, 모둠 활동, 질문 문항이 포함되어 있는데, 이는 Brown, Harris와 Harnett(2012)의 과정 유형, Pat-El 외(2013)의 발표/질문 및 참여 기회 부여 등의 활동과 관련이 있다. 한편, 모니터링의 향상, 격려, 자기규제는 Brown, Harris와 Harnett(2012)의 구성요인과 일치하고, 동료 관련 문항도 이 연구의 자기규제요인에 1문항이 포함되어 있다. 따라서 초등학교 교사의 형성적 피드백 실행 척도의 모니터링 요인은 Pat-El 외(2013), 그리고 Bransford, Brown와 Cocking(2000)의 평가 요소를 모두 포함하고 있는 것으로 해석된다.

넷째, 하위요인별로 문항 합성점수를 측정변수로 설정하여 상관계수를 산출한 결과, 비계의 계획요인과 과정요인 간의 상관이나 모니터링의 향상요인, 격려요인, 자기규제요인 간의 상관이 높았지만, 비계의 과정요인과 모니터링의 향상요인 및 격려요인의 상관계수도 높은 결과를 보였다. 실제로 Brown, Harris와 Harnett(2012)의 연구에서도 과정요인과 향상 간에 매우 높은 상관관계가 있었다. 한편, 박민애, 손원숙(2018)의 학습을 위한 평가척도(단축형) 하위요인과의 상관계수를 보면, 비계의 하위요인과 기존검사의 비계 간 상관이 모니터링과의 상관에 비해 상대적으로 낮았다. 모니터링의 자기규제요인은 박민애, 손원숙(2018)의 모니터링 변수와 높은 상관을 보였지만, 향상과 격려요인은 오히려 선행연구의 비계와 높은 상관을 보였다. Brown(2011)의 교수개선 및 학습향상과의 상관분석 결과, 비계와 모니터링의 하위요인 간에 높은 상관은 초등학교 교사의 형성적 피드백 실행 척도가 형성적 피드백의 기본 요소인

교수개선 및 학습향상과의 공인타당도 분석 결과가 양호한 것으로 해석할 수 있다.

이 연구의 형성적 피드백 실행 척도 개발과 관련된 제언을 결론으로 제시하면 다음과 같다. 첫째, 교사의 형성적 피드백 실행 척도의 요인 세분화가 필요하다. 박민애(2019), 박민애, 손원숙(2016, 2018)과 같이 국외의 형성적 피드백 척도를 번안하여 비계와 모니터링만으로 구분하면 교사들이 형성적 피드백에 대해 임의적 해석을 내릴 가능성이 높다. 따라서 우리나라 실정에 맞는 형성적 피드백의 하위요인을 구분하고 각 하위요인별로 포함시켜야 할 활동 유형이나 활동 내용을 구체적인 문항으로 설정하여 교실에서 실제로 활용할 수 있게 해야 한다. 또한, 교사의 피드백 신념과 피드백 실제 간의 차이를 분석한 Brown, Harris와 Harnett(2012)의 연구에서 주장한 바와 같이, 교사들이 하위요인별로 구체적인 형성적 피드백 실행을 점검함으로써 미흡한 활동에 대한 자기 성찰을 통해 수업 개선을 도모하도록 도움을 줄 필요가 있다.

둘째, 교사의 형성적 피드백은 양질의 교수-학습 및 평가의 핵심이고 형성적 피드백 실행을 통해 학생들이 성공을 경험할 수 있으므로 교사들이 학교현장에서 형성적 피드백을 실천할 수 있는 교육적 환경을 조성해야 한다. 기존의 형성적 피드백 연구와는 달리 이 연구에서 개발한 초등학교 교사의 형성적 피드백 실행 척도는 우리나라 초등학교 현장의 목소리를 담았고, 척도의 타당성 또한 일정 수준에서 확보했다고 볼 수 있다.

셋째, 이 연구결과는 학생들의 과제에 대해 정답 및 오답만을 알려주는 학습결과 위주의 평가 방식에서 벗어나(강대중, 염시창, 2014), 학생들의 학습향상을 위해 학습결과에 대한 계획, 과정, 향상, 격려, 자기규제 등과 관련된 구체적인 형성적 피드백을 제공하는 평가 방식으로 전환을 시도하는데 시사점을 제공할 수 있다. 이를 위해서는 초등학교 교사의 형성적 피드백에 대한 인식(안희상, 손원숙, 2017) 변화뿐만 아니라 피드백 과정에서 학생들이 주인의식을 갖고 참여할 수 있도록 수업과 평가를 재구성해 나가는 노력이 동반되어야 할 것이다.

참 고 문 헌

- 강대중, 염시창(2014). 초등학교 교사들의 형성평가에 대한 인식 및 실태 분석. **학습자중심 교과교육연구**, 14(2), 27-43.
- 김성숙, 김희경, 서민희, 성태제(2015). **교수·학습과 하나되는 형성평가**. 서울: 학지사.
- 노현중, 손원숙(2015). 교사의 숙제 피드백이 학생의 자기조절학습, 과제가치, 학습태도 및 학업성취도에 미치는 영향. **교육평가연구**, 28(3), 879-902.
- 박민애(2019). 학생용 피드백 리터리시 척도(FLSS) 개발 및 타당화. 경북대학교 박사학위논문.
- 박민애, 손원숙(2016). 학습을 위한 평가 척도(AFL-Q)의 타당화. **교육평가연구**, 29(1), 101-121.
- 박민애, 손원숙(2018). Rasch 평정척도 모형을 통한 학습을 위한 평가 척도(AFL-Q)의 단축형 개발 및 타당화. **교육평가연구**, 31(2), 411-434.
- 설현수(2007). Messick의 타당화 관점에서 Rasch 측정모형 적용을 통한 대학 강의평가 도구 개발의 타당화. **교육평가연구**, 20(1), 31-51.
- 손원숙, 박정, 이빛나(2019). 중등 예비교사의 평가인식 잠재프로파일 분석: 피드백 경험과 성취목표와의 관계. **교육평가연구**, 32(2), 375-395.
- 신종호 역(2013). **학습동기: 이론, 연구, 그리고 지원**. 서울: 학지사.
- D. H. Schunk, P. R. Pintrich & J. Meece. (2008). *Motivation in education: Theory, research, and applications*. NJ: Merrill.
- 신현숙, 염시창(2017). 중학생용 학교생활참여척도(K-SES-K) 단축형의 개발 및 타당화. **한국심리학회지: 학교**, 14(2), 149-180.
- 안희상, 손원숙(2017). 초·중등학교의 학교풍토와 교사의 피드백 인식이 피드백 실행에 미치는 영향. **교육평가연구**, 30(3), 445-465.
- 이빛나(2017). 초등교사의 형성적 피드백이 학생의 기본심리욕구와 수업참여에 미치는 영향: 교사-학생 관계의 조절효과. 경북대학교 석사학위논문.
- 지은림(2009). 교사의 피드백 수행을 위한 구성요인 및 특성에 관한 연구. **아시아교육연구**, 10(3), 77-102.
- 지은림(2010). 교사의 형성평가 피드백 수행 척도 개발 및 타당화-Rasch 모형의 적용. **교육평가연구**, 23(1), 79-100.

Arbuckle, J. L. (2013). *Amos(Version 22.0) computer program*. Chicago, IL: IBM-SPSS.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education*, 18, 5-25.

- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86, 8-21.
- Bloom, B. S. (1977). Favorable learning conditions for all. *Teacher*, 95, 22-28.
- Bloom, B. S., Hastings, J. T., Madus, G. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*(3th ed.). New York: Routledge.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*(Rev. ed.). Washington, DC: National Academies Press.
- Brown, G. T. (2011). Self-regulation of assessment beliefs and attitudes: A review of the Students' Conceptions of Assessment inventory. *Educational Psychology*, 31(6), 731-748.
- Brown, G. T., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28, 968-978.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long(Eds.), *Testing structural equation models*(pp. 136-162). Newbury Park, CA: Sage.
- Florez, M. T., & Sammons, P. (2013). *Assessment for learning: Effect and impact*. CfBT Education Trust. University of Oxford.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (6th ed.). Upper Saddle River, NJ: Pearson.
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Linacre, J. M. (2017). *A user's guide to WINSTEPS: Rasch-Model computer program*. www.winsteps.com.
- Martin, A. J., & Dowson, M. (2009). Interpersonal relationships, motivation, engagement, and achievement: Yields for theory, current issues, and educational practice. *Review of Educational Research*, 79, 327-365.
- Meyer, J. P. (2014). *Applied measurement with jMETRIK*. New York: Routledge.
- Meyer, J. P. (2018). *jMETRIK 4.1.1 computer program*. <https://itemanalysis.com>.
- Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning

- ning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31, 199-218.
- Pat-El, R. J., Tillema, H., Segers, M., & Vedder, P. (2013). Validation of assessment for learning questionnaires for teachers and students. *British Journal of Educational Psychology*, 83(1), 98-113.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne & M. Scriven(Eds.), *Perspectives of curriculum evaluation*(pp. 39-83). Chicago, IL: Rand McNally.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189.
- Stiggins, R. (2010). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek(Eds.), *Handbook of formative assessment* (pp. 233-250). New York: Routledge.
- William, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation*, 37, 3-14.
- Wolfe, E. W., & Smith, E. V., Jr. (2007). Instrument development tools and activities for measure validation using Rasch models: Part II - Validation activity. In E. V. Smith, Jr. & R. M. Smith(Eds.), *Rasch measurement: Advanced and specialized applications*(pp. 202-242). Maple Grove, Minnesota: Jam Press.
- Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, 8, 370-371.
- Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education*, 45, 128-136.

· 논문접수 : 2020.01.02. / 수정본접수 : 2020.02.03. / 게재승인 : 2020.02.12.

ABSTRACT

Development and Validation of the Formative Feedback Practice Scale for Elementary School Teachers

Si Chang Yum

Professor, Chonnam National University

Dae Jung Kang

Instructor, Chonnam National University

The purpose of this study was to develop the formative feedback practice scale for elementary school teachers and to test its reliability and validity. Semi-structured questionnaire surveys, interviews and discussions yielded 50 pilot items. These items were administered to elementary school teachers and the data from 409 teachers were analyzed through Rasch rating scale analyses, confirmatory factor analyses, correlation analyses. The Rasch model was applied to each of the 6 formative feedback practice factors(i.e., planning, process, questioning, improvement, encouragement, and self-regulation).

The results from confirmatory factor analyses indicated that the first-order five-factor model showed reasonable fit. Thus, the formative feedback practice scale for elementary school teachers consisted of planning scaffolding(4 items), process scaffolding(6 items), improvement monitoring(4 items), encouragement monitoring(4 items), and self-regulation monitoring(4 items). Internal consistency reliabilities of the sub-scales were high, and positive correlations among the sub-scales were statistically significant. The 5 sub-scales of the formative feedback practice scale for elementary school teachers were positively correlated with the sub-scales of the assessment for learning questionnaire(short-form) and the conceptions of assessment scale. The formative feedback practice scale for elementary school teachers may be useful for measuring and assessing elementary school teachers' feedback practice in the classroom.

Key Words: Formative feedback practice, Elementary teachers, Scale development