

대화 성찰 태도 측정을 위한 설문도구 개발

민병곤 서울대학교 국어교육과 교수/국어교육연구소 겸무연구원(제1저자)

박성석 서울대학교 국어교육과 박사과정(제2저자)

- * 이 연구는 2015년도 서울대학교 사범대학 간접비 지원으로 이루어졌으며, 서울대학교 생명윤리위원회(IRB)의 승인 하에 수행되었음.

- I. 서론
- II. 이론적 배경
- III. 연구 방법
- IV. 연구 결과
- V. 결론

I. 서론

2015 개정 국어과 교육과정에 ‘자기 성찰·계발 역량’이 국어과 핵심 역량 중 하나로 포함된 이후 국어과 교육 전반에서 성찰과 성찰적 교육에 대한 관심이 높아지고 있다. 국어과 하위 영역 중 문학, 작문 교육 등에서는 교육 과정 개정 전부터 이미 성찰적 교육에 관한 논의가 어느 정도 축적되어 왔으며 현 교육과정에 성찰 역량이 포함된 것은 그러한 노력의 결과일 것이다.¹⁾ 상대적으로 다소 늦기는 하였지만 그동안 화법 교육에서도 학습자 스스로 자신의 화법을 돌아보고 성찰하는 일이 중요하다는 인식과 논의가 없지는 않았다.²⁾ 그러나 화법에 대한 성찰을 어떻게 정의할 수 있을지에 관한 기본 논의가

-
- 1) 가령, 문학 교육에서는 최홍원(2006, 2008), 장지혜(2013)에서 성찰적 사고에 대한 체계적인 논의를 바탕으로 시조 읽기나 비평적 에세이 쓰기를 위한 성찰적 교육에 관해 논의한 바 있다. 작문 교육에서는 박영민(2006)에서 작문 평가의 방안으로 학습자가 작문 과정을 돌아보며 이에 관해 글로 쓰는 반성적 쓰기를 활용할 것을 제안하였으며, 박영민·최숙기(2008)에서는 이러한 반성적 쓰기가 중학생의 설명문 쓰기 수행을 향상시킬 수 있음을 실험적으로 검증한 바 있다.
 - 2) 화법 교육에서 초·중등 학습자 대상의 성찰적 교육에 관해 논의한 경우로는 서종훈

전제되지 못한 탓에 논의 결과가 상호 공유되기 어려웠던 것 또한 사실이다.

사전적으로만 보자면 성찰이란 ‘자기의 마음을 반성하고 살핌’으로 간단히 정의될 수 있다. 그러나 지금까지 국내·외의 다양한 학문 분야에서 이루어져 온 논의들을 살펴보면 성찰이라는 개념은 간단히 정의되기 어려운 매우 추상적이고 포괄적인 의미를 담고 있음을 확인할 수 있다.³⁾ 따라서 성찰적 화법 교육에 관해 논의하고자 한다면 본격적인 연구 진행에 앞서 성찰의 개념을 범학문적 기준에서 타당하게 정의할 수 있어야 할 것이다. 이에 본 연구에서는 문헌 검토를 통해 화법 교육에서 다루는 담화 유형 중 가장 기본적인 ‘대화’를 성찰한다는 것이 어떻게 정의되어야 하는지를 논의하고, 이를 바탕으로 대화에 대한 학습자의 성찰적 심리를 타당하고 신뢰할 수 있게 측정할 수 있는 자기 보고식 설문 문항을 개발하고자 한다.

국외 연구들의 논의에 따르면 성찰이란 본질적으로 사고 과정에 해당하며 이는 대화 성찰 또한 마찬가지일 것이다. 그러나 본 연구에서는 사고 과정으로서의 대화 성찰 그 자체가 아닌 대화 성찰을 행하는 경향성 즉, 대화 성찰 ‘태도’의⁴⁾ 측정 문항 개발을 목표로 삼았다. 즉, 본 연구는 성찰적 사고 과정이 제대로 이루어지고 있는지를 측정하는 능력검사 개발을 목표하는 것

(2009), 김정란(2014), 박창균(2015)을 확인할 수 있다.

- 3) 이와 관련하여 Rogers(2001)에서는 교육학적 관심에서 ‘성찰’을 논의해 온 주요 문헌들을 검토하고 ‘성찰’의 개념에 관해 메타적 분석을 제시하고 있어 참고할 만하다. 특히, 검토 대상이 된 8편의 주요 문헌들이 비록 용어를 달리 사용하고 있기는 하나(예: ‘성찰(reflection)’, ‘성찰적 사고(reflective thought)’, ‘관리적 성찰(managerial reflection)’, ‘마음챙김(mindfulness)’), 공통적으로 ‘성찰’이 (혹은 이에 상응하는 다른 용어가) 인지적 과정 혹은 활동을 가리키는 것으로 정의하고 있다고 밝히고 있다(pp. 38-42). 본 연구에서도 ‘성찰’이 기본적으로 인지를 중심으로 한 사고 과정이라고 간주하며, 따라서 본 논의에서는 특별한 언급이 없는 한 ‘성찰’과 ‘성찰적 사고’라는 용어를 동의어처럼 사용하고자 한다.
- 4) 본 연구는 Allport(1935: 810)의 정의에 따라 태도(attitude)를 ‘경험을 통해 조직되며, 관계된 모든 대상 및 상황에 직접적이거나 역동적 영향력을 행사하는, 정신적·신경적 준비상태’라고 간주한다.

이 아니라, 평상시 성찰적 사고를 얼마나 반복적으로 행하는지, 혹은 이를 행 하려는 정신적 준비상태의 정도가 얼마나 높은지를 측정하는 ‘성격검사’의⁵⁾ 개발을 목표로 한다는 것이다. 본 연구에서 대화 성찰이 아닌 대화 성찰 태도의 측정을 목표로 삼은 이유는 다음과 같은 두 가지이다.

첫째, 범학문적으로 성찰적 사고의 과정 및 내용에 관한 다양한 이론적·실증적 논의가 이루어져 온 것이 사실이나, 화법에 초점을 둔 성찰적 사고가 어떻게 이루어지는지에 관해서는 현재까지 논의되거나 밝혀진 바가 많지 않다. 특히, 본 연구에서 관심을 갖는 일상 대화에서의 성찰이 어떻게 이루어지는가에 관한 기존 논의는 국내·외를 막론하고 찾아보기가 어렵다. 따라서 현 시점에서 대화에 대한 성찰적 사고와 이에 기반을 둔 대화상의 문제 해결 능력을 측정할 수 있는 도구의 개발은 시기상조일 것이다. 그러나 대화 성찰 태도의 경우 외적으로 관찰 가능하거나 직관적으로 이해 가능한 행동 경향성에 바탕을 두어 개념을 정의하고 접근하는 것이 가능하므로, 비록 대화 성찰에 관한 논의에 의지할 수는 없다 하더라도 성찰 일반에 관한 타 학문의 논의에 근거하여 충분히 타당하고 신뢰할 수 있는 측정 문항의 개발이 가능하다고 보았다. 그리고 이처럼 대화 성찰 태도의 측정이 가능하게 되면 이를 통해 성찰적 태도가 높거나 낮은 학습자를 구별하는 일이 가능해 질 것이고, 이를 근거로 대화 성찰의 사고 과정 그 자체를 규명하기 위한 논리적 접근이 가능하게 될 것이며,⁶⁾ 이러한 연구 결과가 축적된 연후에 대화 성찰 능력을 측정하려는 시도도 가능해질 것이라 판단하였다.

둘째, 본 연구가 선택한 자기 보고식 측정 방식은 대량의 자료 수집을 위한 시간과 비용이 비교적 적게 드는 장점이 있으나, 전통적으로 능력검사

-
- 5) 박경, 김혜은(2017: 18)에 따르면 ‘성격검사’란 주로 정서, 대인관계, 동기, 흥미, 태도 등 의 특성을 측정하는 검사를 가리킨다.
 - 6) 가령, 연구 참여자의 대화 성찰 내용을 드러내는 구두 인터뷰나 작문 활동을 매개로 자료를 수집한 후 대화 성찰 태도가 높은 참여자와 낮은 참여자의 성찰 내용을 비교·대조하는 접근방식을 예로 들 수 있다.

보다는 태도와 같은 개인의 습관적 특성 혹은 성격을 검사하기에 적합하다고 알려져 있다.⁷⁾ 앞서 밝힌 바와 같이 대화 성찰 능력 검사의 측정 시도가 현재로서는 시기상조이기도 하지만 향후 대화 성찰 능력의 측정 시도가 가능해지더라도 이는 성취도 검사에서와 같이 정답이 정해진 여러 문항을 통해 성찰적 문제 해결 능력을 평가하는 방식이 되거나, 채점자가 채점기준표에 입각해 피검자가 성찰한 내용이 담긴 글을 평가하는 방식이 될 것이다.⁸⁾ 그런데 이러한 방식으로 대량의 자료를 수집하려면 상당한 비용과 시간이 소요될 수밖에 없는데 비해, 본 연구가 선택한 자기 보고식 측정 방식은 비교적 적은 비용과 시간을 들이고도 대량의 응답을 수집할 수 있어 개인 연구자가 손쉽게 연구 수행에 활용할 수 있는 이점이 있다. 따라서 성찰적 사고 능력을 정밀하게 측정하는 것보다는 성찰 경향성이 높은 이와 낮은 이를 구별하거나 성찰적 교육 전후에 학습자의 성찰 경향성이 달라졌는지를 객관적으로 점검하는 것만을 목표로 연구를 설계하고자 한다면, 오히려 본 연구가 개발하고자 하는 자기 보고식 측정 도구가 보다 효율적이고 실용적일 수 있을 것이다.

본 연구는 이상과 같은 배경 및 문제의식을 토대로 대화 성찰 태도를 측

-
- 7) 이와 관련하여 박경, 김혜은(2017: 16-18)에서는 심리검사도구를 수행양식에 따라 크게 극대수행검사와 습관적 수행검사로 나누고, 극대수행검사의 경우 응답자가 일정한 시간 내에 자신의 능력을 최대한 발휘하여 반응하도록 제작되는 검사로서 지능검사, 적성검사, 성취도검사 등이 이에 해당한다는 점, 습관적 수행검사의 경우 개인이 평소에 어떠한 행동을 얼마나 보이는지 측정하는 검사로서 성격검사와 흥미검사가 이에 해당한다는 점을 밝히고 있다. 이들은 또한 습관적 수행검사의 하나인 성격검사는 주로 여러 문항으로 구성된 질문지를 읽고 각각의 문항에 대해 응답자 스스로가 반응한 결과를 통해 측정하는 방식을 사용한다고 밝히고 있다.
- 8) 연구자들의 검토 결과, 성취도 검사와 유사한 형태로 성찰 능력을 측정할 수 있는 도구는 확인되지 않았으나, 성찰 내용을 담은 글을 평가하기 위한 채점기준표나 코딩 도식을 개발한 연구들은 다수 확인할 수 있었다(예: Bain, Ballantyne, Packer, & Mills, 1999; Kember, MaKay, Sinclair, & Wong, 2008; Lau, 2015; Wald, Borkan, Taylor, Anthony, & Reis, 2012).

정할 수 있는 타당하고 신뢰 가능한 자기 보고식 설문 도구 개발을 시도하였다. 이러한 ‘대화 성찰 태도 측정을 위한 설문도구’의 개발 결과를 제시하고자 하는 본고에서는 체계적인 보고를 위해 다음과 같은 두 가지 연구문제를 중심으로 논의를 진행하고자 한다.

첫째, 대화 성찰 태도의 3요인 모형은 구인 타당성을 지니는가?

둘째, 대화 성찰 태도의 측정 문항들은 신뢰할 만한가?

II. 이론적 배경

1. 성찰의 개념 및 분류

성찰의 개념 규정에 대한 시도는 일찍이 Dewey(1910/2011)에서부터 시작되었다. Dewey는 성찰적 사고(reflective thought)를 ‘어떤 믿음이나 소위 지식의 형태에 대해 지지하는 근거와 가져올 결과에 비춰 적극적이고 지속적이며 세심하게 숙고하는 것’이라고 정의한다(p. 9). 그는 성찰적 사고가 ‘당혹, 망설임, 의심의 상태’를 불러일으키는 어떤 문제를⁹⁾ 인식하는 데서 출발하여 ‘문제 해결책으로서 가정된 믿음을 지지하거나 부정하는’ 사실 혹은 증거들을 찾는 과정 속에서 발현된다고 논의한다(p. 13). 그는 ‘만일 생겨난 가정이 즉시 받아들여진다면, 우리는 성찰이 거의 없는 무비판적인 생각을 하게 된다’라고 언급하며(p. 17), 좋은 사고와 이에 따른 최선의 가정 생

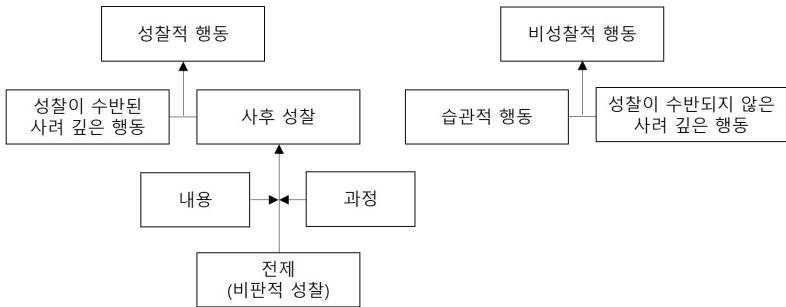
9) Dewey는 자신이 ‘문제(problem)’라 가리키는 것에 대해, ‘성격상 아무리 사소하고 일상적인 것이라 할지라도 마음을 당혹하게 하고 마음에 도전해서 믿음을 조금이라도 불확실하게 만드는 모든 것’이라고(p. 13) 밝히고 있다. 따라서 Dewey가 말하는 문제 해결적 사고로서의 ‘성찰적 사고’는 인간이 자신이 품게 된 의문을 해결하기 위해 의식적으로 행하는 모든 사고 과정과 관련되는 매우 일반적인 용어라 할 수 있다.

성 및 문제 해결을 위해서는 잠시 결론 내리기를 보류하고 의심의 상태를 유지하며 계속해서 증거를 찾고자 하는 성찰적 사고가 수반되어야 함을 역설하였다.

이와 같은 Dewey의 성찰 개념이 탈맥락적이고 일반적인 의미에서의 문제 해결을 위한 사고 과정으로서 규정되었다면, Schön(1983)에서 제안하는 성찰은 무질서하고 이론화되기 어려운, 맥락 속 ‘실천(practice)’ 문제의 해결을 위한 사고라는 보다 구체적인 의미를 담고 있다. Schön은 과학적으로 일반화된 이론 및 기술을 엄격히 적용함으로써 가장 효율적이고 효과적인 실천이 가능하다고 여기는 소위 ‘기술적 합리성(technical rationality)’의 관점이, 실천 맥락의 무질서함과 역동성은 물론, 인간이 수많은 경험을 축적하는 과정에서 계발하게 되나 일반화되기는 어려운 직관적 문제 해결 능력을 무시하고 있음을 지적하며, 무질서하고 역동적인 맥락 속에서 이루어지는 실천이 최선의 방식으로 이루어지려면 단순히 주어진 이론이나 기술을 이해하고 적용하는 것만으로는 부족하고 스스로 자신이 처한 상황과 맥락에 맞는 문제 해결 방법을 찾고자 사고하는 성찰이 필요하다고 보고 있다.

나아가 Schön(1983)에서는 성찰이 이루어지는 시기를 기준으로, ‘행위 중 성찰(reflection in action)’과 ‘행위 후 성찰(reflection on action)’을 구별하고 있다. Schön은 실천 상황에서 한 번도 경험해 보지 못한 당혹스러운 문제에 봉착했을 때, 전문가나 초보자 모두 자신이 취한 문제 해결 방식이 적절했는지를 실천이 종료된 후 돌아보는 ‘행위 후 성찰’은 가능하지만, 전문가일수록 초보자와 달리 실천 상황 중 즉각적으로 최선의 해결 방안을 찾아내는 ‘행위 중 성찰’ 또한 가능하다고 밝히고 있다. 이상과 같은 Schön의 논의는 성찰적 사고를 실천 문제의 해결을 목표로 하는 사고로 한정하여 그 의미를 명확히 규정하고자 하였을 뿐 아니라, 성찰이 실천이 종료되고 난 후 시간을 들여 이루어지는 것만이 아니라 실천의 과정 중에도 나타날 수 있음을 부각하였다는 점에서 의의를 지닌다.

한편, Mezirow(1990, 1991)에서는 성찰 개념에 관한 앞선 논의들을 바



〈그림 1〉 성찰의 의미역(Mezirow, 1990: 7)

탕으로 성찰의 의미역을 체계적으로 분류하여 제시하고 있다(〈그림 1〉 참고). Mezirow는 우선 ‘비성찰적 행동’과 ‘성찰적 행동’을 나눈 후, ‘비성찰적 행동’에는 자신이 지닌 자동화된 기대나 행동 방식에 따라 별다른 생각 없이 이루어지는 ‘습관적 행동(habitual action)’과, 자신의 행동 파악을 위해 잠시 멈추거나 자신의 생각과 느낌에 주목하기는 하나 그러한 일이 성찰과 함께 이루어지는 것은 아닌 ‘성찰이 수반되지 않은 사려 깊은 행동(thoughtful action without reflection)’이 존재한다고 보았다. 한편, ‘성찰적 행동’에서는 행동이 진행되는 중에 성찰이 이루어지는 ‘성찰이 수반된 사려 깊은 행동(thoughtful action with reflection)’과, 행동이 종료된 이후에 성찰이 이루어지는 ‘사후 성찰(ex post facto reflection)’을 구분하였다. ‘성찰이 수반된 사려 깊은 행동’의 일부로서 이루어지는 성찰에는 ‘내용 성찰(content reflection)’이나 ‘과정 성찰(process reflection)’ 혹은 둘 모두가 포함된다. 그리고 사후 성찰의 경우 ‘내용 성찰’이나 ‘과정 성찰’뿐만 아니라 그러한 기본 성찰들의 밑바탕에서 이러한 과정에 영향을 미치는 근본 가정에 대한 성찰인 ‘전제 성찰(premise reflection)’까지 포함될 수 있다고 밝힌다.

이상과 같은 Mezirow(1990, 1991)에서의 논의는 성찰과 비성찰을 구별함은 물론, 성찰을 시기(성찰이 수반된 사려 깊은 행동, 사후 성찰) 및 초점(내용 성찰, 과정 성찰, 전제 성찰)에 따라 상세히 분류하여 규정하고 있다는

〈표 1〉 기존 연구에서의 성찰 분류 체계

구분	출처	성찰의 분류
3분법	Jay & Johnson(2002)	묘사적 성찰, 비교적 성찰, 비판적 성찰
	Kreber & Castleden(2009), McAlpine 외(2004)	내용 성찰, 과정 성찰, 전제 성찰
4분법	Hatton & Smith(1995)	묘사적 작문, 묘사적 성찰, 대화적 성찰, 비판적 성찰
	Kember 외(2000), Kember 외(2008)	습관적 행동(비성찰), 이해, 성찰, 비판적 성찰
	Larrivee(2008)	미성찰, 피상적 성찰, 교수법적 성찰, 비판적 성찰

점에서 의의를 지닌다. 또한 기존 논의들에서 성찰의 개념을 주로 문제 해결을 위한 사고로만 한정하여 논의한 것과 달리, 애초의 문제 제기에 영향을 미치는 주체의 믿음, 가정, 가치관 등 소위 ‘전제’에 대한 ‘비판적 성찰(전제 성찰)’로까지 성찰의 의미역을 확장함으로써 이후 연구에서 성찰을 수준에 따라 분류할 수 있게 하는 이론적 토대를 제공하고 있다.

이상과 같은 Dewey(1910/2011), Schön(1983), Mezirow(1990, 1991)에서의 논의에 바탕을 두어 다양한 학문 분야의 여러 학자들이 성찰을 이론적 혹은 실증적으로 분류하기 위한 체계를 제안해 왔다(〈표 1〉). 그런데 이들 중 상당수는 성찰의 분류 근거를 마련하기 위한 자료를 성찰 일지와 같은 글쓰기 결과물 형태로 얻거나(예: Hatton & Smith, 1995; Jay & Johnson, 2002), 성찰의 초점 영역을 예비 교사 교육 등 특정 전공 분야로 한정하여(예: Kember 외, 2000; Kember 외, 2008; Larrivee, 2008) 분류를 시도한 탓에, 성찰적 사고 본연의 특징과는 다소 거리가 멀어 보이는 글쓰기 상의 특징이 반영된 범주(예: 묘사적 성찰)를 제안하거나, 전공 분야의 특수성이 반영된 범주(예: 교수법적 성찰)를 제안하고 있기도 하다. 그러나 비록 이처럼 분류 범주에서 다소 차이가 있다고는 하나, 이들 모두는 공통적으로 Mezirow(1990, 1991)에서 제안한 ‘전제 성찰’(혹은, ‘비판적 성찰’)이 자신들의 성찰 분류 범주 가운데서도 최상위 수준을 표상한다고 간주하는 등의 공통점도 지니고 있다.

성찰의 분류 방식을 제안한 이상의 연구 중 특히, Kember 외(2000)에서는 본 연구가 목표하는 것과 같은 자기 보고식 심리 측정 문항의 개발 결과를 보여주고 있어 참고할 만하였다.¹⁰⁾ 이들은 전문가 양성을 목표로 하는 대학 강좌를 통해 학습자가 성찰적 사고를 어느 정도 수행할 수 있었는지를 학습자의 자기 보고를 통해 측정하는 방식을 선택하고 있다. 그리고 이들이 최종적으로 제안하는 성찰의 구인은 습관적 행동, 이해, 성찰¹¹⁾, 비판적 성찰의 네 가지였다. 이러한 네 가지의 구인은 대체로 Mezirow가 논의한 성찰의 의미역을 따라 도출된 것들이나(p. 383), 좀 더 면밀히 살펴보게 되면 이들의 구인은 Mezirow의 논의와 적어도 두 가지 사항에 있어 차이가 있는 것으로 보인다.

첫째는 ‘습관적 행동’을 성찰의 하위 구인으로 설정한 점이다. 앞서 제시한 바와 같이 Mezirow는 습관적 행동을 비성찰적인 것으로 간주하고 있으나 Kember 외(2000)에서는 대학 강좌 중 교육 목적으로 반복적 활동을 행함으로써 그러한 활동의 수행이 자동화된 상태를 성찰의 결과라고 보고 이를 측정하기 위한 구인으로 ‘습관적 행동’을 설정하고 있다. 둘째는 Mezirow 가 언급하지 않은 ‘이해(understanding)’를 성찰의 하위 구인으로 설정한 점이다. 이들은 Mezirow가 제안한 ‘사려 깊은 행동’을 재해석하여 ‘이해’ 구인을 설정하였다고 밝히고 있으며(p. 384) 실제로 Mezirow(1991: 107)에서는

10) Larrivee(2008) 또한 자기 보고식 측정 문항 개발을 목표로 하고 있기는 하나, 개발한 문항에 대한 전문가 내용 타당도 점검 결과만을 보고하고 있을 뿐 본격적인 설문조사와 구인 타당도 검증까지는 보고하고 있지는 않다.

11) Kember 외(2000)는 ‘성찰’의 하위 구인으로 다시 ‘성찰’을 설정하고 있어 그 용어법이 다소 혼란스러울 수 있다. 연구자들 스스로가 명시적으로 밝히고 있지는 않으나 하위 구인으로서의 ‘성찰’은 Mezirow의 용어법으로는 내용 성찰과 과정 성찰을 통합한 구인으로서, Dewey, Schön 등이 논의해 온 전통적이고 좁은 의미에서의 성찰 즉, 문제 해결에 초점을 둔 사고를 가리키는 것으로 보인다. 한편, 이를 포괄하는 상위어로서의 ‘성찰’은 문제 제기와 관련된 전체 성찰(비판적 성찰)까지를 포괄하는 넓은 의미의 성찰을 가리키는 것으로 보인다.

사려 깊은 행동이 선행 학습에 의지하는 경우도 있다고 밝히고 있다. 그러나 이러한 ‘이해’ 구인은 추상화된 전문 개념의 학습이 반드시 전제되는 대학 강좌에서 본격적 성찰 수행에 이러한 전문 개념의 이해 정도가 얼마나 관여 하였는가를 측정하기 위한 구인일 뿐 사려 깊은 행동 그 자체를 반영하는 구인은 아니라고 보인다.

이상과 같은 선행 연구 검토 결과, 본 연구는 구인 타당도 검증에 초점을 둔 본 연구의 전반적 설계 및 검증 방식에서만큼은 Kember 외(2000)에서처럼 자기 보고식 심리 측정 도구 개발의 일반 절차를 따르되, 구인 설정 방식에 있어서는 Kember 외(2000)의 구인을 그대로 따르는 대신 이들이 기반으로 삼은 Mezirow(1990, 1991)의 논의에 직접 근거하여 대화 성찰에 적합한 구인을 새로 설정하고자 하였다. 이는 일상적인 삶에서 이루어져야 하는 대화 성찰의 맥락이 전문가 양성을 목표로 하는 대학 교육의 맥락과는 다를 수 있으므로 대화 성찰의 구인에 관해서는 원점에서부터 다시 숙고할 필요하다고 보았기 때문이다.

2. 대화 성찰 태도의 정의 및 구인

본 연구에서는 이상의 선행 논의에 대한 검토를 바탕으로 일상 대화에서의 실천에 관여하는 성찰적 태도 즉, ‘대화 성찰 태도’와 그 하위 구인의 개념을 정의하고 이를 측정하기 위한 문항을 개발하고자 한다. 먼저 대화 성찰 태도를 정의하기 위해, Dewey(1910/2011)의 문제 해결을 위한 사고로서의 성찰, Schön(1983)의 실천 문제 해결을 위한 행위 중 성찰과 행위 후 성찰, Mezirow(1990, 1991)의 비판적 성찰(전체 성찰) 개념을 종합적으로 검토하여 대화 성찰 태도를 ‘대화가 진행 중이거나 대화가 종료된 이후 자신의 발화 내용이나 대화 방식은 물론 이에 영향을 미치는 스스로의 대화에 관한 믿음을 신중하고 사려 깊게 점검하고 살펴보고자 하는 태도’라고 정의하였다.

다음으로 대화 성찰 태도의 하위 구인을 설정하기 위해, 성찰의 범주 혹

은 구인 분류를 시도한 여러 연구들이(예: Kember 외, 2000, 2008; Kreber & Casteleden, 2009; McAlpine 외, 2004) 이론적 기반으로 삼고 있는 Mezirow(1990, 1991)를 주요 근거로 삼아 대화 성찰 태도의 하위 구인을 ‘사려 깊은 행동’, ‘내용·과정 성찰’, ‘전제 성찰’의 세 가지로 설정하였다. Kember 외(2000)에서처럼 습관적 행동 또한 하위 구인으로 포함시킬 것인지에 대한 검토가 있었으나, 성찰적 교육의 효과를 측정하고자 했던 이들의 연구와 달리, 본 연구는 교육적 개입과 상관없이 피조사자의 대화 성찰 태도가 어느 정도인지를 측정하는 데 관심을 두고 있으므로, 습관적 행동을 대화 성찰 태도의 구인으로 포함시키는 것은 불필요하고 논리적으로 맞지 않다고 판단하였다.¹²⁾

첫 번째 구인인 ‘사려 깊은 행동’과 관련하여, Mezirow는 내가 무엇을 하고 있는지를 물음으로서 잠시 멈추는 것이 이에 해당하며(Mezirow, 1990: 6), 자신의 생각이나 느낌에 관하여 사고함을 가리키는 내성(introspection)이 여기에 포함된다고 밝히고 있다(Mezirow, 1991: 107). 그리고 이러한 사려 깊은 행동이 비록 그 자체만 분리하여 생각한다면 이미 학습된 것에 의지하여 진행 중인 행동에 그저 주목하는 것에 불과하다는 점에서 본격적인 성찰로 보이지 않을 수 있지만, 체스를 두거나 논쟁을 벌이거나 재치 있는 행동을 하는 등에서와 같이(Mezirow, 1990: 6), 일련의 행동에서 다음 행동을 어떻게 취할지 심사숙고하거나 새로운 경험을 어떻게 이해할지 고민하는 등 의 문제 해결 과정에서 성찰과 사려 깊은 행동이 통합될 수 있다고 밝히고

12) Mezirow(1991: 106)에서는 ‘습관적 행동’이 타자 치기, 자전거 타기, 자동차 운전하기와 같이 자신이 행하는 것에 대한 자각이 없이 일어나므로 다른 것에 관심을 두더라도 자유롭게 행할 수 있는 비성찰적 행동이라고 밝히고 있다. Kember 외(2000)는 성찰적 교육의 효과를 측정하고자 하는 목적에서 도구를 개발하였으며, 따라서 강좌에서 이루어진 성찰적 학습의 결과로서 나타나는 습관적 행동에 관심을 갖고 이를 측정 구인으로 포함하고 있었으나, 본 연구에서는 교육 여부와 상관없이 대화 성찰 태도를 측정하는 것을 목표로 하므로 측정 시기를 기준으로 이미 습관적 행동으로서 자동화된 상태는 더 이상 성찰과 상관이 없다고 보아 Mezirow의 논의에서 규정한 바와 같이 이를 비성찰로 간주하였다.

있다(Mezirow, 1991: 107). Mezirow(1990)에서 ‘성찰이 수반된 사려 깊은 행동’과 ‘성찰이 수반되지 않은 사려 깊은 행동’을 따로 구별한 것은 바로 이러한 까닭이다. 따라서 연구자들은 여럿이 함께 결합하여 하나의 포괄적 심리 현상을 구성하는 ‘구인(construct)’의 특성을 고려할 때, ‘사려 깊은 행동’ 또한 대화 성찰 태도를 구성하는 하나의 하위 구인이 될 수 있다고 보았다. 한편, Mezirow(1990: 7)가 제시한 성찰의 의미역에서도 확인할 수 있듯이(<그림 1> 참고), 이러한 사려 깊은 행동은 실천 종료 후가 아닌 실천이 이루어지는 도중의 성찰 즉, 행위 중 성찰과 통합되어 발현된다. 따라서 본 연구에서는 ‘사려 깊은 행동’을 ‘대화중 어떤 문제를 느꼈을 때 즉각적으로 반응하기보다는 잠시 숙고한 뒤 신중하게 반응을 보이고자 하며, 특히 감정을 둘러싼 문제에 세심하게 대응하려는 대화 성찰 태도’라고 정의하였다.

두 번째 구인인 ‘내용·과정 성찰’은 Mezirow(1990, 1991)에서 내용 성찰과 과정 성찰의 둘로 구별한 것을 하나로 통합한 구인이다. Mezirow가 내용과 과정을 구별하여 언급하고 있는 것은 사실이지만, 이 둘 각각의 정의를 제시하거나 이 둘의 차이를 명징하게 드러내고 있지는 않다. 다만 문제 해결을 위해 자신이 행한 실천이 적절했는지 살피거나 미래의 유사 문제 상황에서 자신이 생각하는 실천 방식이 적절할지를 살피는 과정에서, 내용 성찰은 문제의 ‘내용’(즉, ‘무엇’)에, 과정 성찰은 문제 해결의 ‘과정’(즉, ‘어떻게’)에 초점을 두는 것이라고 말할 뿐이다(Mezirow, 1990: 6; Mezirow, 1991: 107-108). 문제의 내용과 문제 해결의 과정을 숙고하는 일은 사실상 상당한 유기성을 지녀 그 경계를 나누기가 어려울 수 있다. Kember 외(2000)에서 내용 성찰과 과정 성찰을 별도의 구인으로 설정하는 대신 이를 통합하여 단일 구인인 ‘성찰(reflection)’로 제시한 것도 이러한 이유 때문으로 보인다.¹³⁾ 이

13) Kember 외(2000: 383)에서는 내용 성찰과 과정 성찰을 통합한 까닭을, 그저 ‘보다 유용한 도구를 만들기 위해(to produce a more usable instrument)’라고만 간단히 언급하고 있다. 본 연구의 경우 사전에 내용 성찰과 과정 성찰을 구별하여 예비 문항을 작성하기도 하였으나 전문가들의 내용 타당도 점검 과정에서 서로 대등하고 밀접히 관련된 것으로

에 본 연구에서도 내용 성찰과 과정 성찰을 구별하는 대신 이 둘을 통합하여 ‘내용·과정 성찰’이라는 하나의 하위 구인을 설정하고자 하였다. 대화에서의 문제가 주로 실천을 둘러싼 문제라고 볼 때, 이러한 ‘내용·과정 성찰’은 Schön(1983)에서 제시한 실천 문제에 대한 행위 중 성찰(본 연구에서는 ‘대화 중 성찰’)과 행위 후 성찰(본 연구에서는 ‘대화 후 성찰’)을 포괄적으로 지칭하는 것으로 볼 수 있다. 실제로 Mezirow는 내용 성찰과 과정 성찰이 ‘성찰이 수반된 사려 깊은 행동’과 ‘사후 성찰’ 모두에서 가능하다고 밝히고 있다(Mezirow, 1991: 108). 이에 본 연구에서는 ‘내용·과정 성찰’을 ‘현재 진행 중이거나 이전에 종료된 대화에서 최선의 대화 방법이 무엇일지를 숙고하려 하며, 주로 감정을 중심으로 자신의 발화 내용을 검토하려는 대화 성찰 태도’라고 정의하였다.

끝으로 세 번째 구인인 ‘전제 성찰’의 경우, 문제의 내용이나 문제 해결의 방법을 성찰하기 이전에 애초 자신이 무언가를 문제라고 인식하하게 한 이유가 되는(즉, ‘왜’) 자신의 근본적인 믿음, 가정, 가치관을 성찰하는 것을 가리킨다. 이와 관련하여 Mezirow는 우리는 누구나 스스로의 의미 관점에 붙들려 있기 때문에 결코 편견으로부터 벗어나 스스로의 경험을 해석할 수 없으며(Mezirow, 1990: 6), 따라서 성찰의 과정에서 최초에 문제를 제기하거나 문제라고 규정하는 바탕이 된 전제의 정당성에 대한 의문 제기인 전제 성찰이 필요하다고 밝힌다(Mezirow, 1990: 12). 이러한 전제 성찰에는 이미 형성되고 습관적인 우리의 기대 패턴들, 세계·타인·자신과의 만남에서 형성해 온 의미 관점들에 의문을 제기하는 것이 포함되며(Mezirow, 1990: 12), 이는 성찰을 통한 문제 해결의 방향이 올바르게 설정되도록 하는데 반드시 필요한 성찰이라 할 것이다. 또한 Mezirow는 전제 성찰이 즉각적 행동 과정

보이는 이 둘을 구별하는 것이 대체로 성찰의 수준 차와 관계된 것으로 보이는 구인의 전체적 구조에 비추어 볼 때 어색하다는 지적을 받기도 하였고, 실제로 실증 국면에 들어가서도 이 둘이 분리되기 어렵다는 점을 확인하게 되어 이 둘을 통합하기로 결정하였다.

속에 통합된 요소가 될 수 없다고 밝히고 있는데(Mezirow, 1990: 13), 이는 전제 성찰이 비교적 높은 수준의 숙고를 필요로 하는 까닭에 즉각적으로 행동해야만 하는 상황에서는 발현되기 어렵기 때문인 것으로 보인다. 이와 같은 논의를 근거로 본 연구에서는 ‘전제 성찰’을 ‘자신의 대화 성찰에 영향을 미치는 스스로의 기대 습관이나 가정이 바람직한지, 그것이 주변으로부터 무비판적으로 동화된 것은 아닌지를 비판적으로 점검하려는 대화 성찰 태도’라고 정의하였다.

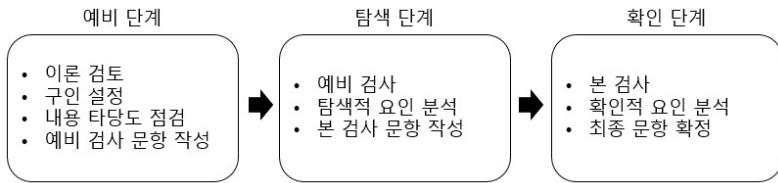
III. 연구 방법

1. 절차

본 연구는 심리 측정 도구 개발의 일반 절차에 따라 <그림 2>와 같이 세 단계로 나뉘어 진행되었다.

먼저 예비 단계에서는 대화 성찰 태도를 측정하기 위한 문항 개발을 위해 성찰에 관한 기존의 이론적 논의(예: Dewey, 1910/2011; Mezirow, 1990, 1991; Rogers, 2001; Sschön, 1983)와 자기 보고식 성찰 측정 도구 개발 연구(Larrivee, 2008; Kember 외, 2000)들을 검토하고 이론적 근거에 부합하도록 대화 성찰 태도를 정의하는 한편 그 하위 구인들을 설정하였다(II장 참고). 그리고 대화 성찰 태도 및 그 하위 구인의 조작적 정의에 부합하도록 예비 문항을 작성한 후 박사학위를 소지하고 있는 화법교육 연구자 5인으로부터 내용 타당도를 점검받았다.

최초에는 대화 성찰 태도의 하위 구인을 ‘습관적 행동’, ‘사려 깊은 행동’, ‘과정 성찰’, ‘내용 성찰’, ‘전제 성찰’로 설정하여 예비 문항을 작성하였으나 전문가들로부터 구인 설정과 이에 따른 하위 문항 작성 방식에서 두 가



〈그림 2〉 본 연구의 절차

지 문제점을 지적받았다. 첫째는 ‘습관적 행동’과 ‘사려 깊은 행동’이 개념적으로는 구별될 수 있을지 몰라도, 이를 구인을 측정하기 위해 작성된 예비문항들이 서로 역(逆)의 관계에 불과한 것으로 보이며, 성찰의 구인에 습관적 행동을 포함시키는 것이 타당치 않아 보인다는 점이었다. 둘째는 구인들이 전반적으로 성찰의 수준 차를 표상하는 것으로 보이는 가운데, ‘과정 성찰’과 ‘내용 성찰’의 두 개 구인만은 서로 다른 수준 차와 관련된다기보다는 동일 수준에서 초점의 차이만을 표상하는 것으로 보인다는 점이었다. 이에 연구자들은 ‘습관적 행동’의 측정을 위해 작성한 예비문항을 ‘사려 깊은 행동’의 역변별 문항으로 삼아 두 개 구인을 ‘사려 깊은 행동’의 단일 구인으로 통합하고, ‘과정 성찰’과 ‘내용 성찰’의 구인을 구별하지 않고 ‘내용·과정 성찰’이라는 단일 구인으로 통합하였다.¹⁴⁾ 이에 따라 본 연구에서는 ‘사려 깊은 행동’, ‘내용·과정 성찰’, ‘전체 성찰’의 세 개 구인을 최종적인 가설로 설정하고 27개의 예비 검사 문항을 마련하였다.

다음으로 탐색 단계에서는 2016년 10월에 서울 지역에 소재한 특목고 1곳, 특성화고 1곳, 일반고 2곳을 포함한 총 4개 고등학교에서 1, 2학년 학생 467명을 대상으로 예비 검사를 실시하였다. 예비 검사의 주된 목적은 예비 검사 문항들이 개략적 수준에서나마 본 연구에서 가정한 3요인 구조를 형성

14) 특히, ‘내용·과정 성찰’은 이후 탐색적 요인 분석 단계에서도 두 개의 독립된 구인으로 나뉘기 어렵다는 점이 확인되어 전문가 검토 의견을 반영한 선택이 옳았음을 확인할 수 있었다.

하고 있는지를 살피는 한편, 3요인 구조를 가정할 수 있다고 할 때 신뢰도와 타당도를 갖춘 문항과 그렇지 않은 문항은 어떤 것인지를 가려내는 데 있다. 예비 검사의 응답 결과에 대해 탐색적 요인 분석(exploratory factor analysis)을 실시한 결과 3요인 가정을 유지할 수 있음을 확인할 수 있었던 한편, 이러한 3요인 가정에 적합한 문항으로 27개의 예비 문항 중 12개를 선별하여 본 검사 문항으로 채택할 수 있었다.

끝으로 확인 단계에서는 2016년 12월에 서울 지역에 소재한 특목고 2곳, 특성화고 1곳, 일반고 4곳의 총 7개 고등학교에서 예비 검사에 참여하지 않은 2학년 학생 568명을 대상으로 본 검사를 실시하였다. 검사 결과를 수합한 후에는 확인적 요인 분석(confirmatory factor analysis)을 실시하여 12개의 본 검사 문항들이 타당도와 신뢰도를 갖추고 있는지를 확인하고자 하였다. 본래 계획에서는 확인 단계에서 단 한 차례의 조사만을 수행하고자 하였으나, 최초 본 검사 결과에 대한 분석 과정에서 확인된 타당도와 신뢰도가 연구자들이 설정한 기준에 조금 못 미치는 것으로 나타났다. 이에 일부 문항을 수정한 후 2017년 3월에 특목고 1곳, 일반고 3곳의 총 4개교에서 예비 검사와 1차 본 검사에 참여하지 않은 1학년 학생 487명을 대상으로 2차 본 검사를 새로 실시하고 이를 통해 최종 확증에 이를 수 있었다.

2. 자료 수집

자료 수집은 탐색 단계에서 1회(예비 검사), 확인 단계에서 2회(1차 본 검사 및 2차 본 검사)로, 총 3회에 걸쳐 이루어졌다. 예비 검사에서는 특목고, 특성화고, 일반고가 포함된 총 4개교에서 479명(결측 제외 467명)의 1, 2학년 학생들을 대상으로 검사를 실시하였다. 1차 본 검사에서는 특목고, 특성화고, 일반고가 포함된 총 7개교에서 583명(결측 제외 568명)의 2학년 학생들을 대상으로 검사를 실시하였다. 2차 본 검사에서는 특목고와 일반고 총 4개교에서 498명(결측 제외 487명)의 1학년 학생들을 대상으로 검사를 실시

〈표 2〉 검사 참여 인원(결측 제외)

검사 시기	인원(학교 수)			
	특목고	특성화고	일반고	계
예비 검사	98명(1개교)	157명(1개교)	212명(2개교)	467명(4개교)
1차 본 검사	65명(2개교)	92명(1개교)	311명(4개교)	568명(7개교)
2차 본 검사	35명(1개교)		452명(3개교)	487명(4개교)

하였다.¹⁵⁾ 본 연구에서 이루어진 총 3회의 검사 중 동일 학교를 2회 이상 참여시킨 경우도 있었으나(특목고 1곳, 일반고 1곳), 이 경우 각 검사 시기마다 서로 다른 학급이나 학년을 대상으로 검사를 실시함으로써 동일 참여자가 두 차례 이상 반복 참여하지 못하도록 조치하였다. 각 검사 시기별 참여 인원은 Backstrom과 Hursh(1963)에서 오차범위 $\pm 5\%$, 신뢰한계 95% 수준을 충족하기 위해 최소 384명의 인원이 필요하다고 한 것을 참고하여 결정하였다. 검사 참여 인원에 관한 세부 정보는 〈표 2〉에서 확인할 수 있다.

3. 예비 문항

예비 검사 시 사용한 최초의 27개 예비 문항을 〈표 3〉에서 확인할 수 있다. 해당 문항들은 내용 타당도 점검을 거친 후 구인 조정과 함께 진술 방식을 수정한 것들이다. ‘사려 깊은 행동’에 9개 문항, ‘내용·과정 성찰’에 13개 문항, ‘전체 성찰’에 5개 문항이 예비 문항으로 작성되었다. ‘사려 깊은 행동’에 속한 문항들 중 ‘(-)’로 표시된 3개 문항은 내용 타당도 점검 전 ‘습관적

15) 1차 본 검사와 2차 본 검사에서 1, 2학년들을 고루 포함하는 대신 특정 학년을 선택한 까닭은 본 연구의 관심과는 별도로 학교 유형 간 차이 여부를 확인하고자 하려는 목적으로 조치된 것이다. 또한 3차 조사 시 특성화고 재학생들을 표집 대상으로 포함시키지 않은 점은 앞선 예비 검사 및 1차 본 검사에서 특성화고 학생들과 일반고 학생들에게서 특별한 차별성을 찾기 어려웠을 뿐만 아니라, 불성실 응답자(즉, 줄 세우기)가 지나치게 많아 이들이 신뢰도를 떨어뜨렸던 점을 고려한 것이다.

〈표 3〉 대화 성찰 태도 측정을 위한 예비 문항(27개)

구인	구별번호	문항
사려 깊은 행동	T1	대화에서 상대의 의중을 파악할 수 없을 때는 상대의 행동을 더 자세하게 관찰한다.
	T2	대화에서 화제의 변화를 파악하려고 노력한다.
	T3	대화에서 문제가 생기더라도 그때그때 생각나는 대로 대처하는 편이다.(-)
	T4	대화에서 어떻게 말해야 할지를 고민하는 것은 중요하다고 생각한다.
	T5	대화에서 상대방의 반응이나 반응 방식의 변화에 무관심한 편이다.(-)
	T6	상대의 말에 기분이 상하더라도 바로 감정을 드러내지 않는다.
	T7	평소 대화에서 아무렇게나 생각나는 대로 말하더라도 괜찮다고 생각한다.(-)
	T8	나의 말이 상대의 감정에 미칠 영향을 생각하며 말할 내용을 신중히 고민한다.
	T9	대화가 나의 예상과 달리 전개되거나 문제가 발생했을 때는 더욱 조심해서 말하려고 한다.
내용·과정 성찰	R1	대화에서 문제를 겪고 나면 이후에 그러한 문제를 어떻게 예방할 수 있을지를 생각해 본다.
	R2	대화를 나눈 후 상대가 타당한 의견을 제시했는지 판단한다.
	R3	대화를 나눈 후 상대에게 정확한 사실을 말했는지 확인한다.
	R4	누군가와 갈등을 겪고 나면 상대와 나눴던 대화를 떠올리며 문제의 원인을 찾아보려 노력한다.
	R5	대화를 나눈 후 상대가 솔직한 감정을 표현했는지 점검한다.
	R6	대화를 나눈 후 상대가 정확한 사실을 말했는지 점검한다.
	R7	이전 대화에서의 문제 해결을 위해 이와 유사했던 과거의 다른 대화 경험들을 떠올려본다.
	R8	대화의 흐름이나 상대의 반응에 따라 최선의 대화 방식이 무엇일지 생각하며 대화한다.
	R9	대화를 나눈 후 상대에게 나의 감정을 솔직하게 표현했는지 고민한다.
	R10	대화를 나눈 후 상대에게 타당한 의견을 제시했는지 반성한다.
	R11	대화를 더 잘하기 위해 끊임없이 새로운 방법을 시도해 본다.
	R12	대화에서의 문제 해결을 위해 다른 사람들로부터 의견을 구하거나 책, 인터넷 등을 찾아본다.
	R13	다양한 대화방식 중 나에게 맞는 최선의 대화방법을 찾기 위해 노력한다.
전제 성찰	C1	나의 평소 대화 습관이 바람직한지를 생각해 본다.
	C2	누군가의 잘못된 대화 방식을 무의식적으로 따라하고 있지 않은지 돌아본다.
	C3	바람직한 대화에 관한 여러 관점을 비교해 보고 이를 바탕으로 나의 관점이 적절한지 점검해 본다.
	C4	좋은 대화에 대한 나의 가정에 문제가 없는지 고민해 본다.
	C5	생각 없이 유행을 따라 잘못된 표현을 사용하고 있지 않은지 반성한다.

‘행동’의 측정 문항으로 작성된 것들 중 일부를 진술 방식을 바꾸지 않고 ‘사려 깊은 행동’의 역변별 측정 문항으로 남겨둔 것들이다.

4. 자료 분석

첫 번째 연구문제인 대화 성찰 태도의 3요인 모형이 타당한지를 확인하기 위해 요인 분석을 실시하였다. 요인 분석의 일반 절차에 따라(cf. 이순묵, 2000), 예비 검사 결과에 대해서는 탐색적 요인 분석을, 1·2차 본 검사 결과에 대해서는 확인적 요인 분석을 실시하였다.

탐색적 요인 분석은 SAS 9.4를 사용하여 이루어졌다.¹⁶⁾ 축소상관행렬을 분해하는 방법으로는 최대 우도법(Maximum Likelihood method)을 사용하였고,¹⁷⁾ 요인 구조의 회전 방법으로는 사각회전의 대표적 방식인 프로맥

-
- 16) 여러 연구들에서 통계적 분석을 위해 두루 사용하고 있는 SPSS 대신 SAS를 사용한 까닭은 다음과 같다. 첫째, 요인 분석에서는 요인분석 시 단일분해를 하지 못하고 재분해만이 가능하다. 이순묵(2000: 35)에 따르면 단일분해는 과소 추정의 위험이, 재분해는 과대 추정의 위험이 있는데 둘 중 하나의 위험을 선택해야 한다면 과소 추정의 위험이 안전하므로 가능하다면 단일분해를 선택하는 것이 바람직할 것이다. 둘째, 요인분석은 주성분분석이 아닌 공통요인분석을 기초로 하는데, SPSS에서는 공통요인분석을 지정해도 주성분분석에서의 고유치만을 제공한다(이순묵, 2000: 37). 고유치는 요인 수의 결정 등에서 중요한 역할을 하기도 하므로 공통요인분석을 통해 정확한 고유치를 산출하는 것이 바람직할 것이다. 셋째, SAS는 반복주축분해 시 분해가 반복될수록 각 변수의 공통분이 커지는 것을 제어할 수 있도록 HEYWOOD 옵션을 제공하는데 반해, SPSS에서는 이러한 옵션이 제공되지 않는다. 물론, 본 연구에서는 최종 분석 방식으로 주축분해법이 아닌 최대 우도법을 선택하였으므로 SPSS를 사용하였더라도 해당 프로그램이 지닌 많은 한계를 비껴갈 수 있었을 것이다. 그러나 이처럼 최대 우도법을 선택하기까지 다양한 방법들을 비교하며 탐색해 보는 과정이 필요했으므로 분석의 타당성을 높이기 위해 SAS를 사용하는 것은 필수적이었다.
 - 17) 최대 우도법을 사용하기 위해서는 기본적으로 몇 개의 요인이 존재하는지에 대한 가정과 측정변수가 다변량 정규 분포를 따른다는 가정이 전제되어야 한다(이순묵, 2000: 53). 본 연구에서는 연구 설계에서부터 3개의 요인을 가정하고 있으므로 최대 우도법 사용을 위한 첫 번째 가정이 성립한다. 두 번째 가정과 관련하여 탐색적 요인 분석에 투입된 27개

스(Promax) 방법을 사용하였다.¹⁸⁾ 연구에서 가정된 3요인 모형에 적합한 문항을 선별하기 위해 각 문항의 공통분(communality)이 .30 이상을 보이는지, 가정된 구인에 대한 요인 부하량(factor loading)이 .40 이상을 보이는지, 가정되지 않은 구인에 대한 요인 부하량이 .30 미만을 보이는지 등을 종합적으로 점검하였다. 또한, 인접 모형인 2요인 모형 및 4요인 모형에 비해 3요인 모형이 더 적합한지를 검토하기 위해 카이제곱 검정 결과와 터커 - 루이스 지수(TLI)를 비교하였다.

확인적 요인 분석은 Lisrel 9.2를 사용하여 이루어졌다.¹⁹⁾ 모형의 전반적 합치도를 보여주는 지수들 중 좋은 합치도의 기준을 지니고 있는 개략화 오차평균(RMSEA), 표준합치도(NFI), 기초합치도(GFI), 비교합치도(CFI), 원소 간 평균차이(RMR)를 종합적으로 확인하였으며, 좋은 합치도의 기준은 이순묵(2000: 134)의 제안에 따라 RMSEA<.08, NFI>.09, GFI>.09, CFI>.09, RMR<.05으로 삼았다.

두 번째 연구문제인 문항의 신뢰도를 판단하기 위해서는 크론바흐의 알

측정변수의 웨도는 $-1.041 \sim 853$ 의 범위를 나타내었으며, 첨도는 $-848 \sim 936$ 의 범위를 나타내 모두 정규성을 지닌 것으로 볼 수 있어 최대 우도법을 사용할 수 있다고 판단하였다.

- 18) 요인구조의 회전 방법 중 직각회전이 요인 간 각도가 90° 라고 가정하는 것은 그 각도의 코사인(cosine) 값에 해당하는 요인 간 상관계수가 0임을 가정하는 것이다(이순묵, 2000: 74-75). 여러 연구들이 요인 간 상관이 있음을 전제하고 있더라도 결과의 간명함을 위해 관례적으로 직각회전의 방식을 사용하기도 하는 것으로 보이나, 본 연구에서는 성찰의 요인 간 상관이 0이라는 비현실적 가정을 적용하기에 무리가 있다고 판단하였다. 실제로 성찰의 과정에서는 그 하위 구인들이 유기적으로 관련되어 작용할 것이므로 이들을 쉽게 분리하여 구별하기 어려울 것이기 때문이다. 이에 본 연구에서는 요인 간 상관이 존재함을 가정하는 사각회전의 방식을 사용하고자 하였다.
- 19) Lisrel을 통한 구조방정식 모형을 사용하기 위해서도 앞선 탐색적 요인 분석에서의 최대 우도법과 마찬가지로 측정변인들이 다변량 정규 분포를 이룬다는 가정이 필요하다. 1차 본 검사에서의 12개 측정변인들은 웨도가 $-1.148 \sim .837$ 의 범위를, 첨도가 $-.556 \sim 1.928$ 의 범위를 나타냈으며, 2차 본 검사에서의 12개 측정변인은 웨도가 $-1.026 \sim .780$ 의 범위를, 첨도가 $-.683 \sim 1.097$ 의 범위를 나타내어 모두 정규분포를 이루는 것으로 간주할 수 있었다.

파(Cronbach's α) 계수를 산출하여 구인별 문항들의 내적 일관성 신뢰도를 확인하고자 하였다. 특히, 구인별로 각각의 문항을 제거했을 때 구인 전체의 알파 값이 높아지는 문항을 신뢰도가 낮은 문항이라고 판단하였다. 신뢰도 확인에서는 SPSS 22.0을 사용하였다.

IV. 연구 결과

1. 탐색적 요인 분석 결과

탐색적 요인 분석 과정에서 분석 방법을 달리하여도 늘 함께 묶여 본 연구에서 가정하지 않은 단일한 구인을 형성하는 세 개 문항(R2, R5, R6)을 발견하여 이들을 우선적으로 소거하였다.²⁰⁾ 다음으로는 공통분(communality)이 0.3 미만인 문항이 나타나지 않을 때까지 7개 문항을 공통분이 작은 것부터 차례대로(T3, T6, T1, R3, T5, R12, R4의 순으로) 삭제하였다. 끝으로, 공통분이 0.3 이상이지만 복수의 구인에 0.3 이상의 요인 부하량(factor loading)을 보인 5개 문항을 두 번째 요인 부하량의 크기가 큰 것부터 차례대로 (R10, R1, C3, R7, T2의 순으로) 삭제하였다.²¹⁾ 이와 같은 과정을 거쳐 최종적으로 선택된 12개 문항은 <표 4>와 같은 요인 구조 및 신뢰도를 보였다. 구체적으로 살펴보면 다음과 같다.

첫째, ‘사려 깊은 행동’으로 가정된 4개 문항(T4, T8, T7, T9)은 모두 제

-
- 20) 이들 문항은 모두 자신이 아닌 타인의 발화 내용을 판단하거나 점검하는지를 묻는 문항들이다.
 - 21) 해당 5개 문항 제거 후에도 새로이 C4, R8의 두 번째 요인 부하량 크기가 각각 .43, .35로 꽤 크게 나타났으나(<표 4> 참고), 문항을 더 이상 소거할 경우 요인 구조가 형성되지 않음을 확인하였기에 해당 시점에서 문항 소거를 중지하였다.

〈표 4〉 선별 문항의 요인 부하량 및 구인별 신뢰도

구인	구별 번호	요인 부하량			Cronbach α	
		제1요인	제2요인	제3요인	구인별	제거시
사려 깊은 행동	T4	.16	.60	.18	.65	.54
	T8	.26	.57	.14		.55
	T7	-.11	.53	.15		.62
	T9	.30	.45	.09		.60
내용·과정 성찰	R11	.71	.08	.09	.71	.64
	R13	.59	.15	.27		.63
	R9	.48	.13	.30		.65
	R8	.43	.35	.18		.67
전체 성찰	C2	.10	.18	.66	.67	.58
	C5	.19	.22	.49		.61
	C1	.31	.21	.45		.60
	C4	.43	.04	.44		.63

* 각 구별 번호가 가리키는 실제 문항은 〈표 3〉을 참고.

2요인에 .40 이상의 요인 부하량을 보였다. 구인 신뢰도는 .65로 비교적 양호하게 나타났으며, 문항 제거 시 해당 구인의 신뢰도를 높이는 문항은 없었다. 둘째, ‘내용·과정 성찰’로 가정된 4개 문항(R11, R13, R9, R8) 또한 모두 제1요인에 .40 이상의 요인 부하량을 보였다. 구인 신뢰도는 .71로 비교적 양호했으며, 문항 제거 시 신뢰도를 높이는 문항은 없는 것으로 나타났다. 셋째, ‘전체 성찰’로 가정된 4개 문항(C2, C5, C1, C4)은 모두 제3요인에 .40 이상의 요인 부하량을 보였고, 구인 신뢰도는 .67로 비교적 양호한 가운데 문항 제거 시 신뢰도를 높이는 문항은 없었다.

이상과 같이 최종 선별된 12개 문항을 바탕으로, 본고에서 가정한 3요인 모형이 인접한 2요인 모형 및 4요인 모형과 비교했을 때 더 타당한지를 확인하기 위해 카이제곱 검정 결과와 터커-루이스 계수(TLI)를 확인한 결과는 〈표 5〉와 같다. 구체적으로 살펴보면 다음과 같다.

〈표 5〉 선별 문항의 요인 수효(2요인, 3요인, 4요인) 비교

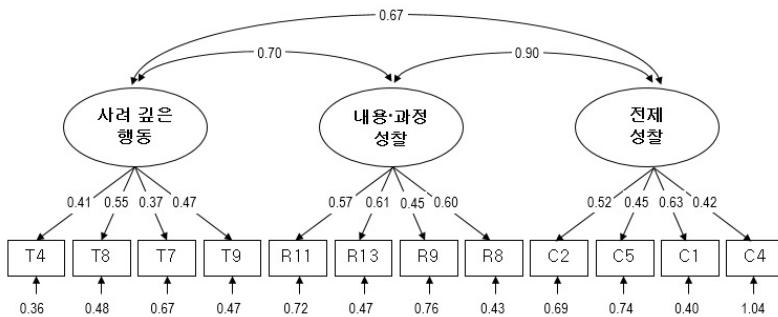
	2요인 모형	3요인 모형	4요인 모형
카이제곱 검정	$\chi^2 = 118.69$, df=43, p<.001	$\chi^2 = 45.09$, df=33, p = .08	$\chi^2 = 23.34$, df=24, p = .50
TLI	.899	.979	1.002

2요인 모형의 경우 카이제곱 검정 결과 ‘2개 요인이 존재한다.’라는 영 가설이 기각되고, 대신에 ‘2개 이상의 요인이 존재한다.’라는 대립가설이 채택되었다(p<.001). TLI 값(.899) 또한 기준치인 .95에 미치지 못하는 것으로 나타났다. 반면, 3요인 모형의 경우 카이제곱 검정 결과 ‘3개 요인이 존재한다.’라는 영가설이 채택될 수 있으며(p=.08), TLI 값(.979)이 기준치를 충족하는 것으로 나타났다. 한편, 4요인 모형의 경우에도 카이제곱 검정 결과 ‘4개 요인이 존재한다.’라는 영가설이 채택될 수 있었으나(p=.50), TLI 값(1.002)이 이론적 최대치인 1.0을 넘기고 있어 지나치게 포화모형에 가깝다는 점을 알 수 있었다.²²⁾ 이에 27개 예비 문항들 중 최종 선택된 12개 문항들은 본고가 가정한 3요인 모형으로 가장 적절히 설명될 수 있다는 결론을 내릴 수 있었다.

2. 확인적 요인 분석 결과

예비 검사 결과에 대한 탐색적 요인 분석 및 신뢰도 분석 결과를 토대로 선별된 12개 문항으로 1차 본 검사를 실시하였다. 검사 결과에 대해 확인적 요인 분석을 실시한 결과, 〈그림 3〉과 같은 요인 구조와 〈표 6〉과 같은 합치도 지수를 확인할 수 있었다. 〈그림 3〉에서 볼 수 있듯 12개 문항이 각 구인

22) 이순복(2000: 52)에 따르면, ‘연구자가 자꾸지꾸 많은 수의 요인을 가정할수록 터커-루 이스 계수의 값은 더욱더 옥 커지는데 이것은 자료 속에 있는 표본 고유의 요인들까지 요인구조에 거두어 넣는 잘못을 범할 가능성을 수반한다.’



〈그림 3〉 1차 본 검사 결과의 요인 구조

〈표 6〉 1차 본 검사 결과의 모형 합치도 지수

	카이제곱	NFI	CFI	GFI	RMR	RMSEA
기준	없음	.90 이상	.90 이상	.90 이상	.05 미만	.08 미만
1차 본 검사 결과	$\chi^2=216.74$, df=51, p<.001	.849	.879	.941	.046	.076

에 대해 지난 요인 부하량의 크기는 한 개 문항(T7)을 제외하고는 모두 .40 이상의 크기를 보였으며, T7 문항 또한 .37의 요인 부하량을 보여 용인 가능한 수준이라고 판단할 수 있었다. 그러나 〈표 6〉에서 볼 수 있듯 모형의 전반적 합치도 지수들 중 NFI(=.849)와 CFI(=.879)가 연구자들이 따르는 기준인 .90에 못 미치는 것으로 나타나 1차 본 검사에서 사용한 12개 문항을 최종 문항으로 확정하기에는 무리가 있을 수 있다고 판단하였다.

이처럼 모형 합치도 지수가 양호하지 않았던 이유는 1차 본 검사 결과에 대한 신뢰도 분석을 통해 일정 부분 확인할 수 있었다. 〈표 7〉에서와 같이 ‘내용·과정 성찰’의 측정 문항인 R9와 ‘전제 성찰’의 측정 문항인 C4가 제거되었을 때 구인의 신뢰도가 각각 .67에서 .69로, .59에서 .62로 높아지는 것으로 나타났다. 이와 관련하여 1차 본 검사에 협조한 교사들과 협의한 결과, 이들 두 개 문항의 의미가 잘 이해되지 않아 질문을 한 경우가 있었다는 사

〈표 7〉 1차 본 검사 문항의 신뢰도

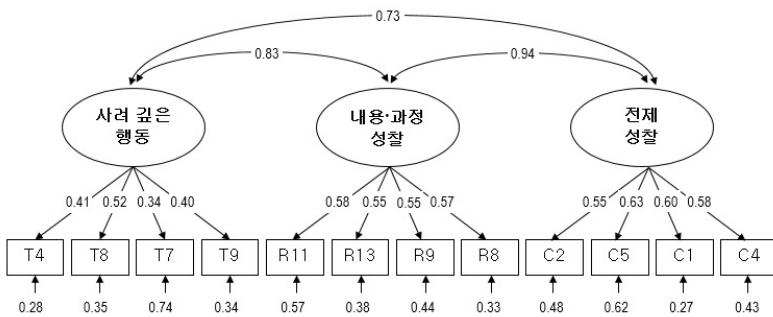
구인	구별 번호	문항	Cronbach α	
			구인별	제거시
사려 깊은 행동	T4	대화에서 어떻게 말해야 할지를 고민하는 것은 중요하다고 생각한다.	.62	.52
	T8	나의 말이 상대의 감정에 미칠 영향을 생각하며 말할 내용을 신중히 고민한다.		.60
	T7	평소 대화에서 아무렇게나 생각나는 대로 말하더라도 괜찮다고 생각한다.(-)		.56
	T9	대화가 나의 예상과 달리 전개되거나 문제가 발생했을 때는 더욱 조심해서 말하려고 한다.		.54
내용·과정 성찰	R11	대화를 더 잘하기 위해 끊임없이 새로운 방법을 시도해 본다.	.67	.57
	R13	다양한 대화방식 중 나에게 맞는 최선의 대화방법을 찾기 위해 노력한다.		.56
	R9	대화를 나눈 후 상대에게 나의 감정을 솔직하게 표현했는지 고민한다.		.69
	R8	대화의 흐름이나 상대의 반응에 따라 최선의 대화 방식이 무엇일지 생각하며 대화한다.		.60
전제 성찰	C2	누군가의 잘못된 대화 방식을 무의식적으로 따라하고 있지 않은지 돌아본다.	.59	.47
	C5	생각 없이 유행을 따라 잘못된 표현을 사용하고 있지 않은지 반성한다.		.51
	C1	나의 평소 대화 습관이 바람직한지를 생각해 본다.		.47
	C4	좋은 대화에 대한 나의 가정에 문제가 없는지 고민해 본다.		.62

* 음영은 신뢰도가 양호하지 않은 문항임.

실을 확인할 수 있었다.²³⁾

이에 R9 문항을 ‘대화를 나눈 후 나의 생각과 감정이 상대에게 왜곡 없이 잘 전달되었는지 돌아본다.’로 수정하고, C4 문항을 ‘어떤 대화가 좋은 대화인지에 관한 나의 믿음이나 생각에 문제가 없는지를 고민한다.’로 수정하

23) 가령, C4 문항의 경우, ‘가정’이라는 단어를 ‘가정(假定)’이 아닌 ‘가정(家庭)’으로 이해하고 질문하는 학생이 있었다고 한다.



〈그림 4〉 2차 본 검사 결과의 요인 구조

〈표 8〉 2차 본 검사 결과의 모형 합치도 지수

	카이제곱	NFI	CFI	GFI	RMR	RMSEA
기준	없음	.90 이상	.90 이상	.90 이상	.05 미만	.08 미만
2차 본 검사 결과	$\chi^2=194.40$, $df=51$, $p<.001$.897	.921	.936	.037	.076

여 새로이 설문지를 마련하였다. 그리고 새로운 학교 및 학생들을 대상으로 2차 본 검사를 실시하였으며, 이렇게 얻은 측정 결과에 대해 다시금 확인적 요인 분석을 실시하였다. 그 결과 이번에는 〈그림 4〉와 같은 요인 구조와 〈표 8〉과 같은 모형의 전반적 합치도 지수를 확인할 수 있었다. 〈그림 4〉에서 확인할 수 있듯이, 이번에도 1차 본 검사에서와 같이 T7 문항의 요인 부하량만이 .34로 나타날 뿐, 나머지 11개 문항은 모두 .40 이상의 요인 부하량을 보였다. 또한 〈표 8〉에서 확인할 수 있듯이, 이번에는 NFI를 제외한 모든 합치도 지수가 기준을 충족하였으며, NFI 또한 .897로 나타나 기준에 거의 근접함을 확인할 수 있었다.

〈표 9〉에서 확인할 수 있듯이, 1차 본 검사 결과에서는 R9 문항과 C4 문항이 각 구인의 신뢰도를 떨어뜨리는 것으로 나타난 데 반해, 2차 본 검사에서는

〈표 9〉 2차 본 검사 문항의 신뢰도

구인	구별 번호	문항	Cronbach α	
			구인별	제거시
사려 깊은 행동	T4	대화에서 어떻게 말해야 할지를 고민하는 것은 중요하다고 생각한다.	.64	.55
	T8	나의 말이 상대의 감정에 미칠 영향을 생각하며 말할 내용을 신중히 고민한다.		.57
	T7	평소 대화에서 아무렇게나 생각나는 대로 말하더라도 괜찮다고 생각한다.(-)		.62
	T9	대화가 나의 예상과 달리 전개되거나 문제가 발생했을 때는 더욱 조심해서 말하려고 한다.		.57
내용·과정 성찰	R11	대화를 더 잘하기 위해 끊임없이 새로운 방법을 시도해 본다.	.75	.69
	R13	다양한 대화방식 중 나에게 맞는 최선의 대화방법을 찾기 위해 노력한다.		.69
	R9'	대화를 나눈 후 나의 생각과 감정이 상대에게 왜곡 없이 잘 전달되었는지 돌아본다.		.72
	R8	대화의 흐름이나 상대의 반응에 따라 최선의 대화 방식이 무엇일지 생각하며 대화한다.		.68
전제 성찰	C2	누군가의 잘못된 대화 방식을 무의식적으로 따라하고 있지 않은지 돌아본다.	.76	.70
	C5	생각 없이 유행을 따라 잘못된 표현을 사용하고 있지 않은지 반성한다.		.70
	C1	나의 평소 대화 습관이 바람직한지를 생각해 본다.		.68
	C4'	어떤 대화가 좋은 대화인지에 관한 나의 믿음이나 생각에 문제가 없는지 고민한다.		.73

* 음영은 2차 본 검사에서 수정된 문항임.

수정된 R9' 문항과 C4' 문항(수정된 문항임을 나타내기 위해 구별 번호 뒤에 콤마(,)로 표시함.)이 구인의 신뢰도를 높이는 데 기여하는 것으로 나타났다.

이상과 같은 결과를 토대로 연구자들은 〈표 9〉에서 제시한 최종 12개 문항이 본 연구가 가정한 대화 성찰 태도의 3개 구인을 측정하기에 신뢰할 만하고 타당하다는 점을 확증할 수 있었다. 최종 확증된 12개 문항을 통한 측정 결과는 〈표 10〉에서 확인할 수 있다.

〈표 10〉 2차 본 검사 문항의 측정값

구인	평균(표준편차)				
	전체(N=487)	일반고			특목고(N=35)
		전체(N=452)	남고(N=240)	여고(N=212)	
사려 깊은 행동	4.17(.54)	4.17(.54)	4.16(.56)	4.19(.51)	4.17(.54)
내용·과정 성찰	3.77(.65)	3.75(.66)	3.81(.67)	3.67(.63)	4.03(.57)
전제 성찰	3.65(.68)	3.63(.68)	3.63(.70)	3.64(.66)	3.86(.68)

V. 결론

본 연구는 성찰에 관해 범학문적으로 이루어져 온 이론적, 실증적 논의에 대한 검토를 바탕으로 대화 성찰 태도를 정의하고 그 측정을 위한 하위 구인을 설정하였다. 그리고 이를 기초로 대화 성찰 태도를 측정할 수 있는 27 개의 예비 문항을 개발한 후 이 가운데 양호한 구인 타당도와 신뢰도를 보여 주는 문항을 선별하고자 하였다. 이를 위해 총 1,522명의 고등학교 1~2학년 학생들을 대상으로 예비 검사를 1회, 본 검사를 2회 실시하였으며, 이 과정에서 문항의 선별 및 수정을 거쳐 최종 12개 문항이 모든 조건에 부합한다는 확증에 이를 수 있었다. 구체적인 확증의 근거는 다음과 같다.

우선, 본 연구가 예비 문항 중 최종 선정한 12개 문항은 ‘사려 깊은 행동’, ‘내용·과정 성찰’, ‘전제 성찰’로 구성된 3요인 모형의 가정에 부합하는 구인 타당도를 보여 주었다. 구인 타당도가 양호하다는 점은 첫째, 탐색적 요인분석 결과 인접한 2요인 모형의 경우 ‘2개 요인이 존재한다.’라는 영가설이 기각된 것과 달리, 3요인 모형의 경우 ‘3개 요인이 존재한다.’라는 영가설이 채택되었다는 점, 둘째, 인접한 4요인 모형의 경우 지나친 포화 모형으로 나타난 데 반해, 3요인 모형의 경우 적정한 모형 합치도를 보여주었다는 점

을 통해 1차적으로 확증되었다. 또한, 확인적 요인 분석 결과 대부분의 모형 합치도 지수가 양호함으로써 최종 확증에 이를 수 있었다.

다음으로 최종 12개 문항은 모두 각 문항이 측정하고자 하는 구인의 측정 신뢰도를 높이는 데 기여한다는 점을 확인할 수 있었다. ‘사려 깊은 행동’의 알파 값은 .64, ‘내용·과정 성찰’의 알파 값은 .75, ‘전체 성찰’의 알파 값은 .76으로 나타난 가운데, 각 구인의 측정 문항들은 모두 해당 문항을 제거 했을 때 오히려 구인의 알파 값을 떨어뜨리는 것으로 나타났다.

이상과 같이 확증된 12개 문항으로 측정한 3개 구인의 측정값은 사려 깊은 행동보다 내용 및 과정 성찰이, 내용 및 과정 성찰보다 비판적 성찰이 더 높은 수준의 성찰에 관계한다고 보는 Mezirow(1990, 1991)에서의 논의에 대체로 부합하는 모습을 보였다. ‘사려 깊은 행동’의 측정값 평균은 4.18로 가장 높았고, ‘내용·과정 성찰’의 측정값 평균은 3.77로 중간 수준이었으며, ‘전체 성찰’의 측정값 평균은 3.65로 가장 낮게 나타났다.²⁴⁾ 이러한 결과는 평균적으로 보아 ‘전체 성찰’, ‘내용·과정 성찰’, ‘사려 깊은 행동’의 순으로 학습자가 도달하기가 어려운 구인일 것임을 시사하는 것으로 보인다. 또한 대체로 학력 수준에 차이가 있다고 간주되는 특목고 학생과 일반고 학생 간에도 어느 정도 차이가 확인되었다. 사려 깊은 행동(특목고 M=4.17, 일반고 M=4.17)의 측정 결과에서는 차이가 없었으나, 특목고 학생들은 일반고 학생에 비해 내용·과정 성찰(특목고 M=4.03, 일반고 M=3.76)과 전체 성찰(특목고 M=3.86, 일반고 M=3.63)의 평균 측정값이 더 높은 것으로 나타났다. 이러한 결과는 비록 본 연구에서 개발한 문항들이 대화 성찰 ‘태도’를 측정하는 것이기는 하지만 대화 성찰의 사고 수준이 측정 결과에 어느 정도 영향을 미치고 있음을 시사하는 것으로 보인다.

24) 반복측정 분산분석 결과, 3개 구인 간 개인 내적 차이는 유의한 것으로 확인되었으며 ($p<.001$), 대비 검정 결과 ‘사려 깊은 행동’과 ‘내용·과정 성찰’, ‘사려 깊은 행동’과 ‘전체 성찰’, ‘내용·과정 성찰’과 ‘비판적 성찰’ 간 차이 또한 모두 유의한 것을 확인할 수 있었다($p<.001$).

본 연구에서 제시한 대화 성찰 태도 측정 문항은 향후 성찰적 대화 교육을 위한 질적·양적 연구들에서 다양한 목적으로 사용 가능할 것으로 보인다. 가령 질적 접근 방식을 통해 대화 성찰 태도가 높은 학습자와 낮은 학습자의 사고 특성 차이를 탐색하고자 한다면, 최초의 사례 선정 단계에서 대화 성찰 태도가 높거나 낮은 학습자를 선별하는 기준을 마련하기 위해 본 연구가 개발한 설문 문항을 활용할 수 있을 것이다. 혹은 성찰적 대화 교육의 설계나 교육 내용을 개발하여 실험적으로 적용하고 그러한 새로운 적용이 효과적이었는지를 확인하고자 한다면, 학습자의 대화 성찰 태도를 교육 전과 후에 반복 측정한 뒤 유의미한 향상이 있었는지를 검증하기 위해 본 연구의 설문 문항을 활용해 볼 수도 있을 것이다. 그 밖에도 본 연구가 개발한 설문 문항은 향후 성찰적 대화 교육에 관한 다양한 연구를 추동할 수 있는 연구 도구로서 다양한 활용 가능성을 지니고 있으므로, 향후 본 연구를 기반으로 한 다양한 후속 연구가 이어지기를 기대해 본다.

- * 본 논문은 2017. 10. 31. 투고되었으며, 2017. 11. 21. 심사가 시작되어 2017. 12. 11. 심사가 종료되었음.

참고문헌

- 김정란(2014), 「반성적 말하기 지도 방법 연구: ALACT 교수 모형 적용을 중심으로」, 『화법연구』 25, 75-101, 한국화법학회.
- 박경·김혜은(2017), 『심리평가의 이해와 활용』, 학지사.
- 박영민(2006), 「반성적 쓰기를 활용한 작문 평가 방안」, 『새국어교육』 73, 33-59, 한국국어교육학회.
- 박영민·최숙기(2008), 「반성적 쓰기가 중학생의 설명문 쓰기 수행에 미치는 영향」, 『국어교육』 125, 137-161, 한국어교육학회.
- 박창균(2015), 「화자의 스파치 수행에 대한 반성적 사고 양상 연구」, 『화법연구』 30, 217-248, 한국화법학회.
- 서종훈(2009), 「말하기 수행에 드러난 반성적 사고 양상: 고1 학습자들의 공적인 자리에서의 말하기를 중심으로」, 『새국어교육』 83, 239-260, 한국국어교육학회.
- 이순목(2000), 『요인분석의 기초』, 교육과학사.
- 장지혜(2013), 「성찰적 사고를 위한 비평적 에세이 쓰기 교육 연구」, 서울대학교 석사학위논문.
- 최홍원(2006), 「성찰적 사고의 문학 교육적 구도: <도산십이곡>과 <여부사시사>를 중심으로」, 『문학교육학』 21, 385-423, 한국문학교육학회.
- 최홍원(2008), 「시조의 성찰적 사고 교육 연구」, 서울대학교 박사학위논문.
- Allport, G. W. (1935), "Attitudes," In: C. Murchison. (ed.) *A Handbook of Social Psychology*, 798-844, London: Clark university press.
- Backstrom, C. H., & Hursh, G. D. (1963), *Survey research*. Illinois: Northwestern University Press.
- Bain, J. D., Ballantyne, R., Packer, J., & Mills, C. (1999), "Using journal writing to enhance student teachers' reflectivity during field experience placements," *Teachers and Teaching* 5(1), 51-73.
- Dewey, J. (2011), 『하우 위 싱크: 과학적 사고의 방법과 교육』, 정희숙 역, 학이시습(원서출판 1910).
- Hatton, N., & Smith, D. (1995), "Reflection in teacher education: Towards definition and implementation," *Teaching and Teacher Education* 11(1), 33-49.
- Jay, J. K., & Johnson, K. L. (2002), "Capturing complexity: A typology of reflective practice for teacher education," *Teaching and Teacher Education* 18(1), 73-85.
- Kember, D., Leung, D. Y. P., Jones, A., Loke, A. Y., McKay, J., Sinclair, K., Tse, H., Webb, C., Wong, F. K. Y., Wong, M., & Yeung, E. (2000), "Development of a questionnaire to measure the level of reflective thinking," *Assessment & Evaluation in Higher Education* 25(4), 381-395.
- Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008), "A four-category scheme for

- coding and assessing the level of reflection in written work," *Assessment & Evaluation in Higher Education* 33(4), 369-379.
- Kreber, C., & Castleden, H.(2009), "Reflection on teaching and epistemological structure: reflective and critically reflective processes in 'pure/soft' and 'pure/hard' fields," *Higher Education* 57(4), 509-531.
- Larrivee, B.(2008), "Development of a tool to assess teachers' level of reflective practice," *Reflective Practice* 9(3), 341-360.
- Lau, K.(2015), "Assessing reflection in English enhancement courses: teachers' views and development of a holistic framework," *Assessment & Evaluation in Higher Education* 41(6), 854-868.
- McAlpine, L., Weston, C., Berthiaume, D., Fairbank-Roch, G., & Owen, M.(2004), "Reflection on teaching: Types and goals of reflection," *Educational Research and Evaluation* 10(4-6), 337-363.
- Mezirow, J.(1990), "How critical reflection triggers transformative learning," In: J. Mezirow (ed.) *Fostering Critical Reflection in Adulthood*, 1-20, San Francisco: Jossey-Bass.
- Mezirow, J.(1991), *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.
- Rogers, R. R.(2001), "Reflection in higher education: A concept analysis," *Innovative higher education* 26(1), 37-57.
- Schön, D. A.(1983), *The Reflective Practitioner: How professionals think in action*. NY: Basic Books.
- Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P.(2012), "Fostering and evaluating reflective capacity in medical education: developing the REFLECT rubric for assessing reflective writing," *Academic Medicine* 87(1), 41-50.

대화 성찰 태도 측정을 위한 설문도구 개발

민병곤·박성석

본 연구는 성찰적 교육에 관한 법학문적 논의에 대한 검토를 바탕으로 ‘대화 성찰 태도’를 정의하고, 이를 타당하고 신뢰할 수 있는 방식으로 측정 할 수 있는 설문 문항을 개발하는 것을 목표로 한다. 예비 단계에서 전문가 5 인의 내용 타당도 점검을 거쳐 ‘사려 깊은 행동’, ‘내용·과정 성찰’, ‘전체 성 찰’의 3구인을 설정하고, 총 27개 예비 검사 문항을 작성하였다. 탐색 단계 에서 4개 고등학교 467명의 1~2학년 학생이 예비 문항에 응답하게 하고 그 결과를 수집하였다. 확인 단계에서 2회에 걸쳐 총 11개 고등학교 1,055명의 1~2학년 학생이 선별된 12개 본 검사 문항에 응답하게 하고 그 결과를 수집 하였다. 탐색 단계에서 수집한 응답 결과에 대해 탐색적 요인 분석(최대우도 법, 프로맥스 회전)을 실시한 결과, 27개 예비 검사 문항 중 12개 문항이 공통 분 .30 이상, 가정된 구인에 대한 요인 부하량 .40 이상, 가정되지 않은 구인 에 대한 요인 부하량 .30 미만의 조건을 충족하는 것이 확인되었다. 또한 3요 인 모형은 2요인 모형과 달리 요인 수 가정이 기각되지 않았으며($\chi^2=45.09$, $df=33$, $p=.08$) 4요인 모형에 비해 적정한 TLI 값(.979)을 보이는 것이 확인 되었다. 확인 단계에서 최초 수집한 응답 결과(568명)에 대해 확인적 요인 분석을 실시한 결과, 2개 문항의 신뢰도가 양호하지 않고 모형 합치도 지 수 중 NFI(.849)와 CFI(.879)가 기준(.900)에 미달하는 것이 확인되었다. 이에 신뢰도에 문제가 있었던 2개 문항의 진술 방식을 수정한 후 재수집한 응 답 결과(487명)에 대해 확인적 요인 분석을 실시한 결과, 모든 문항의 신뢰 도가 양호하며 NFI(.897)가 기준에 다소 미달할 뿐, CFI(.921), GFI(.936),

RMR(.037), RMSEA(.076)가 모두 기준을 충족함을 확인하였다. 결론에서는 본 연구에서 개발한 설문 문항의 측정 결과와 관련한 시사점은 물론, 본 설문 문항을 사용하여 수행 가능한 후속 연구를 제안하였다.

핵심어 대화 성찰 태도, 사려 깊은 행동, 내용 성찰, 과정 성찰, 전제 성찰, 비판적 성찰, 성찰적 대화 교육, 화법 교육

ABSTRACT

Development of a Questionnaire to Measure Reflective Attitude in/on Conversation

Min Byeonggon · Park Seongsoeg

This study aims to define the concept of *Reflective Attitude in/on Conversation* (RAC), based on a review of discussions in the field of reflective education, and develop a valid and reliable questionnaire to measure RAC. In the preliminary stage, three constructs, namely, *Thoughtful Action* (TA), *Content/Process Reflection* (CPR), and *Premise Reflection* (PR), were established, and content validity was verified by five specialists. In the exploratory stage, 467 students, who were in their 1st or 2nd year of school, were sampled from four high schools. In the confirmatory stage, 1,055 students who were in their 1st or 2nd year of school were sampled from 11 high schools across our two sessions. The results of exploratory factor analysis show that 12 among 27 preliminary items have communality greater than 0.3, factor loading to the hypothesized construct greater than 0.4, and factor loading to un-hypothesized constructs lower than 0.3. Further, the 3-factor-model's Qui-square test does not reject the null hypothesis that there are only three factors, while the 2-factor-model's Qui-square test rejects the null hypothesis that there are only two factors. The 3-factor-model also shows suitable TLI value (.979), while the 4-factor-model revealed a highly saturated TLI value. The results of confirmatory factor analysis with our first confirmatory stage samples (568 students), show that two items among 12 are not reliable, and NFI (.849) and CFI (.879), among five model fitness indices, do not satisfy their criteria ($>.900$). After revising the two unreliable items, the results of confirmatory factor analysis with our second confirmatory stage samples (487 students), reveal that all the 12 items are reliable, and although NFI (.897) still does not satisfy the criterion, CFI (.912), GFI (.936),

RMR (.037), and RMSEA (.076) satisfy their criteria. In the conclusion, we suggest some implications of measured results, and some further research directions.

KEYWORDS Reflective Attitude in/on Conversation, Thoughtful Action, Content Reflection, Process Reflection, Premise Reflection, Critical Reflection, Reflection, Reflective Conversation Education, Oracy Education

‘고등학생의 일상 대화에서의 성찰 태도’ 조사를 위한 설문지

♣ 일상 대화에서의 성찰 태도에 관한 다음 문항들을 읽고 자신에게 해당하는 곳에 V표를 해 주십시오. 정답이 있는 물음이 아니므로 솔직하게 응답해 주시기 바랍니다.

1: 전혀 그렇지 않다. 2: 그렇지 않다. 3: 보통이다. 4: 그렇다. 5: 매우 그렇다

문항	1	2	3	4	5
1. 대화에서 어떻게 말해야 할지를 고민하는 것은 중요하다고 생각한다.					
2. 다양한 대화방식 중 나에게 맞는 최선의 대화방법을 찾기 위해 노력한다.					
3. 어떤 대화가 좋은 대화인지에 관한 나의 믿음이나 생각에 문제가 없는지 고민한다.					
4. 대화를 나눈 후 나의 생각과 감정이 상대에게 왜곡 없이 잘 전달되었는지 되돌아 본다.					
5. 대화가 나의 예상과 달리 전개되거나 문제가 발생했을 때는 더욱 조심해서 말하려고 한다.					
6. 생각 없이 유행을 따라 잘못된 표현을 사용하고 있지 않은지 반성한다.					
7. 대화를 더 잘하기 위해 끊임없이 새로운 방법을 시도해 본다.					
8. 나의 말이 상대의 감정에 미칠 영향을 생각하며 말할 내용을 신중히 고민한다.					
9. 누군가의 잘못된 대화 방식을 무의식적으로 따라하고 있지 않은지 돌아본다.					
10 평소에 대화에서 아무렇게나 생각나는 대로 말하더라도 괜찮다고 생각한다.					
11. 나의 평소 대화 습관이 바람직한지를 생각해 본다.					
12. 대화의 흐름이나 상대의 반응에 따라 최선의 대화 방식이 무엇일지 생각하며 대화한다.					

■ 구인별 결과 계산 방법

- 사례 깊은 행동 = {1번+5번+8번+(6-10번)}/4
- 내용·과정 성찰 = (2번+4번+7번+12번)/4
- 전제 성찰 = (3번+6번+9번+11번)/4