

## 중 · 고등학생들의 비형식적 통계적 추리의 수준 연구

이 정 연\* · 이 경 회\*\*

통계교육 연구자들은 형식적 추리 방법을 지도하기에 앞서 비형식적 추리를 지도할 것을 강조하며 통계적 추리의 발달 과정에 주목하고 있다. 본 연구는 표본 비교하기 과제와 모집단의 그래프 추측하기 과제를 해결하는 과정에서 나타나는 중 · 고등학생들의 비형식적 통계적 추리의 수준과 각 수준별 특징을 분석하였다. 연구 결과, 표본 비교하기 과제에서는 개인적인 의견에 기초하여 타당하지 않은 추리를 하는 수준, 자료에 대한 국소적 관점을 가진 수준, 자료에 대한 전체적 관점으로 전환되는 수준, 분포의 다각적인 측면에 주목하는 수준, 통계적 개념들을 통합하여 추리하는 수준이 확인되었다. 모집단의 그래프 추측하기 과제에서는 개인적인 의견에 기초하여 타당하지 않은 추리를 하는 수준, 표본대표성에만 주목하고 표집변이성을 고려하지 않는 수준, 표본대표성과 표집변이성을 모두 고려하며 분포의 한 측면에 주목하여 부분적으로 타당한 추리를 하는 수준, 분포의 다각적 측면에 주목하는 수준, 통계적 개념들을 통합하여 추리하는 수준이 확인되었다.

### 1. 서론

형식적인 추측통계의 이해가 어렵기 때문에 통계교육 연구자들은 비형식적인 방법으로 추측통계를 도입할 것을 제안하였다(Makar & Rubin, 2009; Pfannkuch, 2006; Park, 2015). 연구자들은 학생들이 비형식적 추리를 일찍부터 발달시키기 시작한다면 추측통계의 형식적 방법들을 더 잘 학습할 수 있음을 확인하였다.

비형식적 추리에 대한 연구는 비형식적 추리의 정의와 하위 요소를 분석한 연구(Makar & Rubin, 2009; Pfannkuch, 2006; Zieffler et al., 2008), 비형식적 추리의 발달을 다룬 연구(Huey, 2011; Makar, 2016), 비형식적 추리를 지도하고 평가하기 위한 과제에 대한 연구(Goss, 2014; Madden,

2011)로 이루어져왔다. 최근에는 비형식적 추리와 형식적 추리의 단절이 여전히 존재한다는 지적에 따라 비형식적 추리와 형식적 추리의 관계에 대한 연구(Jacob, 2013)가 이루어지고 있다. NGA Center와 CCSSO(2010)는 고등학교에서 형식적 추리 방법을 지도하기에 앞서 중학교에서 비형식적 추리를 지도할 것을 요구함으로써 수학과 교육과정에서 통계적 추리의 발달에 특히 주목하고 있다.

국외에서 비형식적 추리의 중요성을 인식하고 이에 대한 연구와 지도가 활발히 이루어지고 있는데 반해 국내에서는 대푯값, 변이, 표본, 표집 등의 개별적인 개념에 대한 이해를 다룬 연구가 주로 이루어져왔다. 통계적 추리를 다룬 연구(박민선 외, 2011; Park, 2015)에서는 자료집합 비교 활동의 유용성이 강조되었고 비형식적 추리의

\* 성수고등학교, khskme@hanmail.net (제1 저자)  
\*\* 서울대학교, khmath@snu.ac.kr (교신저자)

평가 모델이 제안되었다. 우리나라에서는 비형식적 추리를 다루지 않고 고등학교 ‘확률과 통계’에서 통계적 추리를 형식적으로 도입하여 지도하고 있다. 따라서 통계적 추리의 발달에 도움이 될 수 있는 학습지도 방법 및 자료의 개발이 필요하며 이를 위해서는 우선 학생들의 비형식적 추리의 특징을 이해할 필요가 있다.

본 연구에서는 비형식적 통계적 추리 과제를 해결하는 과정에서 나타나는 중·고등학생들의 비형식적 통계적 추리의 수준이 어떠한지, 각 수준별 특징이 어떠한지를 분석한다. 학생들의 비형식적 추리의 특징을 범주화하고 수준을 구분하는 것은 학습 위계에 대한 정보를 제공한다는 점에서 추측통계 지도에 대한 시사점을 제시하였다.

## II. 선행 연구

비형식적 통계적 추리의 정의는 연구자들마다 조금씩 다르다. Zieffler 외(2008, p. 44)는 관찰한 표본에 기초하여 알려지지 않은 모집단에 대한 추정을 정당화하기 위해 비형식적 통계 지식을 사용하는 방식을 비형식적 통계적 추리라고 정의하였다. Makar와 Rubin(2009, p. 85)은 비형식적 통계적 추리를 자료의 정보를 이용하여 일반화하고 그것을 검증하는 추론 과정이지만 비형식적인 과정, 즉 반드시 표준적인 통계적 절차를 따를 필요는 없는 추론 과정이라고 설명하였다. 이와 같이 정의의 세부적인 차이는 있지만, 표본으로부터 모집단에 대한 정보를 도출한다는 것과 이를 논증하는 데에 비형식적 통계적 지식을 사용한다는 것이 공통적으로 강조되고 있다.

Makar와 Rubin(2009, p. 85)은 비형식적 통계적

추리의 본질적인 세 가지 원리로, 주어진 자료를 기술하는 것을 넘어서 예측하고 모수를 추정하고 결론을 이끌어내는 것을 포함한 일반화하기, 자료를 일반화에 대한 증거로 사용하기, 이끌어낸 결론의 확실성의 정도에 대해 비형식적 표현을 포함하여 일반화를 기술하는 데에 확률 언어를 사용하기를 제시하였다. Zieffler 외(2008, p. 52)는 비형식적 추리의 연구에 사용할 수 있는 세 가지 유형의 과제로서, 표본을 토대로 모집단에 대해 추정하거나 그래프 그리는 과제, 모집단 간에 차이가 있는지를 추리하기 위해 두 개 이상의 표본을 비교하는 과제, 두 개의 대립되는 모델이나 설명 중에서 어떤 것이 더 옳은지를 판단하는 과제를 제안하였다.

박민선 외(2011)는 자료집합 비교 활동에서 중학생들이 자료값, 비교기준, 퍼짐, 표본과 같은 통계적 요소에 주목함을 확인하고 자료집합 비교 활동이 형식적 추리를 위한 토대가 될 수 있음을 확인하였다. Ciancetta(2007)는 분포 비교활동에서 나타나는 대학생들의 추론 수준을 주관적, 국소적, 과도기적, 초기분포적, 분포적 수준의 5수준으로 제시하였다. 과제의 맥락에만 기초하고 자료에 주목하지 않는 반응을 주관적 수준으로, 절대도수 또는 개별적인 자료값들의 비교에 기초한 반응을 국소적 수준으로 구분하였다. 또한 자료에 대한 전체적 관점을 가지고 있지만 자료의 다양한 특징을 완전히 통합하지 못하는 반응을 과도기적 수준으로 구분하고, 과도기적 수준을 분포의 어떠한 측면에 주목하느냐에 따라 형태, 중심, 변이의 세 종류로 구분하였다<sup>1)</sup>. 초기분포적 수준은 특정값을 기준으로 위나 아래의 도수를 비교하는 비율적 수준과 전체적인 조망을 하지만 분포의 다각적 측면을 세련되게 통합시키지는 못하는 초기전체적 수준으로 구분

1) 한국통계학회에서는 Variation을 ‘변동’, ‘변동성’이라는 용어로 번역하여 사용하지만, 국내수학교육연구에서는 변화의 경향을 지닌 다양한 현상의 관찰 가능한 실체의 특징과 이러한 특징을 기술하거나 측정하는 것이라는 의미에서 ‘변이’, ‘변이성’이라는 용어로 번역하여 사용하고 있다.

되었다. 자료집합에 대해 전체적으로 조망하고 형태, 중심, 변이, 비율과 같은 측면들 중 적어도 두 개를 조합하고 통합하여 추론하는 반응을 분포적 수준으로 구분하였다.

Goss(2014)는 비형식적 통계적 추리의 수준을 평가할 수 있는 과제의 예시들을 제시하고 SOLO 모델<sup>2)</sup>에 근거하여 과제에 대한 반응의 수준을 분석하였다. Huey(2011)는 통계 수업에서 예비교사들의 통계적 추리의 발달을 확인하였으며, SOLO 모델을 토대로 비형식적 통계적 추리의 수준을 전구조적 수준, 단일구조적 수준, 다중구조적 수준, 관계적 수준으로 구분하였다. Huey(2011)에 따르면, 자료집합을 비교할 때 전구조적 수준에서는 자료나 맥락에 기초하지 않은 추리를 하고, 단일구조적 수준에서는 자료에 기초하여 추리하지만 한 측면을 정확하게 비교하거나 몇 가지 측면들을 부분적으로 정확하게 비교한다. 다중구조적 수준에서는 다각적인 측면들을 정확하게 비교하고, 자료와 맥락에 기초하여 부분적으로 타당한 추리를 한다. 관계적 수준에서는 다각적인 측면들을 정확하게 비교하고, 자료와 맥락에 기초하여 타당한 추리를 한다. 이와 같은 수준 구분은 비형식적 통계적 추리에 대한 평가를 용이하게 해주며 교육과정과 학습 지도 설계에 중요한 역할을 한다.

### III. 연구 방법

#### 1. 연구 참여자

본 연구의 참여자는 서울특별시에 소재한 중학교 2학년 학생 11명과 고등학교 2학년 학생 86명이다. 실험 과제를 해결하기 위해서는 통계

그래프를 그리고 해석할 수 있어야 하므로 중학교 1학년의 통계 단원을 학습한 학생들이 참여자로 적합하였다. 연구 목적에 완전히 부합되는 다양한 수준의 참여자 집단을 선정하는 것이 현실적으로 어려웠기 때문에 중학교 2학년과 고등학교 2학년 학생들을 연구 대상으로 선정하였다. 참여 학생들은 수학 성취도와 표현 수준이 다양하여 비형식적 추리 활동에서 다양한 반응과 수준을 확인하는데 유용할 것으로 판단하였다.

#### 2. 실험 과제

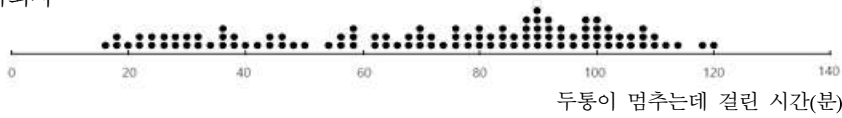
연구에 참여한 중학생을 대상으로 교수 실험을 실시하였고, 고등학생을 대상으로 설문조사를 실시하였다. 실험 과제로 ‘표본 비교하기’ 과제와 ‘모집단의 그래프 추측하기’ 과제를 사용하였다. 실험에 사용한 과제는 [그림 III-1], [그림 III-2]와 같다. 과제 1-1, 과제 2는 각각 Bright 외(2003), Bakker(2004)의 연구에서 사용한 과제를 수정한 것이고, 과제 1-2는 본 연구에서 개발한 과제이다. 과제 1은 서로 독립인 표본들을 이용하여 모집단들을 비교하는 과제로 두통약이라는 현실적인 맥락을 제공하고 하나의 두통약을 추천할 것을 요구하였다. 주어진 조건을 어떻게 해석하고 어떠한 근거에 기초하여 표본을 선택하는지를 알아보기 위한 과제였다. 과제 2는 표본에 기초하여 모집단의 그래프를 추측하여 그리는 과제이다. 이 과제는 표본의 크기가 변화하는 과정의 안정적인 특징을 추론하는지를 알아보기 위한 과제였다. 설문조사에 사용한 과제는 과제 1은 같았으나 과제 2의 경우는 한 학급과 세 학급의 그래프를 모두 주고 모집단의 그래프를 추측하여 그리는 과제로 수정하여 사용하였다.

2) Biggs와 Collis(1991)의 SOLO(Structure of the Observed Learning Outcome) 모델은 학생들의 반응 양식에 초점을 두고 있으며 학습자의 발달 수준 및 학습주기 파악에 유용하기 때문에 통계적 추론 발달을 연구하는 데에 폭넓게 사용되고 있다.

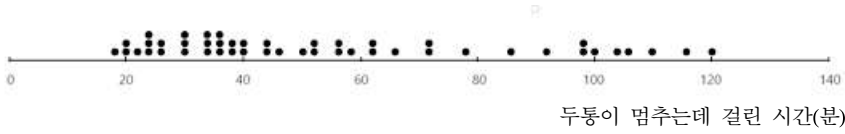
### 3. 자료 수집 및 분석

[과제 1] A 제약회사와 B 제약회사는 새로운 두통약을 개발했습니다. 두 제약회사에서는 새로운 두통약의 효과를 알아보기 위해서 환자들에게 실험을 하였습니다. 실험에 참여한 환자들은 두통을 느끼자마자 바로 약을 먹고 나서 두통이 멈출 때까지 걸린 시간을 기록했습니다. 두 제약회사에서는 이 자료를 가지고 다음과 같은 그래프를 만들었습니다.

A 제약회사

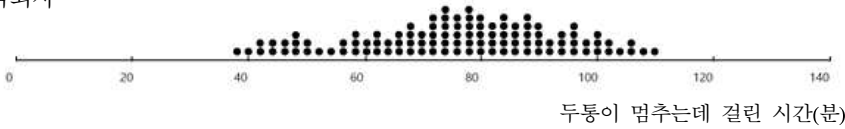


B 제약회사



1. 두통이 심한 친구가 있다면 어느 회사의 두통약을 추천하겠습니까? ( ) 제약회사 그 이유를 적어봅시다.
2. C 제약회사에서도 새로운 두통약을 개발하고 실험을 한 결과를 가지고 다음과 같은 그래프를 만들었습니다.

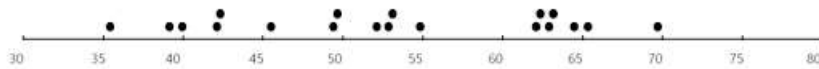
C 제약회사



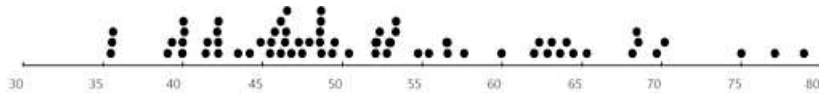
- A, B, C 세 회사의 두통약 중에서 어떤 약을 추천하겠습니까? ( ) 제약회사 그 이유를 적어봅시다.

[그림 III-1] 표본 비교하기 과제

[과제 2] 민지는 서울시에 있는 중학교 2학년 학생입니다. 민지는 같은 반 친구들 19명의 몸무게를 조사해서 다음과 같은 그래프를 만들었습니다.



1. 민지가 같은 학년 세 학급의 학생들의 몸무게를 조사해서 그래프를 만든다면 그래프의 모양이 어떻게 될까요? 그래프를 추측해서 그려봅시다.
2. 다음은 민지가 실제로 같은 학년 세 학급의 학생들의 몸무게를 조사해서 만든 그래프입니다.



위 그래프를 보고, 서울시에 있는 중학교 2학년 학생 1000명의 몸무게를 나타내는 그래프를 추측해서 그려봅시다. 위와 같이 그린 이유를 적어봅시다.

[그림 III-2] 모집단의 그래프 추측하기 과제

2017년 5월에 교수 실험을 실시하였다. 한 차시에 45분씩, 두 차시 동안 과제를 순서대로 다루었다. 학생들에게 개별적으로 활동지를 해결하게 한 후에 전체적으로 토론하도록 수업을 진행하였다. 모든 학생들이 의견을 발표할 수 있게 하였고, 다른 친구의 의견에 대해 보충 설명을 하거나 의견이 다르다면 반박할 것을 요구하였다. 주장에 대한 근거를 제시할 수 있으면 어떠한 주장도 허용될 수 있음을 강조하였다. 교사는 주로 학생의 설명을 다른 학생들이 이해할 수 있도록 요약하여 재진술하는 역할을 했으며 연구자가 직접 수업을 진행하였다. 수업은 녹화·녹음되었고 분석을 위해 모두 전사하였다. 같은 시기에 고등학생들을 대상으로 설문조사를 실시하였다. 고등학생의 활동지를 토대로 학생들의 반응 특징을 예비적으로 범주화하여, 각 범주를 대표하는 전형적인 반응을 보인 학생과 어느 범주에 속하는지 구별이 어려운 학생들을 면담 대상으로 선정하였다. 면담 과정을 모두 녹음한 후 전사하였다.

연구의 자료는 학생 활동지, 영상 파일, 음성 파일이다. 자료 분석의 첫 번째 단계에서 학생들의 표현을 가장 잘 설명해주는 용어를 중심으로 약호화 하였다. 약호화 과정은 철저히 학생들의 반응을 토대로 자료에 근거하는 귀납적 접근법을 사용하였다. 두 번째 단계에서는 자료의 내용을 전체적으로 포섭하여 설명할 수 있는 의미와 범주를 생성하였다. 선행연구를 토대로 범주들을 위계화하여 비형식적 추리의 수준을 구분하고자 하였다. 자료분석의 타당도와 신뢰도를 위해 삼각검증법과 동료검토의 방법을 활용하였다. 삼각검증법은 결과를 분석하기 위해 연구자 2인, 통계교육연구자 1인 등 여러 명의 연구자가 교차 검증하는 방식으로 이루어졌으며, 동료검토는 자료분석 결과에 대한 통계교육연구자 1인, 현직교

사 1인의 조언과 평가를 구하는 방식으로 이루어졌다.

## IV. 연구 결과

### 1. 표본 비교하기 과제

#### 가. 비형식적 추리의 특징 분석

표본 비교하기 과제에서 학생들의 반응은 개인적인 의견에 기초하여 타당하지 않은 추리를 하는 경우, 개별적인 자료값에 주목하는 경우, 분포의 한 측면에 주목하는 경우, 분포의 다각적인 측면에 주목하는 경우로 특징지을 수 있었다.<sup>3)</sup> <표 IV-1>은 표본 비교하기 과제에서 각 수준에 속하는 학생들의 반응 특징을 요약한 것이다.

0수준의 학생들은 자료나 맥락에 기초하기 보다는 개인적 의견에 기초하여 판단하였다. M1과 H22는 두 표본의 크기가 같을 것이라는 추측으로부터 140분 이후의 자료가 존재하지만 기록되지 않았다는 주관적인 판단을 하였다. 이들은 그래프에 나타나지 않은 자료값이 있다고 가정하고 접근하였다. 이는 비교실험한 표본의 대표성을 의심하고 표본을 부정하는 언급이기도 하다. H72와 같이 주어진 자료를 전혀 고려하지 않은 학생도 있었다.

148 M1: 제 생각에는 A 제약회사랑 B 제약회사에서 똑같은 사람을, 똑같은 사람 명수를 토대로 약을 먹게 하였는데, A 제약회사는 이만큼 나왔는데 B 제약회사는 효과가 없거나 아니면 140분 이상 걸려서 표에 기록이 안됐다고 생각을 했어요.

3) 연구에 참여한 중학생은 M1~M11, 고등학생은 H1~H86으로 표시하였다.

실험을 하기 위해 같은 수의 환자를 대상으로 조사를 했을 텐데 140분까지의 수에서 B 제약회사가 더 적은 것으로 보아 그 이후까지 두통이 지속된 환자가 존재할 것이다. 따라서 120분까지 두통이 멈춘 A 회사의 두통약을 추천한다.

[그림 IV-1] 과제 1-1에서 H22의 반응

각 환자에 맞는 두통약이 있는데, 내가 그 환자에게 더 맞는 약을 결정해주기에는 지식이 부족하다. 다만, 이미 A나 B 둘 중 하나를 복용하여 효과가 없는 것을 확인한 친구의 경우에는, 친구가 복용하지 않은 회사의 제품을 추천하겠다.

[그림 IV-2] 과제 1-1에서 H72의 반응

1수준의 학생들은 개별적인 자료값을 비교하거나 절대도수에 기초하여 추리하였다. 이들은 자료집합을 개별적인 자료값들의 합으로 인식하는 자료에 대한 국소적인 관점을 가지고 있었다. 다음은 1수준에 속하는 학생 반응의 예이다.

M11은 0분에서 60분 사이의 자료값의 개수를 비교하여 절대도수가 더 큰 A를 선택하였다. M6은 60보다 작은 자료값에 대해서는 절대도수를 비교하고 60보다 큰 자료값에 대해서는 최댓값을 비교하였다. 이와 같은 가법적 사고는 크

가 같은 표본을 비교할 때에는 적절하지만 과제 1과 같이 크기가 다른 표본을 비교할 때에는 부적절하다. H51은 A를 선택한 이유로 표본의 크기가 크면 '신뢰도'가 높다는 것을 제시하였다. 이는 표본의 크기에 대해서만 인식하는 0수준과는 달리 표본의 크기의 영향력을 인식하고 통계적으로 해석한 반응이다. 그러나 중심, 변이, 형태에 주목하지 않고 자료값의 개수에만 주목하였으므로 1수준으로 분류되었다.

A 제약회사랑 B 제약회사의 그릇에서 0분에서 60분일 때 A 제약회사는 3명, B 제약회사는 3명이 나왔고 그래서 A 제약회사가 더 좋은 것 같다.

[그림 IV-3] 과제 1-1에서 M11의 반응

B 제약회사와 C 제약회사 둘 다 60을 기준으로 사람 수를 세었는데 C 제약회사가 더 적었다. 그리고 60분 이후도 비교를 해보았는데 B 제약회사는 120분 가까이나 120분도 꽤 있었지만 C 제약회사는 120분 가까이 있는어도 120분은 없어서 C 제약회사를 선택하였다.

[그림 IV-4] 과제 1-2에서 M6의 반응

많은 사람들을 실험할 경우, 결과에 대한 신뢰도가 높아질 수 있다.

[그림 IV-5] 과제 1-1에서 H51의 반응

<표 IV-1> 표본 비교하기에 대한 비형식적 추리 수준

수준	각 수준에서의 특징
0	자료나 맥락에 기초하기 보다는 개인적인 의견이나 경험에 기초하여 타당하지 않은 추리를 함.
1	개별적인 자료값의 비교 또는 절대도수에 기초하여 비교함. 자료집합을 개별적인 자료값들의 합으로 인식하며 자료에 대한 국소적 관점을 가짐.
2	중심, 변이, 형태와 같은 분포의 여러 측면들 중 한 측면에 주목하여 비교함. 자료에 대한 국소적 관점에서 전체적 관점으로 전환됨. 자료와 맥락에 기초하여 추리함.
3	분포의 다각적인 측면들에 주목하여 비교하고 비올적 추론을 사용함. 자료에 대한 전체적인 관점을 가지지만 통계적 개념들을 통합하지 못함. 자료와 맥락에 기초하여 추리함.
4	분포의 다각적인 측면들에 주목하고 통계적 개념들을 통합하여 추리함. 자료에 대한 전체적인 관점을 가지며 자료와 맥락에 기초하여 추리함.

처음에는 0수준을 나타냈지만 토론을 거치면서 1수준으로 향상된 학생들도 있었다. M7은 개인적인 추측에 근거해서 주어진 표본을 신뢰하지 않는 관점에서 자료값에 주목하는 것으로 관점의 변화를 보여주었다. M7이 수업 초반부터 그래프에 140분 이후의 기록이 누락되었다는 생각을 지속적으로 가지고 있었기 때문에 좀 더 자료에 주목할 수 있도록 교사가 중재하였다(line 235). M7은 초반에는 개인적 의견에 기초한 추리를 하였으나 토론 과정을 거치면서 60분 이내의 절대도수를 비교하는 관점으로 전환하였고, 승법적 사고를 나타낸 M3의 의견을 듣고 나서도 여전히 절대도수에 기초한 판단을 고수하였다(line 236, 240).

- 235 교사: 음. B는 이게 인원이 적은데 140 이후에 넘어가는 뭔가를 이 사람들이 숨겼다고 생각을 했기 때문에 그런 생각을 했는데 만약에 그러면 이 데이터가 전부라면 그러면 어떻게 생각해요?
- 236 M7: 그냥 비교를 해봐도요. 저기 앞부분부터 저기서 (칠판에 적힌 M3의 설명을 가리키며) 31이랑 36정도 써줬잖아요. 이렇게...
- 240 M7: 아 그러면 60분, 1시간 전으로 B 제약회사랑 A 제약회사를 확인해보면요. A 제약회사가요 B 제약회사보다 이게 효과가 느린 게 아니에요 보면. 오히려 B 제약회사보다 60분 이내로 A 제약회사가 좀 더 많은 환자가 더 일찍 두통이 멈췄어요.

2수준의 학생들은 중심, 변이, 형태와 같은 분포의 여러 측면들 중 한 측면에 주목하여 비교하였다. 이들은 국소적 관점에서 벗어나서 자료를 전체적으로 조망하였는데 어떠한 개념에 주목하여 비교하느냐에 따라 학생들의 반응은 2-

중심, 2-변이, 2-형태로 구분되었다.

최빈값의 위치를 언급하거나 자료값들이 많은 부분에 기초하여 비교하는 반응을 2-중심 수준으로 분류하였다. H32는 ‘가장 많이’, ‘대부분’이라는 용어를 사용하여 최빈값에 주목하였고, M3은 ‘평균적으로’라는 용어를 사용하여 평균이 포함된 구간에 주목하여 비교하였다. 많은 학생들이 중심이 포함된 구간을 비교하였는데, 평균값이 아닌 평균값을 포함한 구간으로 비교하는 것은 모평균의 가설검정의 관점에서도 의미 있는 추리활동이라고 할 수 있다.

A는 두통이 90분 정도에 가장 많이 멈추지만 B는 20-40분 사이에 대부분 두통이 멈춘다. 일찍 두통이 멈추는 B가 나은 것 같다.

[그림 IV-6] 과제 1-1에서 H32의 반응

A 제약회사는 평균적으로 60분에서 100분 사이에 분포되어 있고, B 제약회사는 20분에서 60분 사이에 분포되어 있기 때문에 빨리 두통이 가라앉을 가능성이 큰 B 제약회사의 두통약을 추천할 것이다.

[그림 IV-7] 과제 1-1에서 M3의 반응

범위, 퍼짐, 양끝 값을 언급하며 변이에 주목하여 비교한 반응은 2-변이 수준으로 분류하였다. H3은 ‘범위가 더 좁다’는 것에서 ‘약의 효과가 안정적’이라고 추론하였고, H18은 ‘격차’라는 용어를 사용하여 ‘격차가 적은’ 것을 ‘효과가 확실한’ 것으로 해석하였다. 두 학생은 분포의 범위에 주목한 것이다. M2는 분산을 언급하며 ‘사람들에 따라서 차이가 크다’는 것에서 ‘약의 성능’에 대해 추론하였다(line 232). 이는 퍼짐이 작으면 분포가 안정적이라는 인식을 보여준다. 또한 M2는 두 분포의 최댓값의 차이와 최솟값의 차이를 계산하여 양끝 값에 주목하여 비교하였다(line 309).

B가 더 효과를 빨리 일으킨 여지가 있지만 효과가 나타나는데  
까지의 시간 범위가 더 좁기 때문에 C가 A, B, C 중 가장 안정적으로 효과를  
나타내는 약일 것이다.

[그림 IV-8] 과제 1-2에서 H3의 반응

A, B 제약 회사의 제품은 100% 1/20초까지 서둘러  
효과가 큰 편이지만 C 회사의 제품은 100% 1/20초까지  
효과가 확실 하지만 성장이 느린 편.

[그림 IV-9] 과제 1-2에서 H18의 반응

232 M2: 근데, A 제약회사가, 약이 멈추는데  
걸린 시간이 되게 몰려있는데, B 제  
약회사는 너무 분산돼 있잖아요.  
(중략) 불규칙해서 제가 생각한 거  
는 사람들에 따라서 너무 차이가  
크니까 A 회사 약이 더 성능이 좋  
은 거 아니에요?

309 M2: 이게 A가 가장 늦게 멈춘 시간이  
120분이고 C가 110분이라고 했는  
데.. 그 차이가 10분밖에 안 났잖아  
요. 근데 A 제약회사가 가장 빨리  
멈췄을 때는 20분이나 차이가 나요.  
그래서 저는 A 제약회사가 더 낫다  
고 생각해요.

분포의 형태에 주목하여 비교한 반응은 2-형태  
수준으로 분류하였다. M5가 표본 A의 그래프에  
서 ‘뒤로 갔을 때 사람이 더 많다’는 것은 자료  
의 비대칭성의 형태를 인식한 것으로 볼 수 있  
고, 표본 C의 그래프에서 ‘앞이나 뒤에 쳐져 있  
지 않고 중간에 모여 있다’는 것은 형태의 정규  
성을 인식한 것으로 볼 수 있다(line 296). 중학  
생이기 때문에 ‘편포’, ‘정규성’과 같은 용어를  
사용하지는 않았으나 M5는 비형식적으로 분포  
의 형태를 설명하며 표본을 비교하였다. M8은  
‘넓게 분포’, ‘밀집’라는 용어를 사용하여 밀도를  
표현하였고, H40은 ‘모양’을 언급하여 형태에 주  
목하고 있음을 나타내었다.

296 M5: 왜냐하면 A의 약 효과를 봤을 때  
뒤로 갔을 때 사람이 더 많단 말이  
예요. (중략) C 제약회사를 보면 앞  
이나 뒤에 쳐져 있지 않고 중간값,  
중간에 모여 있단 말이에요.

A회사의 B회사는 그래프가 낮게 뻗되어 있는데, C회사는 그래프가  
비정적 뻗되어 있다. 확실한 약 효과를 원한다면 C 제약회사의  
약을 추천하고 싶다.

[그림 IV-10] 과제 1-2에서 M8의 반응

A와 C회사의 두통이 너무 화상과 비슷한 반면  
C회사의 더 민감하다. 분포가 약간씩 모양이 조금  
달라야 알겠지,  
약품이 왜병이다.

[그림 IV-11] 과제 1-2에서 H40의 반응

3수준의 학생들은 분포의 다각적인 측면들에  
주목하여 비교하고 자료에 대한 전체적인 관점  
을 가졌지만 통계적 개념들을 통합하지는 못하  
였다. 이들은 주로 비율적 추론을 사용하였다.  
다음은 3수준에 속하는 학생들 반응의 예이다.

119 M3: B 제약회사가 47명이고 A 제약회  
사가 106명이어서 이걸 60분을 기  
준으로 해서 나눠봤을 때, A 제약  
회사는 36:70이고 B 제약회사는  
31:16인데, 이것을 B 제약회사를  
106명으로 해 보면은 60분을 기  
준으로 71이랑 35로 나눌 수가 있어  
요. (중략)

17 H27: A하고 B를 비교했을 때, A는 여기랑  
여기 인원수는 비슷할지 모르겠지만.  
전체 비율로 봤을 때는 20에서 60이  
60에서 120의 비율보다 더 적게 나타  
나서 오래 걸렸다고 생각, 두통이 주  
는 데까지 걸린 시간이 오래 걸렸다고  
생각했고, 여기는 20에서 60 비율  
이 나머지 비율보다 더 많아가지고  
더 적게 걸렸다고 생각을 했어요.



M3은 두 표본의 크기가 다르다는 것을 인식하고 60을 기준으로 60 전후의 자료값의 개수를 표본의 크기에 비례하여 조정하였다(line 119). H27은 60보다 작은 자료값의 도수와 60보다 큰 자료값의 도수의 비율을 이용하여 비교하였다(line 17). 한편, M4는 60보다 작은 범위의 자료값의 개수와 표본의 크기의 비율인 상대도수를 이용하였다. 이들은 모두 어떤 구간의 절대도수를 다른 구간의 절대도수와 비교하거나 전체 자료집합의 크기와 비교하는 비율적 추론을 사용하였는데 이는 자료를 전체로서 파악하고 있음을 보여준다.

A 제약회사의 실험인원은 106명이고, B 제약회사의 실험인원은 47명이다. 그러나 A, B 제약회사 모두 두통이 멈추기까지 120분, 즉 2시간이 걸렸다. 120분의 절반인 60분 이내의 인원을 보았을 때, A 제약회사는 106명 중 36명이었고, B 제약회사는 47명 중 31명이 두통이 멈췄다. A 제약회사의 실험인원의 절반의 수인 53명보다 적었다. 그래서 B를 추천한다.

[그림 IV-12] 과제 1-1에서 M4의 반응

4수준의 학생들은 분포의 다각적인 측면들에 주목하여 비교하고 통계적 개념들을 통합하여 추리하였으며 자료에 대한 전체적인 관점을 가지고 있었다. 다음은 4수준에 속하는 학생 반응의 예이다.

- 48 H7: (20부분의 눈금을 가리키며) 여기 두 칸이 있었고, (100 부분의 눈금을 가리키며) 여기 두 칸이 있었으니까. 인원수가 (A를 가리키며) 애가 훨씬 많았잖아요. 애가 인원수가 훨씬 많았는데. 작용을 기준으로 봤을 때 빠른 작용이 많고 느린 작용이 적은 B 제약회사의 제품이 표본은 좀 더 적었지만 좀 더 효과가 있을 거라고 생각을 했고. 또 다르게 봤을 때는 극값의 환자들이 좀 약이 잘 받는 그런 사람들일 수도 있잖아요. 이 약만 안 받는 사람이었다든가. 특이한 체질이니까 극값들을 다 제외하고 봐도 B 제약회사의 부분이, 가운데 그래프를 보면 효과가 조금 더 빠르다. (양 끝을 제외하는 손동작을 하며) 이렇게 보면은 빠르다.
- 52 H7: 왜냐하면 20하고 100을 제외하고 봤으니까. (A를 가리키며) 애는 이렇게 뒤로 가있고 (B를 가리키며) 애는 이렇게 앞으로 가있잖아요.

H7은 그래프의 눈금에서 20과 100을 기준으로 분포의 앞부분, 뒷부분, 가운데 부분으로 나누어 비교하였다. 앞부분과 뒷부분에 대해서는 부분과 전체의 도수를 비교하는 비율적 추론을 하였고 (line 48), 극단적인 경우를 제외하기 위해 가운데 부분에 주목하여 ‘뒤로 가있고’, ‘앞으로 가있고’라는 표현을 하며 분포의 형태에도 주목하였다(line 52). 20과 100을 기준으로 범위를 나눈

<표 IV-2> 과제 1에 대한 비형식적 추리 수준

수준		0	1	2			3	4	합계
				중심	변이	형태			
과제 1-1	중학생	2 (18.2)	1 (9.1)	4 (36.4)	1 (9.1)	1 (9.1)	2 (18.2)	0 (0.0)	11 (100.0)
	고등학생	15 (17.4)	21 (24.4)	20 (23.3)	1 (1.2)	6 (7.0)	22 (25.6)	1 (1.2)	86 (100.0)
과제 1-2	중학생	0 (0.0)	3 (27.3)	1 (9.1)	3 (27.3)	2 (18.2)	2 (18.2)	0 (0.0)	11 (100.0)
	고등학생	5 (5.8)	8 (9.3)	22 (25.6)	23 (26.7)	15 (17.4)	12 (14.0)	1 (1.2)	86 (100.0)

것은 비형식적으로 사분위수, 사분위간 범위에 주목한 것으로 볼 수 있다. 비율적 추론뿐 아니라 중심, 변이, 형태의 측면들을 연결시켜 통합적으로 추리하고 있음을 알 수 있다.

#### 나. 비형식적 추리의 수준 분석

<표 IV-2>는 과제 1에서 각 수준에 속하는 학생들의 빈도를 나타낸 것이다. 중학생과 고등학생 모두 2수준에 가장 많은 학생들이 속하는 것으로 나타났다. 고등학생은 2수준 다음으로 3수준, 1수준, 0수준, 4수준의 순서로 많은 학생들이 속했다. 토론 과정을 거치면서 수준의 변화를 나타낸 학생도 있었다. 0수준에 속하는 2명의 중학생은 모두 1수준으로 향상되었다.

비형식적 추리 수준의 평균은 과제 1-1의 경우 중학생이 1.73, 고등학생이 1.69, 과제 1-2의 경우 중학생이 1.91, 고등학생이 1.95로 중학생과 고등학생의 차이, 과제에 따른 차이가 크게 나타나지 않았다. 과제 1-1에서 중학생은 중심에 주목하여 비교한 2-중심 수준에 가장 많은 학생들이 속한 반면, 고등학생은 2-중심 수준뿐 아니라 비율적인 추론을 사용하며 자료에 대한 전체적 관점을 가진 3수준, 자료에 대한 국소적 관점을 가진 1수준에도 많은 학생들이 분포되었다. 과제 1-2에서 중학생은 1수준과 2-변이 수준에 많은 학생들이 속한 반면, 고등학생은 2-중심 수준과 2-변이 수준에 많은 학생들이 속하는 것으로 나타났다.

## 2. 모집단의 그래프 추측하기 과제

#### 가. 비형식적 추리의 특징 분석

표본으로부터 모집단을 추리할 때 필요한 핵심 아이디어는 표본이 모집단에 관한 모든 정보를 제공하는 것도 아니고, 전혀 정보를 제공하지

않는 것도 아니며, 약간의 정보를 제공한다는 것이다(Rubin et al., 1990, p. 314). 학생들은 표본대표성과 표집변이성 사이의 절충점을 찾아야 한다(Shaughnessy, 2007, p. 977). 표본대표성과 표집변이성에 대한 인식은 비형식적 추리 수준을 구분하는 기준이 될 수 있다.

모집단의 그래프 추측하기 과제에서 학생들의 반응은 개인적인 의견에 기초하여 타당하지 않은 추리를 하는 경우, 표본대표성에만 주목하고 표집변이성을 고려하지 않는 경우, 표본대표성과 표집변이성을 모두 고려하고 분포의 한 측면에 주목하는 경우, 분포의 다각적인 측면에 주목하는 경우로 특징지을 수 있었다. <표 IV-3>은 모집단의 그래프 추측하기 과제에서 각 수준에 속하는 학생들의 반응 특징을 요약한 것이다.

0수준의 학생들은 주어진 표본 자료를 고려하지 않고 개인적인 의견이나 경험에 기초하여 판단하였다.

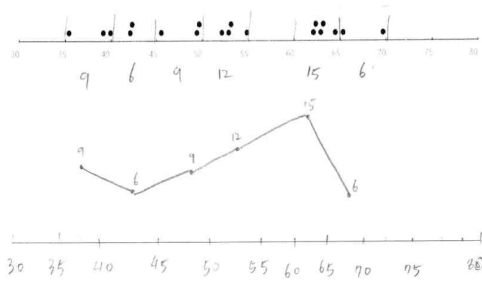
883 M1: 그냥 제 주변에 있는 친구들 몇 명을 데리고 평균 키를 계산을 했는데 남자 여자 합해서 평균키가 170 정도 되거든요. 그래서 그 평균키를 가지고 몸무게를 봤는데요.

885 M1: 제 주변에 되게 잘 먹는 친구들이 많아서 많이 나오더라고요.

M1은 주어진 자료에 기초하지 않고 ‘주변에 있는 친구들 몇 명’에 대한 개인적인 경험에 기초하여 그래프를 그렸다(line 883, 885). 0수준에 속하는 학생들은 주어진 표본의 정보 보다는 개인적 지식에 기초하여 모집단의 그래프를 추측하였다.

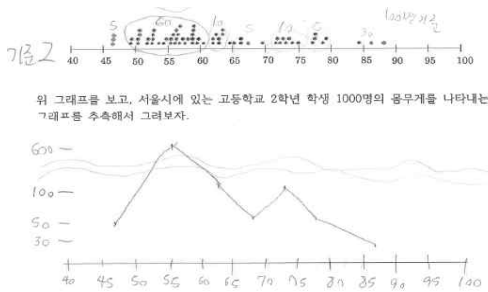
1수준의 학생들은 표집변이성을 고려하지 않고 표본대표성에만 주목하여 비율적 접근법을 사용하였다. 다음은 1수준에 속하는 학생들의 예이다. M9는 과제 2-1에서 각 구간의 도수에 3을 곱하여 그래프를 그렸고, 과제 2-2에서

1000/67의 값을 각 구간의 도수에 곱하여 모집단의 그래프를 그렸다(line 841). H63도 세 학급의 표본에서 비율을 계산하여 모집단의 그래프를 그렸다.



[그림 IV-13] 과제 2-1에서 M9의 반응

841 M9: 아마도 67명인 거 같아요. 점 개수가. 1000 나누기 67을 했더니 14.9 몇몇이 나와 가지고 15씩 곱했어요.

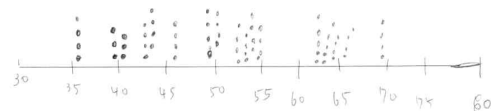


기준 1과 기준 2를 통해 100명 기준으로 몇 kg에 몇 명 있는지 예측한 후, 1000명 그래프를 그렸다.

[그림 IV-14] 과제 2에서 H63의 반응

M4는 표본 그래프 위에 같은 자료값을 갖는 점을 3개씩 추가하여 점그래프를 그렸고(line 629, [그림 IV-15]), M6은 계급의 크기를 10으로 정하고 각 계급의 도수가 가장 큰 순서대로 천명을 비례적으로 할당하여 그래프를 그렸다(line 861). M6은 표본에서 도수가 같은 두 계급에 대해 모집단에서도 도수를 같게 그렸다.

629 M4: 그냥 이 상태에서 세 개를 더 그리면 되는 거 아냐?



[그림 IV-15] 과제 2-1에서 M4의 반응

861 M6: 이게 구간별로 나눠가지고 40에서 50이 제일 많았으니까 40에서 50을 제일 높게 그리고, 그 다음에 50에서 60이랑, 60에서 70이 학생 수가 똑같았어요. 네 그래서 그냥 이것도 둘이 똑같이 했고, 그 다음에 말한 수가 30에서 40이고 마지막으로 이게 제일 적었어요.

<표 IV-3> 모집단의 그래프 추측하기에 대한 비형식적 추리 수준

수준	각 수준에서의 특징
0	자료에 기초하기 보다는 개인적인 의견이나 경험에 기초하여 타당하지 않은 추리를 함.
1	자료에 기초하여 비율적 접근법을 사용함. 표집변이성을 고려하지 않고 표본대표성에만 주목하여 결정론적 사고를 가짐.
2	표본대표성과 표집변이성을 모두 고려함. 분포의 한 측면에 주목하며 부분적으로 타당한 추리를 함. 자료에 맥락에 기초하여 추리함.
3	분포의 다각적인 측면들에 주목하지만 통계적 개념들을 통합하지 못함. 표본대표성과 표집변이성을 모두 고려하고 자료와 맥락에 기초하여 추리함.
4	분포의 다각적인 측면들에 주목하며 통계적 개념들을 통합하여 추리함. 표본대표성과 표집변이성을 모두 고려하고 자료와 맥락에 기초하여 추리함.

이들은 한 학급의 표본보다 세 학급의 표본이 모집단과 더 유사하다고 생각하며 표본의 크기의 통계적 영향을 인식하고 있었으나, 모집단의 크기에 대한 표본의 크기의 비율을 계산하여 추리에 사용하였고 가능성과 불확실성을 언급하지 않았다. 이러한 반응은 표본이 모집단에 대한 완전한 정보를 가지고 있다고 추론한 것으로 표본 대표성에만 주목하고 변이성을 고려하지 않는 결정론적 사고를 보여준다.

2수준의 학생들은 표본대표성과 표집변이성을 모두 고려하고 분포의 한 측면에 주목하여 부분적으로 타당한 추리를 하였다. 이 학생들은 1수준의 학생들처럼 비율적 추론을 사용했으나 변이를 고려하여 추리에 불확실성을 반영하였다.



민지네 한 학급을 조사한 것보다 세 학급을 조사했을 때가 조금 더 2학년 전체 학생 몸무게와 비슷할 것 같아서 3개 학급 조사한 것을 선으로 이었다.

[그림 IV-16] 과제 2에서 H27의 반응

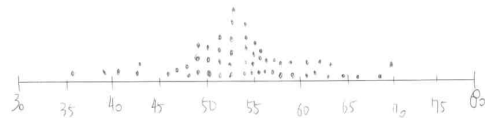
- 6 H27: 점을 거의 이으다시피 해서. 55 하고 60 사이에서 56, 57 이렇게 있을 거 아니에요. 몸무게가. 여기 보면 55 하고 60 사이에서도 56, 7 킬로 대 58, 9 대가 많아가지고 가운데는 별로 없는 걸로 나타나는 것처럼 보여서.

H27은 세 학급 표본의 그래프의 윤곽을 곡선으로 이어 모집단의 그래프를 그렸다. H27은 세 학급 표본의 그래프에서 다수가 위치한 55와 60 사이의 구간에서도 도수가 크고 작은 부분을 지적하며 ‘점을 이으다시피’라는 표현으로 비율적 추론을 나타냈다(line 6). 표본에 나타나지 않은

자료값들을 포함시켰고 곡선으로 나타낸 것은 변이성을 고려했다고 볼 수 있다. 그러나 분포의 울퉁불퉁함을 완만하게 유연화하지 못하였으므로 표본의 크기가 변화하는 과정의 안정적인 특징, 즉 소음 속에 존재하는 신호를 제대로 파악하지 못했음을 알 수 있다.

3수준의 학생들은 표본대표성과 표집변이성을 모두 고려하고 분포의 다각적인 측면들에 주목하지만 통계적 개념들을 통합하지 못하였다. 다음은 3수준에 속하는 학생들 반응의 예이다.

M10은 주어진 한 학급 학생들의 몸무게의 평균값을 계산하여 평균값인 53 근처에 자료값이 더 많이 분포되도록 점그래프를 그렸다. M4가 주어진 그래프에 같은 자료값을 갖는 점을 위로 3개씩 추가했던 것과는 달리 M10은 평균과 변이에 주목하였고 양 끝으로 갈수록 도수가 줄어드는 종 모양의 그래프를 그렸다. 표본 그래프의 점 사이사이에 점을 그려 넣은 것은 표집변이성을 고려한 것으로 볼 수 있다.

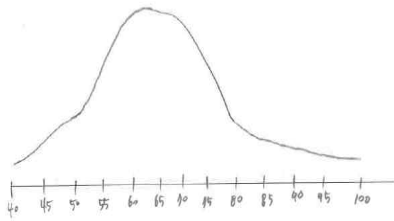


[그림 IV-17] 과제 2-1에서 M10의 반응

H12는 ‘제일 밀집’, ‘중양은 많이 있고 옆으로 짝’이라는 표현을 사용하여 중심, 밀도, 변이에 대한 인식을 나타냈다(line 61, 63). ‘없다고 해서 없는 건 아니고’, ‘많이 없으니까 많이는 없을 거고’라는 표현에서 표집변이성과 표본대표성을 모두 고려하고 있음을 알 수 있다. H78도 ‘55~80에서 가장 집중적으로 분포’라는 표현을 사용하여 밀도와 다수 자료의 위치에 주목함을 나타냈다. 변이에 주목하여 범위를 표본보다 더 크게 했고 양 끝으로 갈수록 밀도가 줄어드는 단봉형

곡선 그래프를 그렸다. 이 학생들은 자료의 신호에 주목하였고 다양한 통계적 개념을 사용하여 추리하였으나 분포의 형태와 적절히 통합시키지는 못했다.

- 61 H12: 55에서 60은 이 세 개 반을 봤을 때, 제일 밀집해 있어가지고. 이 구간에서.
- 63 H12: 세 개 반에서 한 반에서 95에서 100이 없다고 해서 천 명에서 봤을 때 없는 건 아니거든요. 그러니까, 그렇지만 여기서 볼 때도 많이 없으니까 많이는 없을 거고 여기도 없으니까 많이는 없을 거라 해가지고, 중앙은 많이 있고 옆으로 짝이렇게 되는.



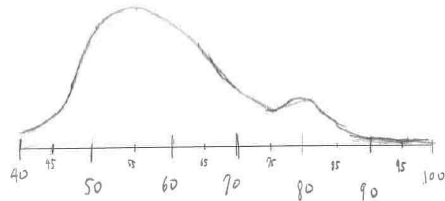
민지의 학급에서나, 같은 학년 세 학급에서나, 55~80kg 대에 학생들의 몸무게가 가장 집중적으로 분포되어 있기 때문이고, 비록 위 두 그래프에서는 40kg 대, 90kg 대 학생들이 존재하지 않지만, 전국적으로 따졌을 때 소수의 학생은 존재할 것이다.

[그림 IV-18] 과제 2에서 H78의 반응

4수준의 학생들은 분포의 다각적인 측면들에 주목하며 통계적 개념들을 통합하여 추리하였고

표본대표성과 표집변이성을 모두 고려하였다. 다음은 4수준에 속하는 학생 반응의 예이다.

H80은 밀도, 다수 자료의 위치, 변이, 분포의 형태 등 다각적인 측면들을 통합시켰으며 쌍봉형 곡선 그래프로 나타내었다. 4수준의 학생들은 다양한 통계적 개념들을 연결하여 표본 그래프의 특징을 적절히 설명했을 뿐 아니라 그러한 특징이 모집단에 대한 추리에 어떻게 사용되는지를 언급하였다.



민지네 반 친구들과 민지와 같은 학년 세 학급의 통계를 보아 50~65kg 사이의 아이들이 가장 많다는 것을 유추할 수 있다. 세 학급의 통계를 볼 때 75~80kg 부분에서 꽤 많은 아이들이 나와서 75~80kg 부분에서 다시 살짝 오르도록 설정해 봤다. 그리고 40kg과 100kg 부분에서 위 두 자료에는 없다고 했지만 1000명 대상으로 볼 때 있을 가능성이 좀 더 높아 일부러 0으로 표시하지 않았다.

[그림 IV-19] 과제 2에서 H80의 반응

나. 비형식적 추리의 수준 분석

<표 IV-4>는 과제 2에서 각 수준에 속하는 학생들의 빈도를 나타낸 것이다. 중학생은 1수준

<표 IV-4> 과제 2에 대한 비형식적 추리 수준

수준		0	1	2	3	4	합계
중학생	과제 2-1	3 (27.3)	5 (45.4)	2 (18.2)	1 (9.1)	0 (0.0)	11 (100.0)
	과제 2-2	1 (9.1)	6 (54.5)	1 (9.1)	2 (18.2)	1 (9.1)	11 (100.0)
고등학생		5 (5.8)	3 (3.5)	9 (10.5)	33 (38.4)	35 (40.7)	86 (100.0)

에, 고등학생은 4수준과 3수준에 가장 많은 학생들이 속하는 것으로 나타났다. 중학생의 경우 과제 2의 활동이 진행됨에 따라 0수준에 속하는 학생은 감소했고 분포의 다각적 측면을 고려하는 3수준 이상의 학생들이 증가했다. 이는 토론 과정에서 개인적 경험에만 주목한 학생들이 자료와 통계적 개념에 주목할 수 있었기 때문이다.

비형식적 추리 수준의 평균은 중학생의 경우 과제 2-1, 2-2에서 각각 1.09, 1.64, 고등학생은 3.06으로 중학생과 고등학생의 차이가 크게 나타났다. 중학생들은 표본대표성만 고려하여 표본의 상대도수, 표본과 모집단의 크기의 비에 주목하였고 표집변이성을 간과하는 경향이 나타났다. 특히, 활동 초반에는 개인적 경험에 근거하는 경향이 있었으나 교사가 자료에 기초하여 판단할 것을 요구하자 정확한 비율을 계산하여 이용하려는 경향이 뚜렷해졌다. 반면 고등학생들은 표본대표성과 표집변이성을 모두 고려하였으며 분포의 형태에 주목하는 경향이 있었다.

## V. 논의 및 결론

비형식적 통계적 추리의 수준을 구분하고 각 수준에 해당하는 학생들의 반응을 확인한 결과, 0수준을 나타낸 학생들의 비율은 과제 1-1에서 중학생이 18.2%, 고등학생이 17.4%이었던 반면 과제 1-2에서는 중학생이 0.0%, 고등학생이 5.8%로 낮아졌다. 일부 학생들은 비교하는 표본의 크기가 반드시 같아야 한다는 생각을 가지고 있었기 때문에 과제 1-1에서 주관적 판단에 의존하는 0수준의 반응이 많이 나타났다. 실생활에서 어떤 두 표본을 비교할 때, 표본의 크기가 정확하게 일치하는 경우는 많지 않다. 교과서에서 다루는 자료집합 비교 과제들은 자료값의 개수가 정확히 일치하는 경우만 다루고 있으므로 비교하는

두 표본의 크기가 같아야 한다는 사고는 고착될 우려가 있다. 실험에서 0수준에 속하던 학생들이 1수준으로 향상된 것은 표본의 크기에 대한 관점의 변화 때문이었다. 표본의 크기가 상대적으로 작기 때문에 표본의 대표성을 의심하는 것에서 점차 표본 크기의 영향력을 통계적으로 해석하는 것으로 관점이 변화하였다.

자료에 대한 전체적 관점을 가진 2수준 이상에 속하는 학생은 중학생의 경우 과제 1-1, 1-2에서 모두 72.7%이었지만 고등학생의 경우는 과제 1-1에서 58.1%, 과제 1-2에서 84.9%이었다. 과제 1-1에서 고등학생이 중학생보다 2수준 이상에 속하는 학생의 비율이 낮았던 것은 표본 크기의 통계적 영향력에 지나치게 주목하여 중심, 변이, 형태 등의 측면에는 주의를 덜 기울였기 때문이었다.

과제 1-1에서는 중심에 주목한 학생이 많았고 변이나 형태에 주목한 학생이 적었지만, 과제 1-2에서는 변이나 형태에 주목한 학생들이 많았다. 이는 과제 설계의 영향으로 발생하였을 가능성이 크다. Shaughnessy 외(2004)는 두 자료집합 비교에서 중심의 역할이 중요하지 않을 때 자료의 변이성에 주목한다는 것을 보여주었는데, 본 연구에서는 학생들이 과제 1-2에서 중심경향치가 다르더라도 자연스럽게 변이를 고려하고 있었다. 학생들은 변이를 고려할 때 퍼진 정도보다는 범위를 더 선호하였는데, 이는 송선아·이경화(2007)의 연구 결과와 유사하였다.

과제 1에서 중학생과 고등학생의 평균 수준의 차이는 거의 나타나지 않았다. 분포의 다각적 측면에 주목하는 3수준 이상의 학생이 적었기 때문에 두 학년 모두 평균 수준이 2수준에 미치지 못했다. 이는 표본 비교 활동에 대한 경험의 부족뿐만 아니라 분포 개념에 대한 이해의 부족 때문이기도 하다. 학교수학에서는 분포, 대푯값, 산포도를 개별적으로 지도하고 그 관련성에 주

목하지 않기 때문에 통계적 개념들 사이의 관계를 인식하고 통합하는 기회가 부족한 것도 이유가 될 수 있다.

표본 비교하기에 대한 비형식적 추리의 0~4수준은 각각 Ciancetta(2007)의 주관적, 국소적, 과도기적, 초기분포적, 분포적 수준과 대응된다. 그러나 3수준에 속하는 모든 학생들이 비율적 추론을 사용하였기 때문에 초기전체적 수준이 따로 구분되지는 않았다. 이는 대학생과 중등학생이라는 연구대상의 차이 때문으로 볼 수 있다. 또한 0, 2, 3, 4수준은 각각 Huey(2011)의 비형식적 통계적 추리의 전구조적, 단일구조적, 다중구조적, 관계적 수준과 유사하지만 개별적인 자료값에 관한 추론은 중심, 변이, 형태와 같은 분포의 한 요소에 관한 추론과 차이가 있기 때문에 1수준을 포함시켜 이를 구분하였다. Goss(2014)는 표본 비교활동에서 SOLO 모델의 네 수준과 함께 학생들이 주목한 측면이 적절한지의 정도를 고려한 3개의 하위 주기를 고려하여 10개의 범주로 구분하였는데, 본 연구에서는 이 부분에 대한 고려가 부족하므로 하위 주기를 포함하는 수준에 대한 후속연구가 이루어질 필요가 있다.

과제 2에서 0수준을 나타낸 학생들의 비율은 중학생의 경우 과제 2-1, 2-2에서 각각 27.3%, 9.1%, 고등학생은 5.8%로 나타났다. 중학생의 경우, 과제 2-1보다 2-2에서 주관적 의견에 기초하여 타당하지 않은 추리를 하는 0수준의 학생들의 비율이 감소했고, 표본대표성을 고려한 학생들의 비율은 증가했다. 크기가 작은 표본의 신호는 확신할 수 없지만 크기가 더 큰 표본의 신호는 신뢰할 수 있다고 인식했기 때문으로 볼 수 있다. 이는 분포에 신호가 불분명할 때는 예외적인 사례에 주목하거나 추리를 기피하지만, 분포에 신호가 분명히 나타날 때는 불확실성에 대한 표현을 거의 하지 않고 추리하는 것으로 나타난 Ben-Zvi 외(2012)의 연구와 유사한 결과였다.

표본대표성과 표집변이성을 모두 고려한 2수준 이상의 학생들의 비율은 중학생이 36.4%, 고등학생이 89.6%로 두 학년의 차이가 크게 나타났다. 중학생과 고등학생의 평균 수준의 차이도 매우 컸다. 약 90%의 고등학생들은 표본대표성과 표집변이성을 모두 고려하여 추측하였으나, 54.5%의 중학생들은 표집변이성을 간과하고 표본대표성에만 주목하였다. 특히, 1수준에 속하는 중학생들은 모두 비율적 접근법을 사용하였다. 중학생들이 자료의 신호에 덜 주목하고 결정론적으로 사고하는 경향이 두 학년의 수준 차이에 큰 영향을 주었다고 볼 수 있다.

4수준에 속하는 학생들의 비율은 중학생이 9.1%, 고등학생이 40.7%로 나타났다. 고등학생의 경우는 Bakker와 Gravemeijer(2004)의 연구와 같이 분포를 세 개의 집단, 즉 평균에 가까운 집단, 높은 값들의 집단, 낮은 값들의 집단으로 구분하는 경향이 나타났지만, 네 개의 집단 즉, 밀도가 가장 큰 집단, 밀도가 두 번째로 큰 집단, 높은 값들의 집단, 낮은 값들의 집단으로 나누는 경향도 나타났다. 중학생들보다 고등학생들이 표본의 신호를 파악하고 다양한 통계적 개념들을 분포의 형태와 적절히 통합시킬 수 있었던 것은 함수 학습 경험과 같은 교육의 간접적 영향 때문일 수 있다.

Goss(2014)는 모집단의 그래프 추측하기 활동에서 모든 자료값을 같은 양만큼 증가시키거나 그래프를 평평하게 나타낸 반응을 1주기로, 도수를 두 배하거나 새로운 자료값을 포함시키지 않고 세로축을 늘려 표본의 형태를 유지하여 나타낸 반응을 2주기로, 적절한 자료값을 포함시켜 가로축까지 확장하여 분포의 개형을 나타낸 반응을 3주기로 구분하였다. 본 연구에서는 도수에 표본과 모집단의 크기의 비를 곱하는 반응은 나타났지만 도수를 같은 양만큼 증가시키는 반응은 없었기 때문에 이를 1수준으로 구분하였고,

표본의 형태를 유지하며 윤곽선으로 나타낸 반응은 2수준으로 구분하여, Goss(2014)의 2주기 반응을 서로 다른 수준으로 나타내었다. 또한 반응 구조의 복잡성을 고려하여 2주기의 반응을 3수준과 4수준으로 구분하였다.

모집단에 대한 추리는 표본을 기초로 모집단에 대한 정보를 알아내기 위한 과정이므로 표집 과정의 영향을 받는다. 따라서 표집에 대한 이해는 모집단의 그래프 추측하기에 대한 비형식적 추리의 수준을 결정하는 요소가 될 수 있다. Watson(2004)의 표집 개념의 발달 연구에서 변이와 대표성, 편의에 대한 인식은 수준 구분의 주요 요소이었다. 본 연구에서도 변이와 대표성에 대한 인식을 토대로 수준을 구분하였다. 과제 2는 그래프를 그리는 활동이었기 때문에 표본 분포를 해석하고 신호를 찾기 위해 다양한 통계적 개념들에 주목할 필요가 있었다. 따라서 어떠한 통계적 개념에 주목하는지를 수준 구분의 또 다른 요소로 사용하였다. 표본 개념과 함께 다양한 통계적 개념들을 관계적으로 이해해야 올바른 모집단의 추리가 가능하다. 추측통계를 현재보다 낮은 학년에서부터 점진적이고 비형식적으로 지도하기 위해서는 표본과 표집에 주목하는 비형식적 추리 활동의 기회를 제공할 필요가 있다.

기존의 통계수업에서는 주로 주어진 그래프를 해석하거나 값을 구하는 폐쇄형 문제를 다루는데 반해 본 연구의 교수 실험에서는 의견을 묻는 문제와 추측하는 문제를 다루었기 때문에 의사소통이 더 활발히 일어날 수 있었다. 폐쇄형 문제는 알고리즘적 절차에 주목하게 하므로 통계적 추리를 지도하는데 적절하지 못하다. 본 연구의 실험 과제는 해법이 다양할 뿐만 아니라 해법에 대한 다양한 평가 준거를 가지며 자료에 근거한 논증에 개인적 신념을 표현해야 하는 개방형 과제였다. 개방형 과제는 추측을 자극하고 이 추측의 불확실성은 정당화의 필요성을 유발

하여 자연스럽게 추측을 정당화하는 탐구 과정을 경험하도록 유도한다. 과제 1과 같이 시각적 표현을 제공하고, 과제 2와 같이 가설적 상황에서 그래프의 형태를 예측하는 과제는 하나의 자료값에 주목하기보다 전체적인 특성과 요소들 사이의 관계에 주목할 수 있게 하므로 통계적 추리를 지도하는데 유용하였다. 또한 두통약이나 몸무게와 같이 학생들에게 친숙한 맥락은 통계적 추리에 필요한 정신적 모델 구성에 도움을 주었다.

수업을 통해 새롭게 깨달은 점이 무엇인지를 묻는 질문에 중학생들은 통계적 추리에 타당한 근거가 필요함을 인식하게 되었다고 답했다. 모든 학생들의 의견을 듣고 친구의 의견에 대해 보충하거나 반박할 것을 요구하였기 때문에 의견의 공유가 이루어졌고, 추리의 근거에 주목하였으며, 추측통계의 본질적인 개념 요소에 초점을 둘 수 있었다. 통계적 추리의 지도에 개방형 과제, 시각적 표현, 맥락이라는 요소뿐 아니라 교실 토론에 대한 사회-수학적 규범의 정립이 중요하다는 시사점을 얻을 수 있었다.

본 연구는 중학생 참여자의 수가 적으며, 고등학생을 대상으로 교수 실험을 실시하지 못했고, 중학생과 고등학생에 대해 서로 다른 조사 방법을 사용하였다는 제한점이 있다. 교육과정에서 다루지 않는 내용을 입시를 앞둔 고등학생들에게 지도하는 것에 현실적인 제약이 있었기 때문에 제한된 형태의 연구가 이루어졌다. 비록 이러한 제한점이 있으나 중·고등학생들의 통계적 소양에 대한 단상을 드러낼 수 있었다는 데에서 그리고 향후 관련 교육의 방향을 수립하는 데에 관련된 시사점을 도출하였다는 데에서 연구의 의의를 찾을 수 있다. 참여자의 학령을 다양화하고 참여자의 수를 확대하여 수준 구분에 대한 논의를 보완하는 후속연구가 이루어져야 한다. 또한 Zieffler 외(2007)가 제시한 비형식적 추리



과제 중 두 종류의 과제를 다루었으므로 다른 유형의 과제에서는 학생들의 비형식적 추리의 양상과 수준이 어떠한가에 대한 연구가 이루어질 필요가 있다. 비형식적 추리의 수준에 대한 연구를 바탕으로 통계적 추리 발달을 위한 학습 위계 설정 및 학습지도 방법에 대한 연구도 이루어져야 할 것이다.

## 참고문헌

- 박민선, 박미미, 이경화, 고은성(2011). 자료집합 비교 활동에서 나타나는 중학교 학생들의 통계적 추리(statistical inference)에 대한 연구. **학교수학**, 13(4), 599-614.
- 송선아, 이경화(2007). 중학교 3학년 학생들의 변이성 이해에 대한 사례 연구. **학교수학**, 9(1), 29-44.
- Bakker, A. (2004). Reasoning about shape as a pattern in variability. *Statistics Education Research Journal*, 3(2), 64-83.
- Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp.147-168). Dordrecht: Kluwer Academic Publishers.
- Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM Mathematics Education*, 44, 913-925.
- Biggs, J., & Collis, K. (1991). Multimodal learning and the quality of intelligent behavior. In H. Rowe (Ed.), *Intelligence, Reconceptualization and Measurement* (pp. 57-76). New Jersey: Erlbaum.
- Bright, G. W., Brewer, W., McClain, K., & Mooney, E. S. (2003). *Navigating through data analysis in grades 6-8*. Reston, VA: National Council of Teachers of Mathematics.
- Ciancetta, M. A. (2007). *Statistics students reasoning when comparing distributions of data*. Unpublished doctoral dissertation, Portland State University, Oregon.
- Goss, J. M. (2014). *A method for assessing and describing the informal inferential reasoning of middle school students*. Unpublished doctoral dissertation, Western Michigan University, Michigan.
- Huey, M. E. (2011). *Characterizing middle and secondary preservice teachers' change in inferential reasoning*. Unpublished doctoral dissertation, University of Missouri-Columbia, Missouri.
- Jacob, B. L. (2013). *The development of introductory statistics students' informal inferential reasoning and its relationship to formal inferential reasoning*. Unpublished doctoral dissertation, Syracuse University, New York.
- Madden, S. R. (2011). Statistically, technologically, and contextually provocative tasks: supporting teachers' informal inferential reasoning. *Mathematical Thinking and Learning*, 13, 109-131.
- Makar, K. (2016). Developing young children's emergent inferential practices in statistics. *Mathematical Thinking and Learning*, 18(1), 1-24.
- Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1),

- 82-105.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards: Mathematics standards*. Retrieved July 17, 2010 from <http://www.corestandards.org>.
- Park, M. S. (2015). *Assessment of Informal Statistical Inference*. Unpublished doctoral dissertation, Seoul National University, Seoul.
- Pfannkuch, M. (2006). Informal inferential reasoning. In A. Rossman & B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil.
- Rubin, A., Bruce, B., & Tenney, Y. (1990). Learning about sampling: Trouble at the core of statistic. In *Proceedings of the Third International Conference on Teaching Statistics* (pp. 314-319). Dunedin, New Zealand.
- Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1009). Charlotte, NC: Information Age Publishing.
- Shaughnessy, J. M., Ciancetta, M., Best, K., & Canada, D. (2004). Students' attention to variability when comparing distribution. Paper presented at *the Research Presession of the 82nd Annual Meeting of the National Council of Teachers of Mathematics*, Philadelphia, PA.
- Watson, J. M. (2004). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking* (pp. 277-294). Dordrecht: Kluwer Academic Publishers.
- Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40-58.

# Study on the Levels of Informal Statistical Inference of the Middle and High School Students

Lee, Jung Yeon (Seong-su High School)  
Lee, Kyeong Hwa (Seoul National University)

The statistical education researchers advise the student's view of the data is transitioning from instructors to educate informal statistical inference local to global, responses that hold a global view and they are paying close attention to the progress but do not clearly integrate multiple aspects of the statistical inference in general. This study the distribution, and responses that integrate multiple aspects of the distribution. Another five levels of the informal statistical inference were identified for comparing the samples of data and estimating the graph of a population: responses that are distracted or misled by an irrelevant aspect, responses that focus only on representativeness, responses that consider both representativeness and variability and focus on one particular aspect of the distribution, responses that focus on multiple frequencies of individual data points and hold a local view of the sample data sets, responses that aspects of distribution but do not clearly integrate them, and responses that integrate multiple aspects of the distribution.

\* Key Words : informal statistical inference(비형식적 통계적 추리), levels of statistical inference(통계적 추리 수준), comparing the samples of data(표본 비교하기), estimating the graph of a population(모집단의 그래프 추측하기)

논문접수 : 2017. 8. 10

논문수정 : 2017. 9. 14

심사완료 : 2017. 9. 15