

중등수학 예비교사들의 통계적 소양 : 표본 개념에 대한 이해를 중심으로

탁 병 주 (서울대학교대학원)
구 나 영 (안양고등학교)
강 현 영 (목원대학교)
이 경 화 (서울대학교)[†]

I. 서론

2015 개정 수학과 교육과정에서도 밝히고 있듯이, 자료를 수집, 정리, 해석하는 과정에서 필요한 통계적 소양은 현대 정보화 사회의 불확실성을 이해함으로써 미래를 예측하고 합리적인 의사 결정을 하는 민주 시민으로서의 기본 소양이다(교육부, 2015, p.35). 특히, 각종 매체를 통해 범람하는 각종 통계적 조사들이 대부분 모집단 전수 조사가 아니라 일부만을 대상으로 실시된 표본조사라는 점을 고려할 때, 표본 개념은 통계적 소양의 기초로서 학교 통계교육에서 매우 강조되어야 하는 개념임을 이해할 수 있다(탁병주 외, 2014, p.277). 실제로 Watson(2013)은 통계 조사에서 표집의 중요성을 간과하면 나머지 절차가 완전히 쓸모없는 것이 되기 때문에, 표본과 표집의 특성을 이해하는 것이 통계적 소양의 한 영역이라 주장한 바 있다(p.40).

표본 개념은 반복가능성, 대표성, 임의성, 변이성, 분포 등 매우 다양한 통계적 개념들이 상호 관련하여 하나의 스키마를 형성하고 있다(Pfannkuch, 2008; Saldanha & Thompson, 2002). 이러한 개념의 복잡성으로 인해, 학생뿐만 아니라 교사들 또한 표본 개념에 대한 종합적인 이해 없이 통계적 추론 과정에서 부정확한 직관에 의존하곤 한다. 일례로, Watson(2013)은 공식이나 절차를

내포하지 않고 대상의 특성을 기술하는 표본 개념의 표현적 특성이 일부 수학교사들에게 어려움으로 다가오기 때문에 그래프 그리기나 평균 계산과 같이 기술과 알고리즘을 필요로 하는 절차적 지식에 더 주목한다고 언급하였다(p.40). 통계 교수를 위해 교사에게 필요한 지식에는 Hill et al.(2008)이 제시한 수학 교수를 위한 지식(mathematical knowledge for teaching)에 포함되지 않는 비수학적인 지식이 존재한다는 선행연구에 비추어볼 때(Groth, 2007), 비결정론적이고 자료기반적인 통계 고유의 사고를 요구하는 표본 개념이 예비교사교육에서 충분히 다루어져야 현직교사들이 겪는 위와 같은 어려움을 해소할 수 있다. 따라서 예비교사교육의 방향을 결정하고 유의미한 시사점을 도출하기 위해서는 예비교사의 이해를 조사하는 기초 연구가 필요하다. 실제로 국내 연구는 예비교사와 현직교사를 연구 대상으로 하는 통계교육 연구가 많이 부족한 실정이다(이영하·심효정, 2003, p.212; 이은희·김원경, 2015, p.256).

국외에서는 표본 개념에 대한 예비교사의 이해를 조사한 연구가 일부 이루어져 왔는데(예를 들어, Watson, 2000; Groth & Bergner, 2005; Noll, 2011), 예비교사들이 통계학의 비결정론적이고 경험적인 성격을 이해하기 위해서는 직접적인 표집 경험을 통해 수학과 대비되는 통계학의 핵심 속성인 변이성(Franklin et al., 2007, p.6)을 인식하게 해야 한다는 것이 공통된 결론이었다. 우리나라에서는 고은성·이경화(2011)가 예비교사를 대상으로 표본 개념에 대한 이해를 조사한 바 있으며, 우리나라의 예비교사들 역시 표본대표성에 비해 표집변이성과 표집 분포에 대한 이해가 부족하다고 나타난 바 있다. 다만, 연구에 참여한 예비교사의 수가 다소 적었던 데 비해 연

* 접수일(2016년 9월 25일), 수정일(1차: 2016년 12월 11일, 2차: 2017년 1월 16일), 게재확정일(2017년 02월 21일)

* ZDM분류 : B5

* MSC2000분류 : 97C70

* 주제어 : 통계적 소양, 표본대표성, 표집변이성, 통계교육
† 교신저자

구에 활용된 질문지의 문항이 대부분 선다형이었기 때문에, 내용 지식의 유무 외에 다른 논의를 이끌어내기 어려웠다는 점에서 한계가 있다.

본 연구에서는 이러한 선행연구의 한계점을 개선하고자 열린 형태의 문항으로 구성된 질문지를 새로 설계하여 예비교사들을 대상으로 조사한 후, 질적 분석을 통해 표본 개념에 대한 예비교사의 이해가 가지는 특징들을 범주화한다. 특히 예비교사를 대상으로 하는 현행 통계교육 체계를 반성하기 위해, 본 연구에서는 예비교사를 대상으로 하는 현행 통계 과목의 수강 여부를 하나의 변수로 상정한다. 이렇게 범주화된 반응들은 기존의 연구들에 비추어 각 범주의 통계적 소양을 고찰함으로써, 종래에는 예비교사의 통계적 소양 수준을 논의한다. 우리나라의 수학과 교육과정에는 민주 시민으로서 갖추어야 할 통계적 소양이 명확히 제시되지 않아 통계적 추론(reasoning) 대신 기술(skill)에 치중한다는 비판이 제기된 바 있으므로(강현영 외, 2014), 기존 연구에서 충분히 다루지 않았던 예비교사의 통계적 소양에 대한 논의는 예비교사를 대상으로 한 통계교육에 유의미한 시사점을 제시할 것이라 기대한다.

본고의 연구 질문은 다음과 같다.

1. 표본대표성과 표집변이성의 측면에서, 예비교사들은 표본 개념에 대해 어떻게 이해하고 있는가?
2. 표본 개념에 대한 이해에 비추어볼 때, 예비교사의 통계적 소양 수준은 어떠한가?

II. 이론적 배경

1. 표본대표성과 표집변이성

수학에서는 일반적으로 연구 대상에 대해 대표성을 고려할 필요가 없다. 그 이유는 수학이 설명력이 피설명항을 완전히 함축하고 있는 연역법을 학문적 방법론으로 채택하고 있기 때문이다. 그러나 통계에서 연구 대상이 지니는 대표성은 매우 중요한 요소인데, 이는 통계적 방법론이 '일부'에 불과한 경험적 사실을 토대로 '전체'에 대한 일반적인 법칙이나 이론을 세우는 귀납법을 채택하고 있기 때문이다. 통계에서 대개 '일부'는 표본을 의미하고 '전체'는 모집단을 의미하는 바, 귀납법으로서의 통계적 방법에 대한

반성과 극복의 결과는 표본대표성의 이해와 확보를 위한 표집 이론의 발전으로 나타났다(정한영, 1995).

표본대표성이란 모집단에서 추출된 표본이 모집단과 비슷한 특성을 가질 것이라는 아이디어이며(Rubin et al. 1990, p.314), 통계가 단순한 기술이 아닌 과학의 범주에 포함되도록 하는 근거가 된다. 역사적으로는 19세기 말 노르웨이의 통계학자 A. N. Kær가 오늘날의 층화다단계 표집과 같은 '대표기법(representative method)'을 제안함으로써 표본대표성에 대한 이론적 설명이 시작되었고, 미국 Literary Digest지의 1936년 대선 예측 실패 사건을 거쳐 모집단과의 관계를 약화시키는 편의(bias)의 개념과 등확률성을 바탕으로 한 임의표집 개념이 종합된 오늘날의 표본대표성 개념이 완성되었다(탁병주 외, 2014). 표본대표성은 표본 개념이 아니라 표집 개념에 내포된 것이므로, 표본대표성에 대한 설명은 자료를 수집하는 중간 단계에서 이루어져야 한다.

임의표집 외에도 표본대표성을 높이기 위한 여러 가지 기법들이 등장하였으나 표본과 모집단 간의 불일치는 여전히 존재하는데, 표본대표성에 영향을 주는 요소가 편이만 존재하는 것은 아니기 때문이다. Moore & Notz(2009)에 따르면 표본대표성은 편이 외에도 표집변이성의 영향을 받는데, 편이가 표본을 여러 번 선택할 때 모수에서 한 방향으로 일관되고 반복적으로 벗어나는 통계량의 편차를 의미한다면, 표집변이성은 표본을 여러 번 선택할 때 통계량의 값이 얼마나 퍼져 있는가를 의미한다. 표집변이성은 단일 모집단에서 추출된 다수의 표본은 항상 같을 수 없으며, 따라서 모집단과도 온전히 일치할 수 없다는 아이디어이다(Rubin et al. 1990, p.314).

편의는 일관되게 편차가 나타나는 '비자연적 이유'를 설명하고 임의표집을 통해 제어가 가능하지만, 표집변이성은 우연 현상에 내재된 '자연스러운' 현상이기 때문에 임의표집을 반복하여 변이의 예측가능한 패턴을 '설명'해야 한다. 이때, 임의표집의 반복은 표집변이성을 바라보는데 우연의 불확실성에 대한 대응 방식으로서 표집분포라는 렌즈(Wild, 2006)를 필요하게 만들었고(이영하, 2014), 나아가 확률분포 등의 수학적 개념을 활용하는 형식적인 통계적 추리가 발전하는 데 기여하였다. 표집변이성에 대한 추론은 이러한 형식적이고 전문적인 통계

적 사고로 나아가는 경로로서의 역할을 한다(Garfield et al., 2015).

요컨대, 표본대표성과 표집변이성은 통계를 기술이 아닌 과학에 편입하는 데 공헌한 한편, 수학과는 분명히 구분되면서도 한편으로는 수학을 활용하게 만드는 독특한 수리과학으로서의 정체성을 형성해 하였다. 통계학이 귀납추론으로 이루어진 방법론을 구성하고 이러한 통계적 추론의 정당성을 이해하기 위한 연역적 근거를 찾아나가는 독특한 사고 체계를 갖추었다는 점(이영하, 2014)에 비추어볼 때, 표본대표성과 표집변이성은 Rubin et al.(1990)의 주장처럼 표본 개념을 이해하고 나아가 통계적 추론을 학습하는 데 핵심이 되는 속성이라 할 수 있다. Ben-Zvi et al.(2015)과 Shaughnessy(2007) 등도 표본대표성과 표집변이성은 통계적 추리를 이해하기 위해 균형 있게 이해되어야 하는 속성이라 밝히고 있다. 이는 표본대표성과 표집변이성이 표본 개념에 대한 예비교사의 이해와 통계적 소양을 분석하는 데 중심으로 고려할 만한 속성임을 뒷받침한다.

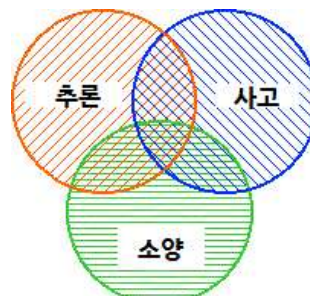
2. 통계적 소양과 표본

현대 정보 사회에서는 각 분야에서 일어나는 각종 불확실한 현상에 대한 합리적인 판단 도구로서 통계적 방법이 놀라울 정도로 광범위하게 사용되고 있다. 어떻게 정보가 처리되며 유용한 지식으로 전이되는가에 대한 이해가 필수적이다. 통계 자료, 통계적 요약, 통계에 대한 추론은 다른 형태의 수학적 분석보다도 일상생활에서 빈번하게 나타난다. 따라서 올바른 예상과 의사결정을 위해 자료 분석에 사용되는 개념과 그 과정을 이해하는 등의 통계적 소양이 현대 사회생활을 영위하는데 필수적으로 요구되고 있다(탁병주 외, 2014, p.734).

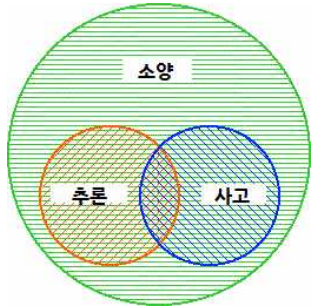
그러나 통계교육의 필요성이나 통계교육의 목적이라는 측면에서 ‘통계적 소양(statistical literacy)’이 주목을 받고 있음에도 이에 대한 합의된 정의가 없어 아직까지 그 의미는 매우 모호하고 다양하다(강현영, 2012, p.122). Wallman(1993)은 통계적 결과의 이해와 비판적 평가 능력을 통계적 소양으로 지칭하였고(p.1), Watson(1997)은 맥락 내에서 통계적 정보의 의미와 시사점, 텍스트를 이해하는 것이 통계적 소양이라 언급하면서 가장 높은 수준의 통계적 소양은 ‘자료에 근거하지 않은 모순된 주장

에 통계적 개념을 적용하여 문제를 제기하는 능력’이라 주장하였다. 또한 Gal(2002)은 통계적 정보와 현상에 대한 해석과 비판적 평가 및 의사소통이 통계적 소양의 구성 요소라 보았다.

[그림 1]처럼 통계적 소양을 통계적 사고, 통계적 추론과 구분하여 정의하고자 시도했던 Garfield et al.(2003)은 ‘통계적 정보나 연구 결과를 이해하는 데 이용되는 기본적인 중요한 기능’으로서 통계적 소양을 정의하였다. 이러한 관점에서 통계적 사고는 통계적 추론보다 더 높은 수준의 사고를 수반하며, 통계적 추론은 통계적 소양보다 더 심도 있는 수준의 개념 이해를 수반하는 것으로 설명된다(Ben-Zvi & Garfield, 2010). 그러나 이는 통계적 소양을 추론과 사고에 필요한 지식과 기술이라는 측면에서만 바라본 전통적인 관점에 기인한 것이다(delMas, 2002). 오늘날 통계적 소양은 민주사회의 시민에게 필요한 통계교육의 목적으로서 다루어지고 있기 때문에 통계적 소양은 [그림 2]와 같이 포괄적인 개념으로 인식된다(Gal, 2002). 이때의 통계적 추론과 통계적 사고는 각각 통계 비전문가, 그리고 전문가로서 통계적으로 유능한 시민의 육성을 목표로 하는 통계교육의 세부 목표이다(고은성, 2015, p.85). 이러한 관점에서 볼 때, 위에서 언급된 통계적 소양에 대한 다양한 정의 중 가장 주목할 부분은 ‘비판’과 ‘의사결정’이며, 실제로 Rumsey(2002)는 교육과정에 통계를 도입해야 하는 첫 번째 이유로 “정보에 대한 비판적 사고와, 정보에 기초한 합리적 의사결정”을 꼽았다.



[그림 1] 통계적 소양의 독립 모형
[Fig. 1] The hierarchical model of statistical literacy



[그림 2] 통계적 소양의 포괄 모형
[Fig. 2] The comprehensive model of statistical literacy

본 연구에서는 예비교사의 통계적 소양을 분석하는 틀로서 <표 1>에 제시된 Watson(1997)의 통계적 소양 수준 체계를 활용한다. Watson(2013)은 학교수학에서 다루어야 하는 통계적 소양에 대해 다음과 같이 언급하면서, 특히 통계적 소양의 “소양(literacy)”은 비판적 소양을 포괄하므로 “정당화 없이 제기된 주장에 대한 비판적인 의문 제기”를 가장 높은 수준의 통계적 소양으로 제시한 Watson(1997)의 체계가 학교수학을 위한 통계적 소양의 적절한 틀이 될 수 있음을 밝혔다(p.298).

[표 1] 통계적 소양의 수준(Watson, 1997)
[Table 1] Tiers of statistical literacy (Watson, 1997)

1수준: 용어 이해하기
2수준: 맥락에서 용어 이해하기
3수준: 정당화 없이 제기된 주장에 대해 비판적인 의문 제기

“통계적 소양은 자료와 우연을 다루는 교육과정과 일상 세계의 교차점이다. 교차점은 미리 연습할 수 없는 맥락과 자동적으로 발생하는 의사결정과 관련이 있는데, 이는 모두 통계적 도구, 일반적인 맥락적 지식, 비판적 소양 기술을 적용하는 능력에 기반을 둔다.” (Watson, 2013)

통계학은 공통적으로 표본 자료를 기반으로 하며, 표본을 어떻게 추출하느냐에 따라서 자료의 질과 통계적 추론에 막대한 영향을 주기 때문에, 표본과 표집에 대한 이해는 통계적 소양의 매우 기초적인 요소이다(Watson & Moritz, 2000b, p.109). 이러한 이유로 표본과 표집에 대한 학생과 예비교사의 이해를 다룬 선행연구들은

Watson(1997)의 통계적 소양 수준을 분석 틀로 활용하곤 하였다. Watson & Moritz(2000a)는 통계적 소양의 각 수준에 따라 과제를 개발하여 3, 6, 9학년 학생들을 대상으로 반응 분석과 면담을 통해 표본 개념의 발달 과정을 분류하고 각 범주들을 통계적 소양 수준에 대응하였다. Groth & Bergner(2005)는 예비 초등교사를 대상으로 표본에 대한 은유를 직접 구성하게 하였고, 그 결과를 바탕으로 표본에 대한 사고의 범주를 나누어 제시한 후, 각 범주가 Watson(1997)의 통계적 소양 수준에서 어느 수준에 해당하는지를 확인하였다. 선행연구의 전례를 통해 Watson(1997)의 통계적 소양 수준은, 본 연구에서 표본 개념에 대한 예비교사들의 이해를 고찰하는데 적절한 틀로 활용될 것이라 기대할 수 있다.

III. 연구방법

1. 연구 대상

본 연구는 광역시 소재의 종합대학에 재학 중인 수학 교육과 학부생 39명을 대상으로 이루어졌다. 이 중 19명은 2학년이고 20명은 3학년인데, 본 연구를 위한 조사가 실행된 시점인 2학기 초를 기준으로 3학년 학생들은 학과 전공과목인 <확률 및 통계 입문>을 모두 수강한 반면, 2학년 학생들은 해당 과목을 수강하지 않은 상태였다. 따라서 연구 참여자 중 2학년과 3학년 학생 사이에는 예비교사를 대상으로 하는 교육과정 중 통계 과목의 수강 여부에 차이가 있음을 확인할 수 있었다.

다만, 본 연구의 참여자들은 편의표집을 통해 추출되었고 소수의 인원을 대상으로 연구가 수행되었기에 본 연구의 결과를 우리나라의 모든 예비교사로 일반화하여 적용할 수는 없다. 다만, 연구 참여자들 모두 동일한 학교에 재학 중이라는 점에서 우리나라의 예비교사 대상 통계교육의 실태를 보여주는 하나의 사례로서 역할을 할 수 있을 것으로 기대된다. 본 연구는 학교 통계교육과 예비교사교육의 개선 방향을 구체화하기 위한 통계적 소양 틀의 활용 가능성을 탐색하는 수준의 목표를 지니고 있다.

2. 자료 수집

본 연구에서는 선행연구를 바탕으로 하여 표본 개념

을 이해하기 위한 두 속성인 표본대표성과 표집변이성을 연구 내용의 중심으로 삼았다. 그리고 각각의 속성에 대한 예비교사들의 이해를 알아보기 위해 검사도구로서 외국의 교과서와 선행연구에서 발췌한 총 4개의 과제를 선정하였다. 본 연구에서는 과제를 수행해야 하는 주체가 예비교사인 점을 고려하여, 단순히 내용 지식으로 문제를 해결하는 것이 아니라 교수학습 상의 맥락 내에서 과제에 대한 메타적인 이해가 이루어지도록 발문을 수정하여 문항을 구성하였다. 교사의 과제 사용 방식과 교실 맥락에 대한 이해는 교사의 이해와 직접적으로 연결되며, 이는 차후 Hill et al.(2008)이 제시한 교과 지식 및 교수학적 내용 지식으로 결정화될 수 있다는 점에서 의미가 있다(Sullivan, Clarke, & Clarke, 2013; Sullivan, Zevenbergen, & Mousley, 2003). 일례로, Noll(2007)은 통계적 소양을 통계 교수 지식(statistical knowledge for teaching) 중, 내용 일반 지식(common content knowledge)으로 인식하여, 과제 활용이나 학생의 오개념 등 다양한 교수학적 맥락에서 대학원 연습교사의 통계적 소양을 확인하였다. 이렇게 설계된 4개의 문항에 대한 간략한 설명은 [표 2]에 제시되어 있으며, 실제 문항은 본고의 맨 뒷면 [부록]에 제시하였다. 이와 같이 선정된 4개의 과제로 이루어진 질문지를 2학기 초에 연구 참여자들에게 제공하였고 문항에 대한 자신의 생각을 자유롭게 서술하도록 하였다.

3. 분석 방법

연구 참여자들이 작성한 답안은 제1 저자가 1차적으로 분류하여 범주화하였다. 범주화는 각 문항의 관련 속성인 표본대표성 또는 표집변이성에 대한 연구 참여자의 인식 수준을 토대로 이루어졌으며, SOLO 모형을 활용한 Watson의 선행연구(e.g., Watson & Moritz, 2000a)를 참

고하였다. 이후 나머지 공저자들과 논의를 거쳐 재분류 및 재범주화 후 본 연구에서 제시된 범주들을 확정하였다. 각 범주에 해당하는 참여자의 수를 확인하여 그 분포표를 작성하였으며, 특히 예비교사를 대상으로 한 통계 과목의 수강 여부가 예비교사의 답안에 유의미한 차이를 가져오는지도 확인하고자 2학년 학생의 분포와 3학년 학생의 분포 또한 비교하였다. 범주별 분포는 비록 단일 학교의 사례이지만 표본 개념에 대한 예비교사들의 이해 실태를 확인하는 데 참고 자료가 될 수 있을 것으로 보인다. 본 연구의 결과를 토대로 논의를 이끌어내기 위해, 표집과 변이 개념에 대한 학생들의 이해를 범주화하여 제시했던 선행연구(Torok & Watson, 2000; Watson et al., 2003; Watson & Moritz, 2000a, 2000b)와도 비교하였고, 이에 따라 연구 참여자들에게 이루어진 예비교사 대상의 통계교육이 이들의 통계적 소양 발달에 어떤 영향을 주는지를 확인함으로써 현행 교사교육에 대한 시사점을 도출하였다.

IV. 결과 분석 및 논의

1. 표본 개념에 대한 예비교사의 이해

1) 표본대표성에 대한 이해

문항 1과 3은 각각, 표본과 표집의 관계에서의 표본 대표성, 그리고 모집단과 표본의 관계에서의 표본대표성에 대한 예비교사의 이해를 확인하고자 설계되었다. 즉, 문항 1은 “표본이 모집단을 대표하기 위해서는 임의표집과 같은 적절한 표집 기법이 필요하다는 사실”을, 문항 3은 “표본은 모집단의 단순한 부분집합이 아니라 유사비례적 축소 모형임(Saldanha & Thompson, 2002)”을 확인하기 위한 것이었다. 두 문항에 대한 연구 참여자들의

[표 2] 질문지 문항의 특성과 출처

[Table 2] Characteristics and resources of tasks in the questionnaire

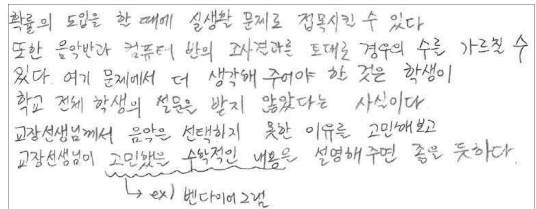
문항	관련 속성	조사 내용	출처
1	표본대표성	과제 활용에 대한 예비교사의 이해	MiC 교과서 『통계로 나타낸 세상』
2	표집변이성	과제 활용에 대한 예비교사의 이해	Coffey(2013)
3	표본대표성	학생의 오개념에 대한 예비교사의 이해	고은성 · 이경화(2011)
4	표집변이성	학생의 오개념에 대한 예비교사의 이해	Kahneman & Tversky(1972)

[표 3] 표본대표성에 대한 예비교사의 이해를 범주화한 결과
 [Table 3] Categories of preservice teachers' understanding of sample representativeness

<p>범주 R1. 표본의 대표성을 인지하지 않음</p> <ul style="list-style-type: none"> - 대표성에 대한 고려 없이 표본과 표집을 언급한다. - 표집 과제에서 통계와 관련이 없는 개념을 언급하기도 한다.
<p>범주 R2. 설명 없이 표본대표성을 언급</p> <ul style="list-style-type: none"> - 설명이나 정당화를 하지 않고 표본대표성이나 임의표집을 언급한다.
<p>범주 R3. 표집에서 발생하는 편의를 예측</p> <ul style="list-style-type: none"> - 표집자의 주관에 의한 표집에 대해 우려한다.
<p>범주 R4. 표본에 대한 불완전한 관점</p> <ul style="list-style-type: none"> - 표본조사 대신 전수조사를 해야 한다고 주장한다.
<p>범주 R5. 표본이 대표성을 가지도록 표집을 고려</p> <ul style="list-style-type: none"> - 표본대표성을 위한 임의표집을 언급한다. - 표본의 크기를 고려하기도 한다.
<p>범주 R6. 편의표본을 비판하고 임의표집을 제시</p> <ul style="list-style-type: none"> - 주어진 맥락 내에서 편의표본을 비판하고 임의표집을 언급한다.

반응에 드러나는 특징들을 나열한 후 이를 표본대표성에 대한 예비교사의 인식이라는 관점에서 범주화한 결과, [표 3]과 같은 범주들을 도출할 수 있었다.

범주 R1에 해당하는 연구 참여자들은 문항 1에서 편의표집 맥락이 제시된 과제를 활용하여 확률, 집합 등 수학적 개념을 가르칠 수 있다고 주장하거나, 문항 3에 대한 답안을 작성하지 못하는 등, 전반적으로 표본대표성에 대한 통계적 이해가 크게 부족한 것으로 드러났다. 일부는 표집에 문제가 있음을 막연하게 인지하고 있었지만 표본의 대표성과 관련된 어떠한 언급도 드러나지 않았다는 점에서 범주 R1으로 분류되기도 하였다. [그림 3]은 범주 R1에 해당하는 연구 참여자의 문항 1에 대한 반응이다. 문항 1에 포함된 과제에 대해 MiC 교과서 『통계로 나타난 세상』의 교사용 지도서에 따르면 “왜 표본조사한 것이 학교 전체 모집단을 대표할 수 없다고 하는지 설명하는” 것이 출제 의도였으나(p.37), 아래의 연구 참여자는 이 과제를 이용하여 확률 개념을 도입하고 벤다이어그램과 같은 수학적 개념을 가르칠 수 있다고 주장하였다.



[그림 3] 범주 R1로 분류된 반응의 예(문항 1)
 [Fig. 3] An example of responses in R1 (Item 1)

범주 R2에 해당하는 연구 참여자들은 문항 1이나 3에서 표본이 편의를 가지지 않고 모집단에 대한 대표성을 가지도록 표집이 이루어져야 한다고 언급하였다. 표본조사에서 표본이 모집단을 대표해야 한다는 문제인식은 분명히 드러나지만, 이 범주에 해당하는 연구 참여자들은 그러한 문제인식만을 언급할 뿐, 어떠한 설명이나 정당화도 시도하지 않고 있다. [그림 4]에서 볼 수 있듯이, 범주 R2에 해당하는 연구 참여자들은 표본대표성이 필요함은 인지하고 있으나 표집에 대한 비판적 소양을 갖출 만큼 표본대표성을 이해하고 있지는 않은 것으로 보인다.

학생들이 좋아하는 음악에 대해 설문 조사를 시행하는데 학생단체를 모집단이라 할 때, 그에 대한 표본집단을 뽑을 때 문제가 생겼다. 왜냐하면 철수와 영호의 반 선택이 잘못되었다. 따라서 표본집단에 대한 이해를 가라질 수 있다.

표본이 모집단의 부분 집합 여긴 확실한, 표본은 모집단을 대표하는 것이므로 모집단의 성질을 가지고 있어야 한다.

[그림 4] 범주 R2로 분류된 반응의 예(문항 1, 3)
[Fig. 4] An example of responses in R2 (Items 1 and 3)

범주 R3에 해당하는 연구 참여자들은 문항 3에서 과제 내의 표본이 대표성을 띠지 못하는 이유로 표집자의 주관을 언급하여 설명을 시도한다는 점에서 범주 R2와 구분된다. 실제로 표본조사에서 표집자의 주관은 임의표집 대신 목적표집이 이루어질 가능성을 높이면서 표본이 모집단을 대표하도록 하는데 장애가 될 수 있다. 다만, 이 범주에 해당하는 연구 참여자들은 [그림 5]에서도 볼 수 있듯이, 표본대표성에 대한 이해가 임의표집이라는 핵심적인 통계적 개념과의 연결까지 나아가지는 못한 것으로 보인다.

우리가 설문조사를 할 예를 보면 크기에나 같은 질문이 설문조사를 한 거. 그것이 필요한 표본이 된다.

하지만 위의 경우에도 크기가 다른 2종이기 때문에 손으로 직접 ~~뽑~~ 꺼낼 때 크나 작기를 ~~2~~ 구별할 수 있기 때문에 ~~구분~~ 표본이 ~~뽑~~ 뽑을 수 없다. 모집단의 표본이나 할 수 없다.

[그림 5] 범주 R3으로 분류된 반응의 예(문항 3)
[Fig. 5] An example of responses in R3 (Item 3)

범주 R4에 해당하는 연구 참여자들은 통계적 조사에 대한 예비교사들의 부정적인 시각을 드러낸다. 이 범주에 해당하는 모든 참여자들은 문항 1, 3에서 제시된 과제를 통해, 표본조사는 신뢰할 수 없으며 정확한 조사를 위해서는 전수조사를 실시해야 한다고 주장한 것이다. Watson(2010)의 연구에서는 장기 면담이 이루어진 38명의 학생들 중 단 한 명만이 표본에 대한 기본적인 개념을 알고 있음에도 모집단 전체가 표집되어야 한다는 주장을 하였다. 그러나 본 연구에서는 39명의 예비교사 중 7명이 [그림 6]과 같은 관점을 취하고 있었다.

많은 학생의 의견이 아닌 일부 학생들의 대답으로 조사한 것 때문에 조사 대상에게서 객관적이지 못하다고 생각한다. 따라서 이 문제를 통해 학생들에게 통계를 내는 때, 모든 경우에 대해 조사해야 한다고 가르칠 수 있다.

[그림 6] 범주 R4로 분류된 반응의 예(문항 1)
[Fig. 6] An example of responses in R4 (Item 1)

범주 R5에 해당하는 연구 참여자들은 표본이 모집단을 대표하기 위해서 임의표집이 이루어져야 한다는 주장을 하였고, 일부는 문항 1에서 과제에 등장하는 두 표본의 크기가 명시되지 않았다는 점을 들어, 표본의 크기에 따라 표본의 대표성이 달라질 수 있음을 언급하였다. 표집 맥락에 대한 비판의 적절성과는 별개로 표본대표성을 위해 임의표집이나 표본 크기와 같은 적절한 통계적 개념을 고려했다는 점이 앞서의 범주들과는 다른 특징이다. [그림 7]의 위는 문항 1에 대해 표본 크기를 통해 표집의 적절성에 문제를 제기한 사례이며, 아래는 문항 3에 대해 표집의 임의성이 전제되어야 함을 주장한 사례이다.

일반 학생과 컴퓨터 학생의 총 인원수가 달라서 결론은 다르다.

조각은 ~~라~~ 결과 ~~부~~ 표본 ~~조사~~ ~~과정~~ 변태

예를 들어 음악반이 10명이었고 10명이 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~

100% 인에 컴퓨터 반이 (10명인) 100% 인에 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~

100% 이므로 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~ 음악반에 10명 반대로 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~

컴퓨터반은 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~ 인원은 인원의 수 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~ ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~

과정에서 차이가 아닌 두칸 ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~ ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~ ~~조~~ ~~조사~~ ~~과정~~ ~~과정~~

(예를 들어, 학생들의 성격을 확인하는데 상위 10%인 학생들을 뽑아 성격을 확인하면 집단이 학생들의 성격보다 훨씬 높은 성격이 된다.)

[그림 7] 범주 R5로 분류된 반응의 예(문항 1, 3)
[Fig. 7] An example of responses in R5 (Items 1 and 3)

범주 R6에 해당하는 연구 참여자는 범주 R5에서 더 나아가 문항에서 제시한 과제의 맥락 내에서 이루어진 편의표집 과정을 적절히 비판하고 임의표집을 대안으로 제시함으로써, 표본대표성을 비롯하여 관련된 주요 개념들을 이해하고 있음은 물론, 맥락 내에서의 적절한 비판적 소양까지 갖추었음을 보여주었다. [그림 8]은 범주 R6에 해당하는 유일한 참여자의 문항별 반응이다.

모집단의 견제, 표본의 견제, 임의추출에 대하여 가르칠 수 있다.
 뒤의 장에서 원래 당의 표본의 특성을 모집단과 다를 수 있고,
 원가가 뽑은 임의 표본은 모든 경우가 뽑은 임의 표본을 대표하지 못한다.
 그런데 위의 견제 받고 표본의 특성을 식당에서 어떤 수밖에
 틀어주어야 한시 견제하지 못하였는데 이는 원가. 당과 표본의 것을 견제에서
 임의적인 추출이 아니라 각 대상의 추출된 특성이 동일하지 못했기 때문이다.
 표본은 추출할 때 각자의 특성을 다르게 각 대상이 추출된 특성이
 동일하게 추출하는 것이 추출된 특성이 동일하지 못했을 경우를 일컫는다.

표본의 모집단과 다른 조건의 원인이 다른 것이지만,
 표본 추출된 모든 조건의 모집단과 다른 각자의 대상은 추출된 특성이 같아야한다.
 뒤에 각 원가 기준은 뽑은 것이나 같은 두 가지 기준이 존재하는데
 이는 식중에 따라 크기가 다른 구별되었다.
 원가가 기준은 뽑은 미 기준은 기준이 되었기에 식중에 뽑은 수는 없었지만
 혼합 반대로 기준 크기를 느낄 수 있었기 때문에 원가의 기준이 틀어질 수 있고
 각 대상이 추출된 특성이 같지 않다.
 원가는 같은 일이고 표본의 개념도 정의하였기 때문에 크게 것은 가치가 되었다.

[그림 8] 범주 R6으로 분류된 반응의 예(문항 1, 3)
 [Fig. 8] An example of responses in R6 (Items 1 and 3)

이와 같이 문항 1, 3에 대한 연구 참여자들의 반응을 토대로 표본대표성에 대한 예비교사의 이해를 6개의 범주로 분류하였다. [표 4]는 각 범주에 해당하는 연구 참여자의 수를 분포표로 나타낸 것인데, 참여자의 약 60%

[표 4] 표본대표성에 대한 예비교사의 이해 범주별 분포
 [Table 4] Number of participants by categories about understanding of sample representativeness

범주	2학년	3학년	전체
R1. 표본의 대표성을 인지하지 않음	5	6	11
R2. 설명 없이 표본대표성을 언급	5	7	12
R3. 표집에서 발생하는 편의를 예측	1	2	3
R4. 표본에 대한 불완전한 관점	5	2	7
R5. 표본이 대표성을 가지도록 표집을 고려	3	2	5
R6. 편의표본을 비판하고 임의표집을 제시	0	1	1
합계	19	20	39

가 범주 R1과 R2로 분류되었음을 확인할 수 있었다. 또한, 범주 R6으로 분류된 참여자는 학부 통계 과목을 수강한 3학년 중에서도 단 1명이라는 점에 주목해 볼 필요가 있다.

2) 표집변이성에 대한 이해

문항 2와 4는 각각, 표집변이성의 편제성, 그리고 통계적 추정에서 표본대표성과 표본 크기의 관계에 대한 예비교사의 이해를 확인하고자 설계되었다. 즉, 문항 2는 “표집 기법과는 별개로 단일 모집단에서 서로 다른 다양한 표본이 추출될 수 있음”을, 문항 4는 “통계적 추정에서 표집변이성의 제어에 필요한 표본의 크기는 모집단의 크기와 무관함”을 확인하기 위한 것이었다. 표본대표성과 마찬가지로 두 문항에 대한 연구 참여자들의 반응에서 드러나는 특징들을 나열한 후 표집변이성에 대한 예비교사의 인식이라는 관점에서 분류하여, [표 5]와 같은 범주들을 도출하였다.

범주 V1에 해당하는 연구 참여자들은 문항 2에서 우연에 의한 변이성 개념과 목적표집에 의한 불공정성 개념을 혼동하거나, 문항 4에서 모집단의 크기가 클수록 표본조사의 결과를 예측하기 쉽다고 주장하는 등, 전반적으로 표집변이성에 대한 통계적 이해가 크게 부족한 것으로 드러났다. 그 외에도 문항에 대해 어떠한 반응도 기재하지 않거나 문제를 잘못 이해한 경우도 범주 V1으로 분류하였다. [그림 9]의 위, 아래는 각각 범주 V1에 해당하는 연구 참여자들의 문항 2, 4에 대한 반응 사례이다.

[표 5] 표집변이성에 대한 예비교사의 이해를 범주화한 결과
 [Table 5] Categories of preservice teachers' understanding of sampling variability

<p>범주 V1. 표집변이성을 문제로 인식</p> <ul style="list-style-type: none"> - 변이성이 표집의 공정성을 방해한다고 인식한다. - 표본의 크기가 아니라 모집단의 크기를 고려한다. - 표집 과제에서 통계와 관련이 없는 개념을 언급하기도 한다.
<p>범주 V2. 통계적 추리를 고려하지 않고 표집변이성을 인식</p> <ul style="list-style-type: none"> - 변이성이 표집에 편재함을 언급한다. - 확률은 절대로 변하지 않는 것이라 주장하기도 한다.
<p>범주 V3. 표본 크기와의 관계를 고려하면서 표집변이성을 인식</p> <ul style="list-style-type: none"> - 통계적 추리에서 표본의 크기가 중요하다고 주장한다. - 큰 수의 법칙을 언급하기도 한다.
<p>범주 V4. 통계적 추리를 위한 분포 개념을 고려하면서 표집변이성을 인식</p> <ul style="list-style-type: none"> - 표집변이성, 표집분포, 신뢰구간의 관계를 이해한다.

(a) 공정하지 않다.
 (b) 주사위를 굴릴 때마다 인차의 범위가 넓다.
 (c) 공정치 않다. 왜냐하면 오차의 범위가 넓기 때문이다.
 인차 많을수록 정확하게 된다.
 표본은 모집단에서 뽑은 자료의 일부인데 표본이 많을수록
 자료의 대상이 많아서 더 때문에 더 정확해지고 믿을 만하고
 그만큼 이해하기 쉬울 것이다.

[그림 9] 범주 V1로 분류된 반응의 예(문항 2, 4)
 [Fig. 9] An example of responses in V1 (Items 2 and 4)

범주 V2에 해당하는 연구 참여자들은 문항 2에 제시된 과제에서 수학적 확률과 통계적 확률의 차이에 주목하였다. [그림 10]에서처럼 표집변이성의 편재로 인해 수학적으로 모델링된 선험적 확률과 통계적으로 산출된 경험적 확률 사이에 차이가 존재함을 언급함으로써, 범주 V1과는 달리 표집변이성이 표본조사에서 자연스러운 것임을 인식하고 있음을 드러내었다. 다만, 이 범주에 속한 참여자들 중 일부는 표집의 공정성을 뒷받침하기 위한 근거로 수학적 확률의 절대성을 제시함으로써, 통계적

추리에 대한 고려가 이루어지지 않고 있음을 보여주었다.

인차의 주사위를 던져서 관념으로 순재가 나오게 하는 것을 직접 경험하게 되어, 순재가 나오는 확률이 더 많지는 않다는 것을 안 수 있게 한다.

[그림 10] 범주 V2로 분류된 반응의 예(문항 2)
 [Fig. 10] An example of responses in V5 (Item 2)

범주 V3에 해당하는 연구 참여자들은 문항 4를 통해 통계적 추리에서 표본의 크기가 중요한 요소임을 주장하였다. 즉, 표본의 크기가 커질수록 표집변이성이 작아지면서 표본조사의 신뢰도가 높아진다는 사실을 언급한 것이다. 다만, 후술할 범주 V4와 다른 점은 이들이 선행연구에서 지적한대로 여전히 표본의 크기를 고려할 때 모집단의 크기를 반영하려 든다는 것이다. [그림 11]에서처럼 표본의 크기 자체가 아니라 표본이 모집단에서 차지하는 비율을 중요하시는 경향을 보이고 있는데 (Bar-Hillel, 1979), 이는 여전히 다수의 표본으로 구성된 표집분포와 단일 표본 내의 변량으로 구성된 표본분포를 혼동하고 있기 때문인 것으로 보인다.

대도시 A와 소도시 B의 경제는 인구를 보라. A에서 1000명보다 B에서 100명이 전체 인구의 몇 퍼센트인지 알지 못한다. 따라서 두 도시의 추세를 비교할 수 없기 때문에 경제적 구분은 옳지 않다. 모집단에 대한 표본의 비율이 직관적으로 잘못된 예측을 내릴 가능성이 높을 수 있다.

[그림 11] 범주 V3으로 분류된 반응의 예(문항 4)
[Fig. 11] An example of responses in V3 (Item 4)

범주 V4에 해당하는 단 한 명의 연구 참여자는 통계적 추리에서 표본의 크기를 고려하는데 모집단의 크기 대신 신뢰구간을 언급함으로써 표집분포에 대한 이해를 보여주고 있다는 점에서 범주 V3과 차이가 있다. 신뢰구간은 표집분포의 중심, 퍼짐, 모양 중 표본의 크기에 의존하는 퍼짐을 수치화한 통계적 개념이다(Lipson, 2003). 표집분포가 통계적 추리에서 표집변이성을 바라보는 렌즈로서의 도구적 성격을 지니고 있다는 점을 고려하면, 이 범주에 해당하는 참여자는 표집변이성을 인식할 때 통계적 추리와 분포 개념을 고려하고 있음을 보여주고 있다고 할 수 있다. [그림 12]가 범주 V4로 분류된 이 참여자가 문항 4에 대해 작성한 반응이다.

모집단에서 표본을 뽑는다. 이 때 표본의 크기가 클수록 신뢰구간이 작아진다. 즉, 오차의 범위가 줄어들기 때문에 더 신뢰할 수 있고, 더 정확하게 측정할 수 있다.

[그림 12] 범주 V4로 분류된 반응의 예(문항 4)
[Fig. 12] An example of responses in V4 (Item 4)

이와 같이 문항 2, 4에 대한 연구 참여자들의 반응을 토대로 표집변이성에 대한 예비교사의 이해를 4개의 범주로 분류하였다. [표 6]은 각 범주에 해당하는 연구 참여자의 수를 분포표로 나타낸 것인데, 참여자의 약 70%가 범주 V1과 V2로 분류되었음을 확인할 수 있었으며,

[표 6] 표집변이성에 대한 예비교사의 이해 범주별 분포
[Table 6] Number of participants by categories about understanding of sampling variability

범주	2학년	3학년	전체
V1. 표집변이성을 문제로 인식	6	8	14
V2. 통계적 추리를 고려하지 않고 표집변이성을 인식	8	6	14
V3. 표본 크기와의 관계를 고려하면서 표집변이성을 인식	5	5	10
V4. 분포 개념을 고려하면서 표집변이성을 인식	0	1	1
합계	19	20	39

이 범주 내에서도 표본대표성과 마찬가지로 통계 과목의 수강 여부는 유의미한 변수로 작용하지 못했음을 알 수 있었다.

2. 표본 개념에 대한 예비교사의 통계적 소양

1) 표본대표성에 대한 통계적 소양

Watson & Moritz(2000a, 2000b)는 질문지와 면담을 바탕으로 표집 개념에 대한 학생들의 이해를 여섯 가지로 범주화하였고, 드러난 범주와 통계적 소양 수준(Watson, 1997) 사이의 관계를 파악해보았다. Watson과 Moritz의 범주는 [표 7]과 같다(Watson & Moritz, 2000a, p.54).

Watson & Moritz(2000a, 2000b)의 연구에서 표집(sampling)이란 통계적 조사(statistical investigation)의 두 번째 단계인 ‘자료 수집’에 해당하는 과정(process)이므로, 통계의 전 과정에서 걸쳐 영향을 미치는 임의성, 독립성, 불규칙적 변이성, 분포적 항상성 등을 상징하는 의미(이영하·신수영, 2011, p.328)로서 표본(sample) 개념을 사용한 본 연구와는 다소 차이가 있다.

그러나 표본대표성이 표본 개념에 선천적으로 내재되어 있는 것이 아니라 표집 기법에 의존하여 제한적으로 확보된다는 사실(탁병주 외, 2014, p.736)에 비추어볼 때, 표집에 대한 이해를 바탕으로 한 Watson & Moritz의 범주들은 표본대표성과 관련 요소들을 내포하므로, 본 연구의 표본대표성 범주와 유의미한 비교가 가능하다고 할 수 있다. 이에 본 연구에서의 표본대표성 범주들과 Watson & Moritz(2000a, 2000b)의 범주, 그리고 Watson(1997)의 통계적 소양 수준을 [표 8]과 같이 대응하였다.

[표 7] 표집 개념 발달에 대한 범주와 통계적 소양 수준

[Table 7] Categories of developing concepts of sampling with respect to tiers of statistical literacy

<p>1수준: 용어 이해하기</p> <p>범주 1. 표집 과정을 거치지 않고 작은 크기의 표본을 선정</p> <ul style="list-style-type: none"> - 식품과 같은 것으로 표본의 예를 제시 - ‘작은 부분’으로 표본을 설명하기도 하며, 드물게 ‘시도/시범’으로 설명하기도 함 - 15보다 작은 크기의 표본을 허용 - 표집 방법을 제시하지 않거나 주관적인 방법을 제안 <p>범주 2. 원시적인 임의표집으로 작은 크기의 표본을 선정</p> <ul style="list-style-type: none"> - 식품과 같은 것으로 표본의 예를 제시 - ‘작은 부분’ 또는 ‘시도/시범’으로 표본을 설명 - 15보다 작은 크기의 표본을 허용 - 설명 없이 ‘임의표집’을 제안하거나 ‘여러 학교들에서~’와 같이 임의적 선정에 관한 간단한 표현 사용 <p>2수준: 맥락에서 용어 이해하기</p> <p>범주 3. 예측 결과를 고려하여 작은 크기의 표본을 선정</p> <ul style="list-style-type: none"> - 식품과 같은 것으로 표본의 예를 제시 - ‘작은 부분’이나 ‘시도/시범’ 모두 이용하여 표본을 설명 - 15보다 작은 크기의 표본을 허용 - 똥똥하고 마른 정도를 고려하거나 표준 체중을 가진 사람들을 고려하는 것처럼, 체중에 따라 사람들을 선정할 것을 제안 <p>범주 4. 불안정한 방법으로 표본을 선정</p> <ul style="list-style-type: none"> - 표본의 예를 제시하고 표본을 설명 - 간혹 부적절한 측면을 근거로 표본의 크기에 관심을 가지지 않기도 함 - 적절한 표집이나 부분적으로 편의에 민감한 방법을 사용하면서 크기가 작은 표본을 제안하거나, 또는 부적절한 방법을 사용하면서 크기가 큰 표본을 제안 <p>범주 5. 임의표집 혹은 분포를 고려한 표집으로 큰 크기의 표본을 선정</p> <ul style="list-style-type: none"> - 식품과 같은 것으로 표본의 예를 제시 - ‘작은 부분’이나 ‘시도/시범’ 모두 이용하여 표본을 설명 - ‘평균’이라는 용어를 언급하기도 함 - 표본의 크기로 20 이상의 수 또는 모집단의 비율을 이용 - 임의표집이나 지역 분포에 기초하여 표집방법을 제안 <p>3수준: 정당화 없이 제기된 주장에 대해 비판적인 의문 갖기</p> <p>범주 6. 편의에 민감하게 큰 크기의 표본을 선정</p> <ul style="list-style-type: none"> - 표본의 예로 조사과정을 포함하는 예를 제시 - ‘작은 부분’이나 ‘시도/시범’ 모두 이용하여 표본을 설명 - ‘평균’이나 ‘대표성’이라는 용어를 언급하기도 함 - 표본의 크기로 20 이상의 수 또는 모집단의 비율을 이용 - 임의표집이나 지역 분포에 기초하여 표집방법을 제안 - 편의를 피하기 위한 표집방법에 관심을 보임 - 조사 결과를 보고하는 신문 기사에서 편의가 있는 표본을 식별

본 연구에서 도출한 범주 R1은 표본의 대표성을 인지하지 않아 적절한 표집 과정에 대한 고려가 이루어지지 않았다는 점에서, 표집 방법을 제시하지 않거나 주관적

인 방법을 제시하는 Watson & Moritz(2000a)의 범주 1에 대응할 수 있다. 범주 R2는 설명 없이 표본대표성과 임의표집을 언급한다는 점에서 범주 2에, 범주 R3은 편

의를 고려한다는 점에서 범주 3에 대응시켰다. 범주 R4의 경우, “표본조사 대신 전수조사를 실시해야 한다”는 대표적인 반응이 실제 Watson(2010)의 연구에서 범주 4로 분류된 바 있다. 범주 R5는 표본이 대표성을 가지도록 다양한 표집 방법을 고려했다는 점에서 범주 5에, 마지막으로 범주 R6은 편의표본에 대한 적절한 비판적 의문이 포함되었다는 점에서 범주 6에 대응하였다.

본 연구에서 예비교사를 대상으로 한 표본대표성의 이해에 대한 범주와 Watson & Moritz(2000a, 2000b)의 연구에서 학생들을 대상으로 한 표집 이해에 대한 범주는, 다루고 있는 내용에 다소 차이가 있어 온전히 일치한다고 볼 수는 없다. 가령, Watson & Moritz의 범주화에서 중요하게 다루어지고 있는 표본의 크기에 대한 이해를, 본 연구에서는 예비교사들의 반응을 범주화하는데 고려하지 않았다. 그러나 Watson & Moritz의 연구에서

표본의 크기와 함께 중요하게 다루어지고 있는 임의표집과 편의는 본 연구에서 표본대표성에 대한 이해를 설명하는데 매우 핵심적인 요소인 바, 이를 바탕으로 본 연구의 범주와 Watson & Moritz 연구의 범주는 구조적 유사성을 지닌다고 볼 수 있다. 따라서 본 연구에서 도출한 범주들은 Watson & Moritz 연구에서의 범주를 매개로 하여 [표 8]에서 제시한 것처럼 Watson(1997)의 통계적 소양 수준과도 대응할 수 있는데, 이에 따라 본 연구의 참여자들이 표본대표성에 대한 이해에서 어느 정도의 통계적 소양 수준에 도달해있는지를 [표 9]와 같이 확인할 수 있다.

위 결과에 따르면 표본대표성에 대한 예비교사들의 통계적 소양은 연구 참여자들 중 약 60%가 1수준에 불과했고, 최고 수준인 3수준에 도달한 참여자는 단 1명뿐이었다. 그리고 예비교사 대상의 통계 과목을 소속 대학

[표 8] 표본대표성에 대한 이해 범주와 Watson & Moritz(2000a) 범주의 관계
 [Table 8] Categories of understanding of sample representativeness and Watson & Moritz(2000a)'s categories

표본대표성에 대한 이해	표집 개념에 대한 이해(Watson & Moritz, 2000a)	통계적 소양 수준(Watson, 1997)
R1. 표본의 대표성을 인지하지 않음	범주 1. 표집 과정을 거치지 않고 작은 크기의 표본을 선정	1수준. 용어 이해하기
R2. 설명 없이 표본대표성을 언급	범주 2. 원시적인 임의표집으로 작은 크기의 표본을 선정	
R3. 표집에서 발생하는 편이를 예측	범주 3. 예측 결과를 고려하여 작은 크기의 표본을 선정	2수준. 맥락에서 용어 이해하기
R4. 표본에 대한 불완전한 관점	범주 4. 불완전한 방법으로 표본을 선정	
R5. 표본이 대표성을 가지도록 표집을 고려	범주 5. 임의표집 혹은 분포를 고려한 표집으로 큰 크기의 표본을 선정	
R6. 편의표본을 비판하고 임의표집을 제시	범주 6. 편의에 민감하게 큰 크기의 표본을 선정	3수준. 정당화 없이 제기된 주장에 대해 비판적 의문 제기

[표 9] 표본대표성에 대한 예비교사의 통계적 소양 수준
 [Table 9] Tiers of participants' statistical literacy about understanding of sample representativeness

범주	2학년	3학년	전체
1수준. 용어 이해하기	10	13	23
2수준. 맥락에서 용어 이해하기	9	6	15
3수준. 정당화 없이 제기된 주장에 대해 비판적인 의문 제기	0	1	1
합계	19	20	39

에서 수강한 3학년 참여자들 중에서 오히려 2학년 참여자들에 비해 통계적 소양의 1수준에 불과한 학생들이 다소 많았다. 이는 확률 및 통계를 비롯한 전공과목에서 학습한 형식적인 개념을 활용하는 과정에서 과제에 주어진 맥락과 유리되었기 때문으로 추측해 볼 수 있다.

대체적으로 3학년 학생들은 2학년 학생들에 비해 확률, 임의성, 큰 수의 법칙과 같은 형식적인 개념들을 활용하여 답을 하려는 경향을 보였다. 반면에 2학년 학생들은 과제의 맥락에 비추어 비형식적인 설명을 시도하곤 하였다. 이전 장의 [그림 4]가 3학년 학생의 반응이었는데, 표본조사라는 현실적 맥락을 고려하지 않고 확률, 벤다이어그램과 같은 추상적인 개념이 언급되어 통계적 소양의 1수준에 해당하는 범주 R1으로 분류되었다.

2) 표집변이성에 대한 통계적 소양

한편, Watson & Moritz(2000a, 2000b)의 연구와 유사하게 변이 개념에 대한 학생들의 이해를 범주화한 연구 또한 Watson과 그 동료들을 중심으로 이루어졌다(Torok & Watson, 2000; Watson et al., 2003). 이에 따르면 변이 개념에 대한 학생들의 이해는 네 가지로 범주화되며 이 역시 Watson(1997)의 통계적 소양 수준에 대응할 수 있는데, 그 결과는 [표 11]과 같다(Torok & Watson, 2000, p.155; Watson et al., 2003, p.20).

Reading & Shaughnessy(2010)는 변이성을 변화하는 실체의 특징으로, 변이를 변이성의 기술, 측정 결과로 구분하지만(p.202), 학생이나 교사가 통계교육에서 인식해야 하는 대상으로서의 변이 개념은 대개 변이성과 구별

되지 않고 사용되기도 한다(지은정, 2011; Moore, 1990). 따라서 학생들의 변이 개념 발달에 대한 Torok & Watson(2000)의 연구는, 본 연구에서 다룬 표집변이성 개념을 포괄하고 있다. 이에 본 연구에서의 표집변이성 범주들과 Torok & Watson(2000), Watson et al.(2003)의 범주, 그리고 Watson(1997)의 통계적 소양 수준을 [표 10]과 같이 대응하였다.

본 연구에서 도출한 범주 V1은 변이성에 대한 이해와 인식이 매우 약하다는 점에서, 변이가 내재된 실험 결과에 중요하고 변이를 명시적으로 언급하지 못하는 Torok & Watson(2000)의 범주 A에 대응할 수 있다. 범주 V2는 변이성의 편재를 인정하나 통계적 추리라는 사고 과정 내에서 변이성을 인식하는 것이 아니라 변이 현상 자체만을 지엽적으로 바라본다는 점에서 범주 B에 대응시켰다. 범주 V3은 표본 크기와 관계를 고려하는 과정에서 강한 비례적 사고가 드러난다는 점에서 범주 C에, 범주 V4은 분포 개념을 통해 군집 내 변이성에 대한 일관되면서도 형식적인 이해를 갖추었다는 점에서 범주 D에 대응하였다.

Torok & Watson(2000)의 연구에서는 표집 맥락에서의 변이 외에도 추정 변이나 자연 변이 등 통계 전반에 걸친 다양한 변이들을 다루어 학생들의 이해를 범주화하였기 때문에, 본 연구에 비해 훨씬 포괄적인 내용 영역을 담고 있다. Torok & Watson의 연구에서 다루는 변이 개념 중 표집변이성에 대한 것만을 토대로 본 연구에서 표집변이성에 대한 예비교사들의 이해를 범주화한 것과 비교해볼 때, 앞서의 표본대표성과 마찬가지로 본 연

[표 10] 표집변이성에 대한 이해 범주와 Torok & Watson(2000) 범주의 관계
[Table 10] Categories of understanding of sampling variability and Torok & Watson(2000)'s categories

표집변이성에 대한 이해	변이 개념에 대한 이해(Torok & Watson, 2000)	통계적 소양 수준(Watson, 1997)
V1. 표집변이성을 문제로 인식	범주 A. 변이에 대한 약한 인식	1수준. 용어 이해하기
V2. 통계적 추리를 고려하지 않고 표집 변이성을 인식	범주 B. 변이와 군집의 일부 측면만 고립적으로 인식	2수준. 맥락에서 용어 이해하기
V3. 표본 크기와 관계를 고려하면서 표집변이성을 인식	범주 C. 변이와 군집에 대한 비일관된 인식	
V4. 분포 개념을 고려하면서 표집변이성을 인식	범주 D. 변이와 군집에 대한 일관된 인식	3수준. 정당화 없이 제기된 주장에 대해 비판적 의문 제기

[표 11] 변이 개념 발달에 대한 범주와 통계적 소양 수준

[Table 11] Categories of developing concepts of variability with respect to tiers of statistical literacy

<p>1수준: 용어 이해하기</p> <p>범주 A. 변이에 대한 약한 인식</p> <ul style="list-style-type: none"> - 변이 인정 - 비례 아이디어에 대해 매우 약한 이해를 보이는 반응 제시 - 전체 집단에 대한 고려 없이 개별 결과에 주목 - 가장 일반적인 개별값으로 평균 언급 - 변이와 군집의 비일관적인 수준의 답변 - 실험 결과에 동요함 - 유의미한 요약 그래프를 그리지 못함 - 변이를 명시적으로 언급하지 못하며, 변이라는 용어에 대해 매우 빈약한 지식을 보임 - 실세계 상황에 대한 일반적인 지식이 부족함 <p>2수준: 맥락에서 용어 이해하기</p> <p>범주 B. 변이와 군집의 일부 측면만 고립적으로 인식</p> <ul style="list-style-type: none"> - 변이를 기꺼이 인정 - 비례 아이디어에 대해 매우 약한 이해를 보이는 반응 제시 - 특정값의 부분범위 관점에서 답변 - 일반적인 값의 범위 내의 값으로서 평균을 언급 - 일관되게 너무 많은 혹은 너무 적은 변이와 군집으로 답변 - 실험 결과에 적당히 동요함 - 요약 그래프를 그리려 시도하나 유의미한 것을 그리지 못함 - 변이를 명시적으로 언급하지 않지만 변이 용어에 대해 합리적인 지식을 가지고 있음 - 실세계 상황에 대한 가변적 지식을 가지고 있음 <p>범주 C. 변이와 군집에 대한 비일관된 인식</p> <ul style="list-style-type: none"> - 변이를 기꺼이 인정 - 강한 비례적 사고를 보여주며 완전표본과 같은 대표성을 시사하는 반응을 제시 - 결과 집합의 맥락에서 특정 결과에 대한 답을 함 - 일관되게 너무 많은 혹은 너무 적은 변이와 군집으로 답변 - 실험 결과에 약간만 영향을 받음 - 자료 요약을 위해 시계열 그래프와 동치인 것을 그림 - 변이를 명시적으로 언급하며 변이 용어에 대한 강력한 지식 - 실세계 상황에 대한 기초적인 일반 지식 <p>3수준: 정당화 없이 제기된 주장에 대해 비판적인 의문 갖기</p> <p>범주 D. 변이와 군집에 대한 일관된 인식</p> <ul style="list-style-type: none"> - 변이를 기꺼이 인정 - 비례 아이디어 간 갈등을 제시하는 반응을 보임; 혹은 강한 비례적 사고를 보이며 대표성을 시사하는 반응을 제시 - 결과 집합의 맥락에서 특정 결과로서 답을 함 - 군집의 적절한 수준을 가지고 일관되게 답을 함 - 실험 결과에 약간만 영향을 받음 - 자료 요약을 위해 빈도그래프 혹은 시계열 그래프를 그림 - 변이를 명시적으로 언급하여 일반적으로 변이 용어에 대한 강력한 지식 - 실세계 상황에 대한 좋은 일반적 지식
--

구의 범주와 Torok & Watson 연구의 범주는 구조적으로 유사하다. [표 12]는 본 연구의 참여자들이 보여준 표

집변이성에 대한 이해를, Watson(1997)의 통계적 소양 수준에 대응하여 그 수준을 표로 제시한 것이다.

[표 12] 표집변이성에 대한 예비교사의 통계적 소양 수준
 [Table 12] Tiers of participants' statistical literacy about understanding of sampling variability

범주	2학년	3학년	전체
1수준. 용어 이해하기	14	14	28
2수준. 맥락에서 용어 이해하기	5	5	10
3수준. 정당화 없이 제기된 주장에 대해 비판적인 의문 제기	0	1	1
합계	19	20	39

위 결과에 따르면 표집변이성에 대한 예비교사들의 통계적 소양은 연구 참여자들 중 70% 이상이 1수준에 불과했는데, 이는 표본대표성보다도 더 높은 수치다. 또한, 통계적 소양의 각 수준에 해당하는 2학년과 3학년 참여자의 수에 큰 차이가 없는 연구 결과가 도출되었다. 이는 예비교사 대상의 통계 과목 수강 여부가 표집변이성에 대한 통계적 소양 수준에 유의미한 변수로 작용하지 않았다는 점을 의미한다. 기존의 선행연구들이 통계교육에서 표집변이성의 중요성을 강조하고 있으나(예를 들어, 고은성, 2012; 구나영 외, 2015; Pfannkuch, 2008; Pfannkuch et al., 2015 등), 대학에서 이루어지고 있는 예비교사 대상 통계교육 체계에서는 이러한 연구 결과들이 잘 반영되지 않는 것으로 보인다.

V. 결론 및 제언

본 연구에서는 예비교사를 대상으로 표본 개념에 대한 이해를 살펴보았다. 이를 위해 표본 개념에 대한 이해의 분석 틀로 표본대표성과 표집변이성을 선정한 후, 선행연구와 외국 교과서를 바탕으로 문항을 설계하여 각 속성에 대한 예비교사의 이해를 확인하였다. 문항 반응에 나타난 예비교사의 표본 개념 이해를 표본대표성 이해와 표집변이성 이해로 나누어 범주화한 후, 각 범주들을 통계적 소양 수준에 비추어 통계적 소양 수준에 대응시켰다.

연구 결과를 통해 표본 개념에 대한 예비교사들의 통계적 소양 수준은 오늘날 학교수학에서 강조되고 있는 비판적 소양에까지 다다르지 못하고 있음을 확인하였다. 특히, 표본대표성보다 표집변이성에 대한 이해가 더욱 부족하다고 나타났으며, 예비교사를 대상으로 한 통계 과목이 통계적 소양에 유의미한 변화를 이끌어내지 못하는 것으로 보인다. 이와 같은 결과들은 학교수학에서 이

루어지는 통계교육의 개선을 위한 예비교사교육에 여러 가지 시사점을 제공한다. 이를 토대로 본고에서는 예비교사를 대상으로 하는 통계교육에 다음과 같은 제언을 함으로써 결론을 대신한다.

첫째, 표본 개념을 이해하는데 예비교사들이 겪는 어려움과 그 원인에 대한 후속 연구가 이루어져야 한다. 이는 고은성·이경화(2011)도 일찍이 지적한 바 있는데, 본 연구에서도 표본 개념에 대한 예비교사들의 이해 수준은 통계적 소양 수준에 비추어볼 때 여전히 맥락이나 비판적 소양과 유리된 채로 낮은 수준에 머물러 있음을 확인할 수 있었다. 특히, 범주 R4에 해당하는 반응에서 확인할 수 있었던 예비교사들의 반응은 기존 선행연구(Watson, 2010)에서 특이한 사례로서 다루어졌던 것인데, 표본조사를 불신하고 전수조사를 수행해야 한다는 주장은 예비교사의 통계적 지식 외에도 통계에 대한 신념이나 태도와 관련된 것이다. 따라서 우리나라 예비교사와 현직교사들이 통계에 대해 어떠한 신념과 태도를 가지고 있는지를 연구해본다면, 향후 예비교사를 대상으로 한 통계교육의 새로운 개선점을 도출할 수 있을 것으로 기대된다.

둘째, 예비교사들이 학교수학에서 표본 개념을 지도할 때 표본대표성뿐만 아니라 표집변이성에도 초점을 두어야 함을 인식해야 한다. 호주, 뉴질랜드, 영국, 미국 등 여러 나라에서는 교육과정에서 표집변이성을 표본대표성과 함께 지도해야 함을 강조하고 있다. 그러나 우리나라 교육과정에서는 고등학교에서 처음으로 모집단과 표본에 대해 간략히 다룬 후 바로 통계적 추정을 학습하고 있다(구나영 외, 2015, p.526). 통계교육을 통해 표집이 이루어질 때마다 표본평균의 값이 달라지면 모평균과 관련하여 각각 특수한 분포 양상을 지닌다는 것을 학생들이 인식하는 기회를 가져야 하는데, 우리나라의 교육과정은 분명한 한계를 지니고 있기 때문에 통계 교수학습의 실

제에서 교사의 역할이 매우 중요하다.

셋째, 예비교사들이 통계적 소양의 교육적 의미를 깨닫고 표본 개념 또한 통계적 소양에 비추어 이해할 수 있어야 하며, 이를 위해서는 예비교사교육 체계의 개선이 시급하다. 전공수학과 일반교육학, 그리고 제한적인 실습 기회로 이루어진 예비교사교육 체계는 통계적 소양과 관련된 적절한 내용 지식과 교수학적 지식을 발달시키는 데 한계가 있다(Ridgway et al., 2011, p.317). 더욱이, 우리나라에서는 예비교사를 대상으로 하는 통계 과목이 1~2개 수준에 불과하며(이강섭, 2003), 그나마도 확률과 묶여 확률론적 성격을 띠거나 아예 수리통계학이 중심이 되는 경우가 대부분이다. 통계적 소양 교육의 중요성이 통계교육 연구에서 아무리 강조되고 교사들이 통계적 소양의 중요성을 인식하더라도, 실제로 통계적 소양을 강조하는 교육을 경험해보지 않은 교사들이 수업에서 이를 실천하는 것은 거의 불가능에 가깝다는 점을 고려한다면, 예비교사교육 체계의 개선을 통해 통계적 소양 교육의 의미와 전략에 대해 다각도로 이해하고 판단할 수 있는 기회를 제공해야 한다(이경화, 2015, p.40).

본 연구의 참여자는 편의표집에 의해 추출되었기 때문에 본 연구 결과를 일반화하는 데에는 한계가 있다. 또한 연구 참여자들의 수가 다소 적고 모두 동일한 학교에서 선별되었다는 점을 고려한다면, 본 연구 결과가 우리나라의 모든 예비교사 대상 통계교육의 현실을 대변할 수는 없다. 그러나 학교 통계교육과 예비교사교육의 개선 방향을, 기존 연구들에 비해 다소 구체화시키고자 통계적 소양의 틀을 활용하였다는 점에서 본 연구는 이후 이루어질 후속 연구의 예비 성격을 지니고 있다. 향후, 본 연구를 토대로 학교수학에서 통계적 소양을 강조하기 위한 예비교사의 전문성 신장 전략과 체계를 개발함으로써, 통계교육의 최전선에 있는 수학교사들에게 실천적인 시사점을 제공할 수 있기를 희망한다.

참 고 문 헌

- 강현영 (2012). 통계적 소양의 교육적 의미 고찰, 한국수학사학회지, 25(4), 121-137.
- Kang, H.Y. (2012). Study of the educational meaning of Statistical Literacy. *The Korean Journal for History of Mathematics*, 23(4), 121-137.
- 강현영, 신보미, 고은성, 이동환, 심송용, 김정자, 구나영, 정인수, 최경식, 홍지혜 (2014). 통계교육 활성화를 위한 수학 교육과정 개선 방안 연구. 한국과학창의재단. 연구보고 2014A039.
- Kang, H.Y., Shin, B., Ko, E.S., Lee, D.H., Sim, S., Kim, J.J., Ku, N.Y., Jeong, I.S., Choi, K.S., & Hong, J.H. (2014). *Research on improvement of mathematics curriculum for activation of statistics education*. Research Report. Korea Foundation for the Advancement of Society and Creativity.
- 고은성 (2012). 통계적 변이성 사고 요소 간의 관계 연구. 학교수학, 14(4), 495-516.
- Ko, E.S. (2012). The Relationships among Components of Thinking related to Statistical Variability. *School Mathematics*, 14(4), 495-516.
- 고은성 (2015). 현 초중고 통계교육 연구 진행 현황 및 과제. 통계청(편). 미래사회, 이제 통계적 소양이다 (pp.83-100). 대전: 통계청.
- Ko, E.S. (2015). The status quo and the issues of statistics education research in school mathematics. Statistics Korea (Eds.). *Statistics literacy for the future society*. Daejeon: Statistics Korea.
- 고은성, 이경화 (2011). 예비교사들의 통계적 표집에 대한 이해. 수학교육학연구, 21(1), 17-32.
- Ko, E.S. & Lee, K.H. (2011). Pre-service Teachers' Understanding of Statistical Sampling. *Journal of Educational Research in Mathematics*, 21(1), 17-32.
- 교육부 (2015). 수학과 교육과정. 교육부 고시 제2015-74호 [별책 8]
- Ministry of Education. (2015). *Mathematics Curriculum*. Bulletin of MOE No. 2015-74 [Seperate Volume #8]
- 구나영, 탁병주, 강현영, 이경화 (2015). 표본 지도에 대한 고찰: 국외 교육과정 분석을 중심으로. 학교수학, 17(3), 515-530.
- Ku, N.Y., Tak, B., Kang, H.Y., & Lee, K.H. (2015). A Study on the Teaching Sample - An Analysis of Foreign Curriculum. *School Mathematics*, 17(3), 515-530.
- 이강섭 (2003). 교원양성기관의 통계분야 교육과정에 대한 기본 자료. 수학교육논문집, 15, 59-64.
- Lee, K.S. (2003). Basic data on the curriculum of the statistics for teacher education. *Communications of Mathematical Education*, 15, 59-64.
- 이경화 (2015). 우리나라 초중고 통계교육의 실제와 방

- 향. 통계청(편.). 미래사회, 이제 통계적 소양이다 (pp.27-43). 대전: 통계청.
- Lee, K.H. (2015). The location and direction of statistics education research in Korea. *Statistics Korea* (Eds.). *Statistics literacy for the future society*. Daejeon: Statistics Korea.
- 이영하 (2014). 인문학으로 풀어 쓴 통계교육 원론. 서울: 이화여자대학교출판부.
- Lee Y. (2014). *Principles of statistics education in terms of humanities*. Seoul: Ewha Woman's University Press.
- 이영하, 신수영 (2011). 초·중·고등학교 확률과 통계 단원에 나타난 표본개념에 대한 분석. 수학교육학연구, 21(4), 327-344.
- Lee, Y. & Shin, S. (2011). Features of sample concepts in the probability and statistics chapters of Korean mathematics textbooks of grades 1-12. *Journal of Educational Research in Mathematics*, 21(4), 327-344.
- 이영하, 심효정 (2003). 확률·통계 연구에 대한 수학교육학적 고찰: <수학교육>에 게재된 논문을 중심으로. 수학교육, 42(2), 203-218.
- Lee, Y. & Sim, H.J. (2003). A Trend analysis on the educational research of the probability and statistics: Focused on papers published in <The Mathematical Education>, the journal of Korea Society of Mathematical Education. *The Mathematical Education*, 42(2), 203-218.
- 이은희, 김임경 (2015). 국내외 통계교육 연구동향 비교 분석. 수학교육, 54(3), 241-259.
- Lee, E.H. & Kim, W.K. (2015). A comparative analysis on research trends of statistics education between Korea and overseas. *The Mathematical Education*, 54(3), 241-259.
- 정한영 (1995). 통계학사 개론. 춘천: 한림대학교출판부.
- Chung, H.Y. (1995). *Introduction to Statistics History*. Chuncheon: Hallim University Press.
- 지은정 (2011). 자료의 분포와 표본추출 상황에서 변이성 개념 지도. 박사학위논문, 한국교원대학교.
- Ji, E.J. (2011). *Instruction on the concept of variability in the data distributions and sampling situations*. Doctoral Dissertation, Korea University of Education.
- 탁병주, 구나영, 강현영, 이경화 (2014). 표본 개념에 대한 고찰: 역사적 분석을 중심으로. 학교수학, 16(4), 727-743.
- Tak, B., Ku, N.Y., Kang, H.Y., & Lee, K.H. (2014). A Study on the Concept of Sample by a Historical Analysis. *School Mathematics*, 16(4), 727-743.
- Bar-Hillel, M. (1979). The role of sample size in sample evaluation. *Organizational Behavior and Human Performance*, 24(2), 245-257.
- Ben-Zvi, D. & Garfield, J. (2010). 통계적 소양, 추론, 사고: 목표, 정의, 난제. In D. Ben-Zvi & J. Garfield (Eds.). 통계적 사고의 의미와 교육 (이경화 외 9인 역), (pp.3-17) 서울: 경문사. (원저 2004년 출판)
- Ben-Zvi, D., Bakker, A., & Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, 88(3), 291-303.
- Coffey, D. (2013). *Pearson Mathematics 8: Student Book*. Frenchs Forest: Pearson Education Australia.
- delMas, R.C. (2002). Statistical literacy, reasoning, and learning: A commentary. *Journal of Statistics Education*, 10(3).
- Franklin, C.G., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education report*. Alexandria: American Statistical Association
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Garfield, J., delMas, R.C., & Chance, B. (2003). The Web-based ARTIST: Assessment resource tools for improving statistical thinking. In *annual meeting of the American Educational Research Association*, Chicago.
- Garfield, J., Le, L., Zieffler, A., & Ben-Zvi, D. (2015). Developing students' reasoning about samples and sampling variability as a path to expert statistical thinking. *Educational Studies in Mathematics*, 88(3), 327-342.
- Groth, R.E. (2007). Toward a conceptualization of statistical knowledge for teaching. *Journal for research in Mathematics Education*, 38(5), 427-437.
- Groth, R.E., & Bergner, J.A. (2005). Pre-service elementary school teachers' metaphors for the

- concept of statistical sample. *Statistics education research Journal*, 4(2), 27-42.
- Hill, H.C., Blunk, M.L., Charalambous, C.Y., Lewis, J.M., Phelps, G.C., Sleep, L., & Ball, D.L. (2008). Mathematical knowledge for teaching and the mathematical quality of instruction: An exploratory study. *Cognition and Instruction*, 26(4), 430-511.
- Kahneman, D. & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430-454.
- Lipson, K. (2003). The role of the sampling distribution in understanding statistical inference. *Mathematics Education Research Journal*, 15(3), 270-287.
- Moore, D.S. (1990). Uncertainty. In L.S. Steen (Ed.). *On the Shoulders of Giants: New Approaches to Numeracy* (pp.95-137). Washington, DC: National Academy Press.
- Moore, D.S., & Notz, W.I. (2009). 논쟁거리로 배우는 통계학 (심규박, 조태영, 신미경 역), 서울: 홍릉과학출판사. (원저 2005년 출판)
- Noll, J.A. (2007). *Graduate teaching assistants' statistical knowledge for teaching*. Ph.D. dissertation. Portland State University.
- Noll, J.A. (2011). Graduate teaching assistants' statistical content knowledge of sampling. *Statistics Education Research Journal*, 10(2), 48-74.
- Pfannkuch, M. (2008). Building sampling concepts for statistical inference: A case study. *ICME-11 Proceedings*, Monterrey, Mexico, July 2008.
- Pfannkuch, M., Arnold, P., & Wild, C.J. (2015). What I see is not quite the way it really is: students' emergent reasoning about sampling variability. *Educational Studies in Mathematics*, 88(3), 343-360.
- Reading, C., & Shaughnessy, J.M. (2010). 변이에 대한 추론. In D. Ben-Zvi & J. Garfield (Eds.). 통계적 사고의 의미와 교육 (이경화 외 9인 역), (pp.239-270) 서울: 경문사. (원저 2004년 출판)
- Ridgway, J., Nicholson, J., & McCusker, S. (2011). Developing statistical literacy in students and teachers. In C., Batanero, G. Burrill, & C. Reading (Eds.). *Teaching Statistics in School Mathematics: Challenges for Teaching and Teacher Education*. New York: Springer.
- Rubin, A., Bruce, B., & Tenney, Y. (1990). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones(Eds.). *Proceedings of the third international conference on teaching statistics*. (Vol. 1, pp.314-319). Voorburg: International Statistical Institute.
- Rumsey, D.J. (2002). Statistical literacy as a goal for introductory statistics courses. *Journal of Statistics Education*, 10(3), 6-13.
- Saldanha, L. & Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Shaughnessy, J.M. (2007). Research on statistics learning. *Second handbook of research on mathematics teaching and learning*, 957-1009.
- Sullivan, P., Clarke, D., & Clarke, B. (2013). *Teaching with tasks for effective mathematics learning*. New York: Springer.
- Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The Contexts of mathematics tasks and the context of the classroom: Are we including all students?. *Mathematics Education Research Journal*, 15(2), 107-121.
- Torok, R., & Watson, J. (2000). Development of the concept of statistical variation: An exploratory study. *Mathematics Education Research Journal*, 12(2), 147-169.
- Wallman, K.K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association*, 88, 1-8.
- Watson, J.M. (1997). Assessing Statistical Thinking Using the Media. In I. Gal & J. B. Garfield (Eds.). *The Assessment Challenge in Statistics Education* (pp.107-121). Amsterdam: IOS Press.

- Watson, J.M. (2000). Preservice mathematics teachers' understanding of sampling: Intuition or mathematics. *Mathematics Teacher Education and Development*, 2(1), 121-135.
- Watson, J.M. (2010). 표본에 대한 추론. In D. Ben-Zvi & J. Garfield (Eds.). *통계적 사고의 의미와 교육* (이경화 외 역 pp.329-351) 서울: 경문사. (원저 2004년 출판)
- Watson, J.M. (2013). *학교에서 어떤 통계를 배워야 하지? 통계적 소양의 성장과 목표* (박영희 역). 서울: 경문사. (원저 2006년 출판)
- Watson, J.M., Kelly, B.A., Callingham, R.A., & Shaughnessy, J.M. (2003). The measurement of school students' understanding of statistical variation. *International Journal of Mathematical Education in Science and Technology*, 34(1), 1-29.
- Watson, J.M. & Moritz, J.B. (2000a). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31, 44-70.
- Watson, J.M. & Moritz, J.B. (2000b). Development of understanding of sampling for statistical literacy. *Journal of Mathematical Behavior*, 19, 109-136.
- Wild, C.J. (2006). The concept of distribution. *Statistics Education Research Journal*, 5(2), 10-26.

Preservice Secondary Mathematics Teachers' Statistical Literacy in Understanding of Sample

Tak, Byungjoo

Graduate School of Seoul National University
E-mail : tbj87@snu.ac.kr

Ku, Na-Young

Anyang High School
E-mail : noz39@snu.ac.kr

Kang, Hyun-Young

Department of Mathematics Education, Mokwon University
E-mail : hykang@mokwon.ac.kr

Lee, Kyeong-Hwa[†]

Department of Mathematics Education, Seoul National University
E-mail : khmath@snu.ac.kr

Taking samples of data and using samples to make inferences about unknown populations are at the core of statistical investigations. So, an understanding of the nature of sample as statistical thinking is involved in the area of statistical literacy, since the process of a statistical investigation can turn out to be totally useless if we don't appreciate the part sampling plays. However, the conception of sampling is a scheme of interrelated ideas entailing many statistical notions such as repeatability, representativeness, randomness, variability, and distribution. This complexity makes many people, teachers as well as students, reason about statistical inference relying on their incorrect intuitions without understanding sample comprehensively.

Some research investigated how the concept of a sample is understood by not only students but also teachers or preservice teachers, but we want to identify preservice secondary mathematics teachers' understanding of sample as the statistical literacy by a qualitative analysis. We designed four items which asked preservice teachers to write their understanding for sampling tasks including representativeness and variability. Then, we categorized the similar responses and compared these categories with Watson's statistical literacy hierarchy.

As a result, many preservice teachers turned out to be lie in the low level of statistical literacy as they ignore contexts and critical thinking, especially about sampling variability rather than sample representativeness. Moreover, the experience of taking statistics courses in university did not seem to make a contribution to development of their statistical literacy. These findings should be considered when design preservice teacher education program to promote statistics education.

* ZDM Classification : B5

* 2000 Mathematics Subject Classification : 97C70

* Key words : statistical literacy, sample representativeness,

sampling variability, statistics education

[†] Corresponding author

[부록] 표본 개념에 대한 예비교사의 이해 조사를 위한 질문지

1. 아래의 과제는 외국의 수학 교과서에 등장하는 문제입니다. 이 과제로 학생들에게 무엇을 가르칠 수 있을까요? 최대한 자세히 서술해주세요.

◆ 당신이 좋아하는 음악은?

철수와 영희는 학교 점심시간에 식당에서 음악을 틀어 주어야 한다고 생각합니다. 교장 선생님도 이들의 의견에 찬성하여 학생들이 어떤 음악을 좋아하는지 알아보라고 하셨습니다. 둘은 열 반 학생들과 복도에서 만난 학생들을 대상으로 설문 조사를 하였습니다. 또, 철수는 음악반, 영희는 컴퓨터반에서도 조사를 하였습니다. 그리고 그 결과를 교장 선생님께 보고했습니다. 교장 선생님은 철수와 영희의 조사 결과를 보고 식당에서 어떤 음악을 틀어 주어야 할지 종류를 결정하지 못하였습니다. 왜 그랬을까요? 설명해 보세요.

2. 아래의 과제는 외국의 수학 교과서에 등장하는 문제입니다. 이 과제로 학생들에게 무엇을 가르칠 수 있을까요? 최대한 자세히 서술해주세요.

◆ 아래 제시된 표는 주사위를 60번 던지는 시행을 다섯 차례 반복한 결과를 기록한 것이다.

결과	1세트	2세트	3세트	4세트	5세트
1	10	10	11	11	7
2	14	7	8	11	9
3	13	12	7	10	11
4	6	11	17	9	11
5	12	10	7	11	9
6	5	10	10	8	13

(a) 각 세트 결과를 이용해 주사위가 공정한지 설명하여라.

(b) (a)에서 각 세트에 따라 내린 결론을 이용하여 주사위의 공정성에 관한 전반적인 결론을 내려라.

(c) 결과들을 모두 통합하고 이를 바탕으로 주사위의 공정성을 판단하고 그 이유를 설명하여라.

3. 아래의 사례에서 밑줄 친 철수의 주장에는 어떤 문제가 있으며, 철수가 표본 개념에 대해 가지고 있는 오개념은 무엇일까요? 최대한 자세히 서술해주세요.

주머니에 색깔에 따라 크기가 다른 흰 구슬과 검정 구슬이 모두 100개가 들어 있다(단, 흰 구슬과 검정 구슬의 개수는 알지 못한다). 철수는 주머니에 흰 구슬과 검정 구슬이 몇 개씩 들어있는지 조사하기 위해 20개의 구슬을 손으로 직접 꺼내기 시작하였다. 그리고 그 구슬이 표본에 해당하느냐는 질문에 다음과 같이 답하였다.

“네. 교과서에는 모집단에서 뽑은 자료의 일부를 표본이라고 한다는데, 그렇다면 모집단의 부분집합이지만 하면 다 표본이 될 수 있는 거잖아요.”

4. 아래의 사례에서 밑줄 친 영희의 주장에는 어떤 문제가 있으며, 영희가 표본 개념에 대해 가지고 있는 오개념은 무엇일까요? 최대한 자세히 서술해주세요.

대통령 선거에서 인구가 많은 대도시 A와 인구가 적은 소도시 B에서 각각 1000명, 100명을 대상으로 출구조사를 실시하였다. 어느 도시의 출구조사가 잘못된 예측을 내놓았을 가능성이 높을까? 이 문제를 접한 영희는 다음과 같이 생각했다고 한다.

“당연히 인구가 많을수록 예측하기 어렵지 않겠어요?”