

Exploratory Study on the Discriminatory Curriculum for School Physical Education using DIF Analysis of Item Response Theory

Tae-Koo Lee¹, Ji-Young Sohn² & Heewon Yang^{1*}

¹Yonsei University & ²Sujoo Middle School

The purpose of this study was to analyze and confirm whether the items used in final paper and pencil test was determined to DIF when school sports clubs in each school operated by discriminatory curriculum in accordance with gender. Participants were 8th middle school students(male=135, female=141). They joined in school sports club every week from freshman to sophomore 1st semester. At that time, boys of them participated in soccer and basketball, and girls played dodge ball. They studied soccer unit at sophomore 1st semester, and had a final examination consisting of 5 soccer items. Using the data, differentially functioning item by the population difference between male and female were analysed quantitatively and qualitatively. The results showed that Mantel-Haenszel method(using classical test theory), comparison of item characteristic curve and likelihood ratio test(using IRT) determined item number 4 and 5 to differentially functioning item. Finally, item number 4 were identified differentially functioning item in favor of male students in intensive qualitative analyses. That item have low content validity and application-level of cognitive behavior classification. The result provides that application-level item can be functioning differentially to female students with little sports experience than male students in paper and pencil test of PE.

Key Words: Item Response Theory, DIF, Curriculum for PE, School Sports Club 

서 론

학교 체육 분야에서 교육 평등에 관한 논쟁은 중요한 정책 이슈이다. 미국에서는 1972년 Title IX을 통해 체육시간 남녀혼성수업을 전면적으로 실시하였고, 여학생들의 체육활동 시간 및 물적·인적 재원의 확보가 이루어지면서 체육교육과정 계획과 실천과정에 있어 성별에 따른 불평등적 관행들이 사라지기 시작하였다(Joo, 2015). 그리고 이때부터 미국사회에서 여학생들의 체육활동 참가비율은 급격히 증가하기 시작하였다(Lee, 2014).

학교 체육에서의 남녀 불평등은 의도적이지 않더라도 남녀 학생에게 동등하게 주어져야 할 교육 경험이 일부 학생들에게만 주어지는 것을 말한다. 이는 교육과정 실천 측면에서 교육적 지원인 교수학습과 언어 및 기타 상호작용이 일부 학생들에게만 지원되는 것을 의미한다(Francis, 2000; Metzler, 2005).

국내에서도 1990년대 이후로 학급이 단성학급에서 혼성학급으로 변화하면서 체육시간의 성차별 연구가 본격적으로 시작되었다. 체육시간관련 성차별 선행연구들은 성차별적인 고정관념과 그에 따른 성역할을 탐색(Cho & Lim, 2012; Kim & You, 2004)하거나 체육시간에 나타는 성차별적인 수업방식 및 양상(Jeong et al., 2008; Kim & Cho 2006; Yeo & Nam, 2009)에 초점을 두고 연구되었다. 그리고 보다 최근에

논문 투고일 : 2015. 12. 24.

논문 수정일 : 2016. 03. 01.

게재 확정일 : 2016. 03. 15.

* 저자 연락처 : 양희원(yhw1627@naver.com)

체육교과 인지적 평가인 지필평가와 영상평가에서 문항의 성차별적인 기능을 탐색하고 그 원인을 규명하는 연구들(Lee et al., 2010; Lee & Yang, 2016)이 이루어졌는데, 문항의 통계적 데이터를 통해 성차별적 문제를 다루는 연구가 차별기능문항(differential item function, DIF) 연구이다.

차별기능문항이란 같은 능력수준을 가진 피검자들이 그들이 속한 집단의 특성 때문에 문항의 답을 맞힐 확률이 다른 문항이다(Roussos & Stout, 1996). 즉, 차별기능문항은 집단에 따라 문항의 기능이 다른 편파성 문항을 말하는데, 결과적으로 검사 문항들에서 차별기능문항의 발견은 검사의 공정성을 훼손하는 중대한 문제이다(Kim, 2001).

차별기능을 판별하는 방법은 고전검사이론(Classical Test Theory, TCC)을 적용하는 방법과 문항반응이론(Item Response Theory, IRT)을 적용하는 방법이 있다. 그러나 고전검사이론에 비해 문항반응이론을 사용하면 문항모수 불변성과 능력모수 불변성으로 인하여 차별기능문항을 더 정확하게 판별할 수 있다. 고전검사이론은 피험자 집단 특성에 따라 문항 특성이 다르게 분석되지만, 문항반응이론에 의한 문항특성 추정과 피험자 집단의 특성에 따라 영향을 받지 않는다(Hambleton et al., 1991). 또한 고전검사이론에 의하면 개인의 고유한 능력 수준이 있음에도 불구하고 검사도구의 난이도에 따라 피험자 능력이 과소 혹은 과대 추정되지만, 문항반응이론은 검사도구의 특성이 피험자 능력추정에 영향을 주지 않는다(Seong, 2009). 고전검사이론의 이러한 단점으로 인해 차별기능문항을 판별할 때 실제 피검자의 능력 차이를 검사에 의한 검사점수의 차이와 혼동할 수 있다. 그러나 문항반응이론에 의하여 판별되는 차별기능문항은 피험자 집단 특성에 따라서 문항이 다르게 작용하고 있음을 명확히 보여줄 수 있다. 따라서 차별기능문항을 추출하는 다수의 선행연구들(Lee et al., 2007; Lee et al., 2010; Lee & Yang, 2016; No et al., 2010; Sohn, 2012; Seong, 1994)은 문항반응이론을 활용하고 있다.

차별기능문항을 판별하는 연구들은 최근에 차별기능문항들로 추출된 문항들(differential bundle functioning, DBF)의 공통점을 탐색하는 연구로 발전하고 있는데,

이는 한 검사에서 차별기능문항군을 추출하여 공통된 특성을 밝힘으로서, 차별기능문항 발생의 원인까지 탐색하려는 시도이다. 차별기능문항군을 최종적으로 판단하기 위해 사용하는 방법은 확인적 방법과 탐색적 방법이 있다(Douglas et al., 1996). 확인적 방법은 내용 전문가들이 질적 분석을 실시하여 최종적으로 차별기능문항을 판단하는 것으로, DIF를 판단하는 것과 유사하며, 국제적인 평가기관인 ETS에서 시행하고 있는 방법이다. 탐색적 방법은 통계적 방법을 통해 검사에서 일정하게 차별적으로 기능하는 문항군(differential bundle functioning, DBF)을 구성하는 방법이다. 확인적 방법(Lee et al., 2010; Lee & Yang, 2016)과 탐색적 방법(Lee et al., 2014)을 통해 차별기능문항을 최종 추출하고 그 원인을 밝힌 연구들은 추후 교수학습이나 검사문항 제작에 유용한 정보를 제공함을 보여 주었다.

학교스포츠클럽의 성차별적 운영

학교스포츠클럽 활동은 2007년에 도입·추진되어 현재까지 그 교육적 타당성을 지지받으면서 이어져오고 있는 학교체육정책이다. 그 도입 목적은 학교폭력예방과 인성교육 강화였으며, 도입 후 시간이 지난 지금 그 효과가 다양한 연구(Ahn & Eom, 2014; Cha & Kim, 2014; Hwang & Kim, 2012; Lee & Won, 2013; Youn, 2013)를 통해 교육적 타당성이 지지받고 있다. 학교스포츠클럽 활동은 2012년에 큰 변화가 있었다. 기존에 교육과정 외 시간인, 아침시간, 점심시간, 방과 후 시간 또는 주말을 이용하던 학교스포츠클럽 활동이 교육과정 내의 정규 시간으로 편성된 것이다. 2012년 7월에 초·중등 교육과정(교육과학기술부 고시 제2012-14호)을 개정·고시하면서 2012년 2학기부터는 전국의 중학교에서 창의적 재량활동의 하나로 정규의 수업시간 안에서 학교스포츠클럽 활동이 중학교 전 학년을 대상으로 실시되어 지금에 이른다. 결론적으로, 학교스포츠클럽 활동이 단위학교의 공식적인 체육교육과정으로 운영되기 시작한 것이다.

현재 전국의 중학교에서는 전교생을 대상으로 학교스포츠클럽 활동을 효과적으로 운영하기 위해 다양한 운영

유형들이 적용되고 있다. 매주 1시간씩 운영되거나 2주에 한 번씩 블록타임제로 운영되기도 한다. 그러나 운영상의 문제도 드러나고 있는데, 가장 크게 지적되는 문제는 학생들의 희망을 반영하지 못한 종목의 선택과 운영 그리고 지도교사의 전문성 부족이다. 학생들의 희망종목을 반영하지 못하는 것은 단위학교가 처한 시설 및 인적 자원의 제한으로 학교스포츠클럽 운영 종목을 결정할 때, 학생들의 희망 보다는 단위학교에서 소수의 체육교사들이 운영 가능하도록 적은 수의 종목만을 정해서 운영하는 것이다(Kim et al., 2013). 예를 들어, 축구, 농구, 피구 등 적은 수의 종목만을 운영하는 것이다. 지도교사의 전문성 부족은 소수의 체육교사들 만으로는 운영이 불가능하여 외부 강사를 채용하거나 전문성이 부족한 일반교과 지도교사가 운영을 하면서 학생들이 방치되는 것이다(Seo et al., 2014). 예를 들어, 남학생은 외부강사나 체육교사가 지도하고, 여학생은 일반교과 선생님들이 지도하면서 여학생들은 방치되는 경우이다. 이렇게 선행연구들(Kim et al., 2013; Seo et al., 2014)에서 지적한 학교스포츠클럽 운영 과정의 문제점들을 성차별적 교육과정 측면에서 종합하면, 소수의 운동 종목을 정해서 운영할 때 남학생들과 여학생들의 운영종목을 분리해서 운영한다는 것이다. 예를 들어, 남학생들은 축구, 여학생들은 피구로 운영하는 것이다. 본 연구대상인 된 행복중학교(가명)는 최근 2년 동안 남학생들은 축구를 비롯한 구기 종목을, 여학생들은 피구 종목만을 학교스포츠클럽 종목으로 운영하였으며, 여학생들은 일반교과 교사가 지도하여 교수학습의 전문성이 부족했다.

차별기능향의 발생 원인을 밝힌 선행연구들(Lee et al., 2010; Lee & Yang, 2016; No et al., 2010)은 학교 내 교육과정의 성차별적인 운영과 성차별적인 사회·문화적 기능에 원인이 있음을 보고하고 있다. 그렇다면 역으로 학교 체육교육과정 즉, 학교스포츠클럽 활동이 성차별적으로 운영되고 있을 때, 체육 지필검사에서 출제된 문항들이 차별기능향으로 추출된 가능성이 높다고 추론할 수 있고, 이는 평가의 공정성이 훼손되는 것이다. 그러나 선행연구(Kim et al., 2013; Seo et al., 2014)에서 지적하고 있듯이, 전국의 많은 중학교들에서 학교스포츠클럽 운영 종목을 남녀 구분하여 운영하는 사례들이 발생하고 있지만, 이로 인해 체육

교과 지필시험에서 차별기능향의 발생으로 평가의 공정성이 훼손되고 있는지를 확인하는 연구는 실천되지 않고 있는 실정이다. 만약 차별기능향이 발생한다면 학교스포츠클럽 운영 방법의 변화가 불가피하다.

따라서 이 연구의 목적은 교육과정 내 스포츠클럽 활동이 성별에 따라 차별적인 교육과정으로 운영되었을 때, 지필검사에 출제된 문항들이 차별기능향으로 추출되는지 분석·확인하는 것이다. 이렇게 지필문항에서 차별기능향을 확인하는 것은 성차별적인 교육과정에 대한 반성과 개선은 물론, 궁극적으로 여학생체육활성화에 기여할 수 있다(Lee & Yang, 2016). 이를 위한 구체적인 연구문제는 다음과 같다.

첫째, 문항들의 일반적인 문항 양호도는 어떠한가?

둘째, 고전검사이론의 MH 방법과 문항반응이론의 문항특성곡선 분석 및 우도비 검정 방법을 적용·분석하여 판별한 차별기능향은 무엇인가?

셋째, 질적 분석을 통해 최종적으로 차별기능향으로 추출된 문항은 무엇인가?

연구방법

연구 참여자

이 연구의 연구대상은 경기도 행복중학교(가명) 2학년 276명(남:135, 여:141) 학생이었다. 본 연구의 목적은 만약 학교스포츠클럽 활동이 성차별적으로 이루어지고 있으며, 그때 체육수업에서 학교스포츠클럽 활동의 종목을 선택하여 지도하였다면, 지필시험에서 차별기능향이 나타날지 확인하는 것이다. 연구자 본인이 근무하는 행복중학교는 이에 부합하는 학교체육교육과정을 운영하였다.

연구 대상은 지난 18개월 동안 학교스포츠클럽 활동에서 남학생은 축구를 비롯한 구기 활동을, 여학생은 피구를 운영 종목으로 경험하였다. 남학생들 축구지도와 관련하여 1학년 때는 축구전문 외부강사가, 2학년 때는 체육교사가 지도하였다. 외부강사와 체육교사는 축구의 기술부터 경기기술을 지도하고 학생들이 축구 경기도 경

협하도록 하였다. 여학생들의 피구지도는 교내 타 교과 교사들이 지도하였다. 2학년 1학기에 진행된 체육수업에서 축구단원 수업은 총 15차시로 진행되었는데, 교수학습은 수행평가인 드리블(시간)과 슈팅(성공개수) 위주로 이루어지고 수행평가가 양적 평가로 이루어 졌다.

연구대상 학생들은 2학년 1학기 기말고사에서 20문항으로 구성된 체육 기말고사에 응답하였다. 기말고사는 육상, 축구, 플라잉 디스크를 범위로 출제되었는데, 축구(풋살)는 다섯 문항 이었다 [축구관련 4문항, 풋살관련 1문항, 이후에는 모두 축구관련 문항으로 표기]. 따라서 본 연구에서는 축구관련 다섯 문항을 하나의 검사도구로 여기고 연구 대상 학생들이 축구관련 문항에 응답한 반응에 대하여 분석하였다. 성별 검사총점의 분표와 점수별 빈도는 <Table 1>, <Table 2>와 같다.

전체학생의 평균점수는 3.3점으로 남학생들과 여학생들의 평균점수는 같았다. 총점의 비율에서 남학생들은 4점을 획득한 학생들(14.9%)이 가장 높은 비율이었으나, 여학생들은 5점 만점을 획득한 학생들의 비율(13.0%)이 가장 높았으며, 5점 만점을 받은 학생들은 남학생(28명)보다 여학생(36명)이 많았다.

차별기능문항의 질적 판단을 위해서는 좋은체육수업 나눔연구회 소속 연구위원 18명의 교사가 참여하였는데, 이 교사들은 교직경력 10년 이상의 교사들로 체육시험 문항을 지속적으로 출제하고 검토한 경험을 가지고 있다.

Table 1. Score Distribution; Mean and SD

	Sample Size	Mean	SD
Total	276	3.3	1.354
Male	135	3.3	1.379
Female	141	3.3	1.334

Table 2. Sample Sizes by Total Score

Score	Sample Size			Percent(%)		
	Total	Male	Female	Total	Male	Female
0	6	5	1	2.2	1.8	0.4
1	22	11	11	8.0	4.0	4.0
2	57	23	34	20.7	8.3	12.3
3	53	27	26	19.2	9.8	9.4
4	74	41	33	26.8	14.9	12.0
5	64	28	36	23.2	10.1	13.0

분석 방법

분석은 세 단계를 거쳐 이루어졌다. 먼저 축구 관련 다섯 문항에 대한 문항 양호도를 분석하였다. 문항 양호도 분석의 항목들에 관하여 중등학교 체육교과 검사문항 양호도 연구를 한 Lee et al.(2010)와 Lee & Yang(2016)의 연구를 참고하여, 타당도, 신뢰도, 변별도, 추측도, 검사정보함수를 살펴보고, 체육교과 지필검사의 문항분석을 시도한 Kim & Lee(2004)와 Park(2011)의 연구를 참고하여 Bloom의 인지행동 유목(Anderson et. al., 2001)을 추가하여 문항 양호도를 분석하였다. 인지행동 유목의 분석은 선행연구들(Kim & Lee, 2004; Park, 2011)의 문항 질적 분석에서 주요하게 다루는 내용이다.

두 번째 분석은 차별기능문항 분석을 위해 고전검사이론의 MH 방법과 문항반응이론의 문항특성곡선의 비교, 그리고 우도비 검정 방법을 적용하였다.

세 번째 분석은 차별기능문항으로 판별된 문항에 대한 질적 분석을 통해 최종적으로 차별기능문항을 추출하는 것이다. 질적 분석은 문항의 개발자 또는 내용의 전문가 등이 토론을 통해 문항의 편파성과 그 원인을 분석하는 것이다. 질적 분석을 위해서는 전문가 집단의 전문성을 확보하는 것이 매우 중요하다. 왜냐하면 차별기능문항의 추출과정 가운데 의견이 일치하지 않는 상황(Roussos & Stout, 1996)도 발생하기 때문에 일정수준의 전문성을 가진 집단의 구성이 의견 불일치가 발생하는 상황을 최소화하는 필요조건 중에 하나이다. 따라서 연구자들은 체육 지필문항 출제에 전문성을 갖춘 체육교사들을 선별하여 질적 분석 위원회를 조직하였다. 이는 ETS에서 운영하는 방식으로 차별기능문항을 최종 추출할 때 깊이 있는 전문성에 기반하기 위함이다. 질적 분석 위원회 구성은 경기도교육청 소속 좋은체육수업나눔연구회 소속 연구위원¹⁾ 교사 18명이 질적 분석에 참

1) 좋은체육수업나눔연구회 온라인 카페에는 2016년 2월 현재 전국의 4,000명 이상의 체육교사들이 회원으로 가입되어 있으며, 이는 전국 체육교사의 1/3에 해당한다. 이중 연구위원 교사들은 교직경력 10년 이상의 교사들 중 선발된 교사들로 구성되며, 이들의 수행한 연구 결과로 인해, 본 연구회는 경기도교육청에서 심사하여 선정된 우수연구회로 2014·2015년 2년 연속 선정되었으며, 이는 체육교과 연구회로서는 유일하다. 연구위원들은 연구대회 수상, 장학자료와 체육교과서 집

여하였다. 질적 분석 위원회 교사들은 먼저 이메일을 통해 축구관련 5문항과 이원목적 분류표를 받고 문항의 내용 타당도를 채점하여 송부하였다. 그리고 오프라인 모임에서 교사들은 문항들에 대한 인지행동 유목분석과 연구자들이 통계적 기법으로 판별한 차별기능문항들에 대하여 질적 분석을 하였다. 분석에 활용된 다섯 개의 축구관련 문항은 <Fig. 1>부터 <Fig. 5>에 제시하였다.

1. Choose the correct explanation about the regulation in football game?

- ① It is possible to change the three players.
- ② The game time is 50 minutes each half.
- ③ There are a chief referee, a assistant referee and a waiting referee.
- ④ The break time of game in middle school student is 20 minutes.
- ⑤ There are 11 substitutes and main players in each

Fig. 1. Item No.2

2. Choose the fouls related with direct freekick a the whole bang in examples about football game.

<Examples>

- ㉠. 2th yellow card
- ㉡. Holding
- ㉢. Off-side
- ㉣. Keeping holding the ball of a goalkeeper over 6 seconds.
- ㉤. Handball of player without a goalkeeper

- ① ㉠, ㉡
- ② ㉠, ㉢
- ③ ㉡, ㉤
- ④ ㉢, ㉣
- ⑤ ㉣, ㉤, ㉤

Fig. 2. Item No.1

필 등에 참여하였으며, 2014년 체육교육한마당(교육과학기술부·스포츠교육학회주최)에서는 수업사례분과 운영을 맡다 현직 교사들 대표로 우수수업사례를 발표하였다.

3. Match the names inserting (㉠) and (㉡) related with the football skills.

- (㉠) is the skill to stop and receive a rolling and flying ball using body parts without hands.
- (㉡) is used in time to move quickly forward.

(㉠)	(㉡)
① Heading	Trapping
② Trapping	Throw in
③ Throw in	Trapping
④ Trapping	Instep dribble
⑤ Instep dribble	Inside dribble

Fig. 3. Item No.3

4. Choose the correct explanation about the football skill.

- ① The name to kick the ball bending toward the outside of body is in-front kick.
- ② The name to kick the ball bending toward the inside of body is out-front kick.
- ③ Instep kick is used in shout distant pass or shot.
- ④ Inside kick is accurate because the contact area of foot and ball is wide relatively.
- ⑤ Shoulder charge is cheating skill for a defender for an offenser's breaking through defence.

Fig. 4. Item No.4

5. Choose the correct explanation about the futsal game.

- ① The number of main player is three.
- ② Tackle and tussle of the player are injuncted.
- ③ Game time is 10 minutes in each half.
- ④ A opponent will receive direct freekick when the player commit 3rd foul.
- ⑤ A player committed a poul after the 6th inning will be withdrawn for 5 minutes.

Fig. 5. Item No.5

연구 결과

문항 양호도

타당도

내용타당도는 검사내용 전문가에 의해 검사가 측정하고자 하는 속성을 제대로 측정하고 있는지를 주관적으로 판단하는 것이다(Seong, 2002). 좋은체육수업나눔연구회 연구위원 18명의 체육교사들은 이원목적 분류표를 근거로 각 문항을 100점 만점에 10점 단위로 하여 내용 타당도 채점을 하였다. 조사 결과는 <Table 3>과 같다.

내용 타당도 조사결과에서 전체 평균점수는 80.4%로 비교적 양호하였다. 문항별로 보면, 직접 프리킥의 사례를 고르는 2번 문항의 내용 타당도가 가장 높았고, 킥의 종류를 비롯한 축구 기술에 대해 물어보는 4번 문항이 가장 낮은 내용 타당도를 보였다.

Table 3. Content Validity (unit: %)

Item	1	2	3	4	5	Mean
Score	77.2	86.7	82.8	76.7	78.9	80.4

Table 4. Results of Bloom's Taxonomy for Items Related Football

Item	Cognitive Process Dimension*					
	①	②	③	④	⑤	⑥
1	○					
2		○				
3	○					
4			○			
5	○					

* ①=Remember, ②=Understand, ③=Apply, ④=Analyze, ⑤=Evaluate, ⑥=Create

인지행동유목

인지행동 유목에서 제시하고 있는 기억부터 창안까지 유목들은 낮은 지적 기능으로부터 높은 지적 기능으로 순차적으로 기술되어 있는데, 인지행동 유목은 원래 검사문항 개발자들에게 필요한 검사문항 틀을 제공하기 위

해 만들어졌다(Anderson et. al., 2001). 축구관련 문항들을 대상으로 인지행동 유목 분석결과는 <Table 4>에 제시하였다. 분석결과, 1번, 3번, 5번 문항이 기억수준이었고, 2번 문항이 이해수준, 4번 문항이 적용수준이었다.

문항특성

문항의 양호도를 검증하기 위한 또 다른 방법으로 고전검사이론과 문항반응이론을 적용하여 문항분석을 실시하였으며, 성별에 따라 문항에 다르게 응답한 양상을 확인하기 위하여 남학생과 여학생 각각의 집단에 대하여 같은 방법으로 문항분석을 실시하였다. 전체 집단을 대상으로 한 문항분석 결과는 <Table 5>와 <Table 6>에 제시하였다. 고전검사이론에서 난이도는 총 피험자 중 답을 맞힌 피험자의 비율로, 변별도는 문항점수와 피험자 총점의 상관관계인 점이연상관계수로 나타낸다. 대체로 변별도는 0.20 미만인 경우 변별도가 매우 낮다고 평가하며, 그 값이 음수인 경우 검사에서 제외하여야 한다(Seong, 2002).

고전검사이론에 의하여 문항을 분석한 결과, <Table 5>에서와 같이 정답률이 가장 높은 문항은 3번이었고, 가장 낮은 문항은 5번이었다. 변별도가 가장 높은 문항은 5번이었고, 가장 낮은 문항은 3번이었다. 3번 문항의 변별도는 0.2 이하로 매우 낮게 판별되었다. 내적 일관성 신뢰도의 추정치인 Cronbach's α 계수는 0.522로 나타났다.

Table 5. Item Analysis using Classical Test Theory

Item	Percentage of correct answers	point-biserial correlation	Reliability
1	74.7	0.330	
2	59.8	0.296	
3	83.7	0.182	0.522
4	62.7	0.266	
5	52.2	0.374	

문항반응이론에 의하여 문항을 분석한 결과는 <Table 6>과 같다. 문항반응이론에 의한 문항분석에서, 난이도는 문항의 답을 맞힐 확률이 50%에 해당하는 능

력수준을 의미하며, -2에서 2 사이에 일반적으로 존재하여 그 값이 클수록 그 문항은 어렵다고 해석된다. 변별도는 문항난이도를 나타내는 피험자 능력수준보다 낮은 능력의 피험자와 높은 능력의 피험자를 변별하는 정도를 나타내며, 대체로 0부터 2의 범위에 있다. 추측도는 능력이 전혀 없는 피험자가 추측에 의해 문항의 답을 맞힐 수 있는 확률을 의미하는데, 일반적으로 '1/(답지수)'보다 낮은 확률을 나타낸다(Seong, 2002).

Table 6. Item Analysis using Item Response Theory

Item	Difficulty	Discriminat-ion	Guessing	Reliability
1	-0.536	1.024	0.196	
2	0.105	0.977	0.235	
3	-1.560	0.603	0.211	0.541
4	-0.101	0.722	0.214	
5	0.187	1.252	0.149	

Table 7. Item Analysis using Classical Test Theory by Gender Group

Item	Male		Femael	
	Percentage of correct answers	point -biserial correlation	Percentage of correct answers	point -biserial correlation
1	71.9	0.345	71.6	0.316
2	55.6	0.395	63.8	0.208
3	85.2	0.242	82.3	0.130
4	70.4	0.304	55.3	0.267
5	44.4	0.349	59.6	0.424

문항반응이론에 의하여 문항을 분석한 결과, 고전검사이론에 의한 변별도가 가장 낮은 3번 문항의 난이도가 -1.560으로 가장 낮게 나타났고, 변별도 역시 0.603으로 가장 낮았다. 추측도의 경우 2번 문항이 0.235로 가장 높고, 5번 문항이 0.149로 가장 낮았으며, 전체 평균은 0.201로 높지 않았다. 문항반응이론에 의한 신뢰도는 피험자 능력 수준에 따라 다르게 산출되는 측정오차의 평균인 주변 신뢰도(marginal reliability)를 산출하였으며(Green et, al., 1984), 그 값은 0.541로 고전검사이론에서의 Cronbach'α 계수와 유사하였다.

다음으로 성별에 따라 문항특성이 어떻게 다른지 살펴보기 위하여 남학생과 여학생 각각의 집단에 대하여 고전검사이론과 문항반응이론을 적용하여 문항분석을 실시하였다(Table 7). 분석결과, 남학생과 여학생 집단의 평균 정답률은 각각 65.5%와 66.5%로 유사하지만, 4번 문항과 5번 문항의 정답률은 15%p 이상의 차이를 보였다.

BILOG-MG를 사용하여 집단별 모수추정을 수행할 경우, 변별도와 추측도 모수를 전체 집단에서 산출되는 값으로 고정하고 난이도 모수를 추정하므로, 여기에서는 남학생과 여학생 집단의 난이도 모수만을 제시하였다(Table 8). 고전검사이론에 의한 평균 정답률과 동일하게 남학생과 여학생의 난이도 모수의 평균 역시 -0.552와 -0.532로 유사하지만, 2번, 4번, 5번 문항은 남학생과 여학생의 난이도 모수에 차이가 나타났다. 남학생은 4번 문항에서 0 이하의 난이도 모수가 추정되었고, 여학생은 1번과 5번 문항에서 0 이하의 난이도 모수가 추정되었다.

Table 8. Item Analysis using Item Response Theory by Gender Group

Item	Difficulty	
	Male	Female
1	-0.676	-0.676
2	0.077	-0.297
3	-1.925	-1.632
4	-0.666	0.136
5	0.431	-0.191
Mean	-0.552	-0.532

검사정보함수

문항정보함수(item information function)는 피험자의 능력 추정에서 표준오차가 어느 정도인지 말해준다. 표준오차가 작다는 것은 능력을 정확하게 추정하였으므로 그 문항은 높은 정보를 지니게 된다고 해석된다. 그러나 하나의 문항으로 피험자의 능력을 추정하는 경우는 거의 없으므로, 문항정보들의 총합인 검사정보함수를 사용한다. 다음 <Fig. 6.>의 검사정보함수 곡선에서 실

선은 남학생 집단을, 점선은 여학생 집단을 나타낸다. 이 연구에 쓰인 문항들로 구성된 검사의 경우 능력수준이 -1에서 1 사이일 때 검사정보가 가장 높다. 즉, 이 능력 수준 범위를 벗어나는 피험자에게는 이 검사가 적당하지 않을 수 있으며 특히 능력수준이 1 이상인 피험자에게는 매우 쉬운 문항으로 구성된 검사가 된다.

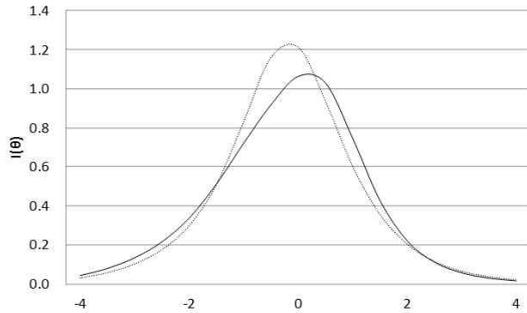


Fig. 6. Test Characteristic Curve

차별기능문항 분석

각 문항이 남학생과 여학생의 집단에 따라 차별적으로 기능하고 있는지 확인하기 위하여, 고전검사이론의 MH 방법과 문항반응이론에 의한 문항특성곡선(Item Characteristic Curve: ICC)의 확인 및 우도비 검정 방법을 활용하였다.

고전검사 활용: MH 방법

MH방법을 적용한 결과는 다음 <Table 9>와 같다. 유의수준 0.05에서 자유도가 1일 때의 χ^2_{MH} 값인 3.84보다 큰 값을 나타내는 문항은 집단 간 동질성을 확보하기 어려운 문항으로 볼 수 있다(Dorans & Holland, 1993). 이 연구에서 사용된 문항들 중 4번이 이에 해당하며, 5번 역시 통계적으로 유의미하지는 않지만 비교적 집단 간 동질성의 확보가 어려운 것으로 볼 수 있다. 4번 문항의 Δ_{MH} 값은 음수이므로 남학생 집단에 유리하고, 5번 문항의 Δ_{MH} 값은 양수이므로 여학생 집단에 유리한 것으로 판별되었다. 남학생 집단에 유리한 것으로 판별된 4번 문항의 정답률은 남학생의 경우 70.4%였고,

여학생의 경우 55.3%로 그 차이가 뚜렷하다. 체육교과에서 차별기능문항을 판별한 Lee et al.(2010)과 Lee & Yang(2016)의 연구에서 차별기능문항의 최대 χ^2_{MH} 값은 각각 4.549와 4.450이었다. 이는 χ^2_{MH} 값 10.663인 4번 문항은 선행연구에서 추출된 차별기능문항들보다 상대적으로 더 크게 남학생에게 유리한 문항이라고 할 수 있다.

Table 9. DIF Using MH Method

Item	Δ_{MH}	χ^2_{MH}	advantageous group
1	-0.334	0.023	
2	0.399	0.178	
3	-1.371	0.890	
4	-2.464	10.663	Male
5	1.461	3.417	Female

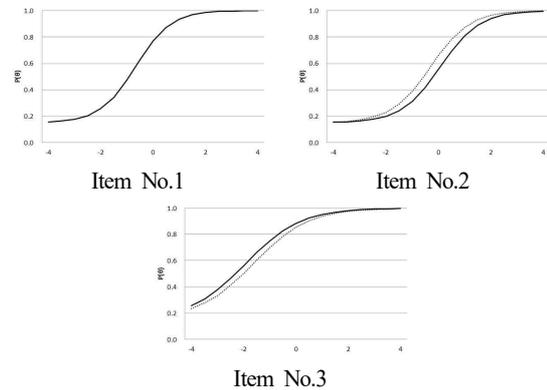


Fig. 7. ICC for Item No.1 to No.3

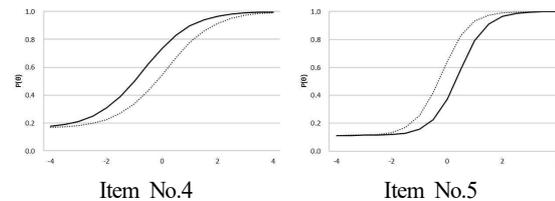


Fig. 8. ICC for Item No.4 to No.5

문항반응이론 활용: 문항특성곡선

<Fig. 7>과 <Fig. 8>은 축구관련 다섯 문항에 대한 문항특성곡선을 남녀 그룹별로 비교한 결과이다. 문항반응이론에 의한 문항특성곡선은 실선이 남학생 집단, 점

선이 여학생 집단을 나타낸다.

(Fig. 7)에서 제시하는 1번부터 3번 문항까지의 문항특성곡선과 (Fig. 8)에서 제시하는 4번과 5번 문항의 문항특성곡선은 남녀별 곡선이 불일치 정도에서 뚜렷이 구별된다. 4번과 5번 문항의 문항특성곡선은 1번부터 3번 문항까지의 문항특성곡선에 비해 두 곡선이 불일치하는 정도가 큰 것을 확인할 수 있다. 4번 문항은 같은 능력 모수일 경우 여학생 집단에 비해 남학생집단이 정답할 확률이 더 높다는 것을 보여주고 있으나, 5번 문항은 반대로 여학생 집단이 정답할 확률이 더 높다. 즉, 4번 문항은 남학생 집단에게 그리고 5번 문항은 여학생에게 더 유리한 문항이다. 4번 문항과 5번 문항의 차별기능의 정도는 4번 문항의 차별기능 정도가 더 컸다. 2번 문항과 3번 문항도 남녀별 두 곡선이 불일치하는 정도가 작지만, 어느 정도의 차별기능을 수행하고 있는 것으로 나타났다. 특히 3번 문항은 낮은 능력모수에서 차별기능을 수행하고 있다.

문항반응이론 활용: 우도비 검정

우도비 검정 방법은 문항반응이론의 문항모수의 차이를 통해 차별기능문항을 추출한다. 이 연구에서는 3모수 문항반응이론 모형을 적용하였고, 분석 결과 G^2 값이 $\alpha = 0.05$ 수준에서 자유도 1인 χ^2 분포의 임계치인 3.84를 넘는 문항은 4번과 5번 문항인 것으로 나타났다 (Table 10). 두 문항 모두 난이도 모수에서 차이가 나타났다으므로 균일적 차별기능을 수행하였다. 균일적 차별기능은 비균일적과 대비되는 표현으로 학생들의 능력모수 전 범위에서 일관되게 차별기능을 수행하는 것을 말한다. 4번 문항에서 남학생 집단과 여학생 집단의 난이도모수는 각각 -0.69와 0.31로 남학생에게 유리하였고, 5번 문항의 난이도 모수는 각각 0.53과 0.02로 여학생에게 유리하였다.

차별기능문항 질적 분석

고전검사이론의 MH 방법과 문항반응이론에 의한 문항특성곡선의 비교와 우도비 검정 방법을 통해 모두 문항 4번과 5번이 차별기능문항으로 판별되었다. 4번 문항은 남학생에게 유리하고, 5번 문항은 여학생에게 유리

Table 10. DIF Using Likelihood Ratio Test Method

Item	G^2	DIF type	advantageous group
1	1.4		
2	1.3		
3	1.6		
4	11.8	uniform	Male
5	7.9	uniform	Female

하였다. 질적 분석은 4번과 5번 문항에 대한 체육교과 전문성을 기반으로 분석을 하여 최종적으로 차별기능문항을 추출하고 그 원인을 추론하는 것이다.

질적 분석 위원회 교사들은 먼저 여학생에게 유리한 문항으로 판별된 5번 문항에 대하여 논의를 하였다. 교사들은 일반적으로 체육지필검사에서 여학생에게 유리한 문항들에 대해서는 학교 현장의 현실성을 고려하여 논의를 해야 한다고 하였다. 현장의 현실성이란? 체육지필검사의 문항들이 단순 기억수준의 유목을 가진 문항들로 출제되었을 때, 학교 현장에서는 여학생들이 남학생들보다 시험성적이 높다는 것이다. 이러한 학교 현장의 상황들은 소수의 선행연구들에서도 확인할 수 있는데 Lee & Yang(2016)과 Cha et al.(2014)의 연구에서는 체육지필검사에서 신체활동 관련 경험이 상대적으로 적은 여학생들의 점수가 남학생들보다 높았다. 이러한 결과는 여학생들이 시험을 위해 교과서 지문 전체를 더 잘 암기하기 때문에 남학생들보다 시험점수가 높은 학교 현장의 현실성이 드러난 것으로 추측해 볼 수 있다. 5번 문항과 관련하여 남학생과 여학생들 모두는 풋살 경기를 실제로 경험한 적이 없고, 시험을 위한 이론수업만 진행한 상태였다. 인지행동 유목이 기억수준인 5번 문항이 여학생에게 유리하게 차별적으로 기능하는 것은 여학생들의 학업적 특성 때문으로 판단되기에, 최종적으로 5번 문항은 차별기능문항에서 제외하였다.

4번 문항은 χ^2_{MH} 값 10.663으로 매우 현격하게 남학생에게 유리한 문항이었다. 4번 문항은 인지행동 유목이 적용수준으로 보기 네 개는 킥의 다양한 동작을 설명하고 있고, 실제로 킥을 해본 경험이 있어 동작을 이해하는 학생들은 쉽게 답을 할 수 있는 문항이었다. 질적 분석 위원회 교사들은 지필검사에서 4번 문항과 같이 특

정동작이나 경기상황관련 지식을 물어보는 문항들은 그 내용을 실제로 경험하지 못한 여학생들에게는 어려운 문제라고 지적하였다. 더욱이 해당 종목에서 동작이나 경기상황관련 지식을 종이위에 표현하여 문항을 전달하는 것은 실제 상황을 표현하는데 한계가 있기 때문에(Choi & Park, 2014), 여학생들에게는 4번 문항이 더욱 어려웠을 것으로 판단하였다. 사실 4번 문항 보기에서 제시한 킥 명칭들은 정답에 대한 힌트가 되는 면이 있음에도 불구하고 여학생들과 남학생들의 정답률과 난이도의 차이가 큰 것(Table 7과 8)은 남학생들이 킥의 대한 경험이 많기 때문으로 교사들은 판단하였다. 따라서 4번 문항은 최종적으로 차별기능문항으로 추출되었다.

4번 문항이 최종적으로 차별기능문항으로 추출된 이유에 대하여 질적 분석 위원회 교사들은 두 가지 원인을 추론하였다. 하나는 차별적인 교육과정 운영과 두 번째는 사회문화적인 배경이다. 차별적인 교육과정 운영과 관련하여 학교스포츠클럽에서 남학생들은 1학년 때 1년 동안 전문 축구강사와 같이 축구를 하였다. 그리고 2학년 1학기 때에는 일반교사가 아닌 체육교사가 전담하여 학생들을 지도하였다. 비록 2학년 1학기 체육수업에서 남학생들과 여학생들이 동일하게 축구단원을 학습하였지만, 차별적인 학교스포츠클럽 교육과정 운영으로 인해 남학생들은 자연스럽게 4번 문항의 보기에서 제시되는 인프런트 킥, 아웃프런트 킥, 인스텝 킥, 인사이드 킥, 솔더 차징 등의 단어를 듣게 되고 자신의 기능적인 축구 실력과는 별개로 축구 용어들을 더 잘 인지적으로 이해할 수 있게 되었을 것이라고 추론할 수 있다. 이렇게 차별적인 교육과정 운영이 최종적으로 4번 문항이 차별기능문항으로 추출되는 원인이 될 수 있다. 질적 분석 위원회 교사들이 두 번째 원인으로 지목한 사회문화적인 배경은 남학생들과 여학생들의 사회화과정의 차이를 말하는데, 교사들은 축구종목의 인지적 수준차이가 이미 매우 커서 연구결과는 이것을 반영한 결과일 수 있다고 지적하였다. 학교 현장에서 축구 종목은 남학생들이 매우 열정적으로 참가하는 스포츠종목이다. 2007년 학교스포츠클럽 활동이 도입되기 이전부터 축구는 학교 현장에서 남학생들이 가장 좋아하고 즐겨 참여하는 스포츠 활동이었으며, 국가대표의 축구경기나 영국이나 스페인 리그의 축구경기 중계를 남학생들은 즐겨본다. 이렇게

축구 중계를 보는 남학생들은 그렇지 않은 여학생들보다 축구경기 용어나 인지적 내용에 대해 더 많이 배우게 된다. 왜냐하면 학생들의 스포츠지식 형성은 실제로 실제 생활 속에서 시청한 중계방송을 통해 이루어 질 수 있기 때문이다(Brooker et. al., 2000). 질적 분석 위원회 교사들은 이런 사회문화적인 배경으로 인해, 4번 문항처럼 특정동작이나 경기상황과 관련된 문항들은 축구 활동 경험이 많은 남학생들에게 매우 유리하게 작용하여 성별에 따라 정답률의 차이가 발생할 수 있다고 하였다.

질적 분석 위원회 교사들은 4번 문항이 남녀 집단별로 차별적 기능을 하는 원인으로 차별적인 교육과정 운영과 사회문화적인 배경을 지적하였지만, 행복중학교처럼 학교스포츠클럽 운영을 남녀 차별적으로 운영하는 것은 남녀학생들에게 공평한 교수학습기회를 제공한다는 측면에서 매우 큰 문제라는데 인식을 같이 하였다. 이미 체육교사로서 우리 사회 안에 있는 성차별적인 사회문화적인 배경을 인식하고 있는 교사라면 학교 체육교육과정을 운영할 때 성차별적인 운영이 되지 않도록 세심한 주의가 필요하다는 것이다. 특히 4번 문항은 현격하게 남학생에게 유리한 문항특성을 보이고 있으므로, 질적 분석 위원회 교사들은 차별적인 교육과정 운영이 그 원인이 되었을 가능성이 매우 높다고 추론하였다.

논 의

이 연구의 목적은 교육과정 내 스포츠클럽활동이 성별에 따라 차별적인 교육과정으로 운영되었을 때, 지필 검사에 출제된 문항들이 차별기능문항으로 추출되는지 분석·확인하는 것이다. 행복중학교는 최근 2년 동안 학교스포츠클럽운영 종목을 성별에 따라 선택적으로 운영하였으며, 기말고사에서 남학생들에게 선택적으로 운영된 축구종목에 대하여 다섯 문항을 시험문항으로 출제하였다. 분석결과, 문항양호도 측면에서 축구관련 다섯 문항은 문항별 차이가 있었으나 전체적으로 타당도, 난이도, 변별도 및 신뢰도는 양호하였고, 인지행동 유목은 기억수준 세 문항, 이해수준과 적용수준은 각각 한 문항이었다. 검사특성곡선은 남녀별 곡선의 형태가 유사했으나 남학생들보다는 여학생들에게서 능력수준이 좀 더 낮

은 범위에서 많은 정보를 주고 있었다. 차별기능문항분석에서 고전검사이론과 문항반응이론을 적용한 분석에서 모두 4번과 5번 문항이 성별에 따른 차별기능을 하고 있었으며, 질적 분석을 통해 4번 문항이 최종적으로 남학생에게 유리한 차별기능문항으로 추출되었다. 그 원인으로서는 사회문화적인 배경과 차별적인 교육과정 운영이 지적되었으나 차별적인 교육과정 운영이 보다 유력한 것으로 추론되었다. 이러한 결과를 바탕으로 논의를 하면 다음과 같다.

첫째, 일반적인 문항 양호도는 적정수준이었으나, 문항별, 성별로는 차이를 보였다. 전체적인 내용타당도는 80.4%로 나타났으며, 4번 문항의 수치(76.7%)가 가장 낮았다. 남학생들에게서 정답률이 가장 낮은 문항은 5번 문항(44.4%)이었으나, 여학생들은 4번 문항(55.3%)이었다. 문항반응이론을 적용한 난이도 역시, 남학생들은 5번 문항(0.431)을 가장 어려워했으나, 여학생들은 4번 문항(0.136)을 어려워해 고전검사이론의 정답률과 유사한 결과가 나타났다. 문항별 인지행동 유목분석결과에 연관이해보면 기억수준의 문항인 1번과 3번은 남녀모두에서 정답률이 높았다. 그러나 5번 문항은 남학생의 정답률(44.4%)이 상대적으로 여학생(59.6%)보다 크게 낮았다. 이는 차별기능문항의 질적 분석에서 언급되었듯이 여학생들의 학업적인 특성이 반영된 것으로 파악되었다. 4번 문항은 적용 수준의 문항인데, 체육지필검사관련 선행연구들(Kim & Lee, 2004; Park, 2011)을 보면 적용수준의 문항은 현장 체육교사들이 잘 출제하지 않는 수준의 문항이다. 4번 문항은 고전검사이론과 문항반응이론을 적용한 분석 모두에서 여학생들에게는 어렵고 상대적으로 남학생들에게는 쉬운 문항으로 적용수준의 지필문항이 운동경험이 적은 여학생들에게 보다 어려운 문항이 될 수 있음을 알 수 있다.

두 검사이론을 적용한 신뢰도 분석결과는 0.522와 0.541로 유사했다. Ebel & Frisble(1991)은 검사 길이와 신뢰도 계수의 관계를 제시하면서 문항수가 증가하면서 신뢰도 계수는 선형적으로 증가한다고 하였다. 이들이 제시한 이론에서는 다섯 문항의 검사에서 기대할 수 있는 신뢰도는 .20인데, 본 축구관련 다섯 문항의 신뢰도가 0.5점대로 나타난 것은 문항길이 대비 좋은 신뢰도라고 할 수 있다.

둘째, 차별기능문항 판별과 관련하여 고전검사이론과 문항반응이론을 적용·분석한 결과 4번과 5번 문항이 차별기능문항으로 판별되었다. 4번 문항은 남학생에게 유리한 문항이었고, 5번 문항은 여학생에게 유리한 문항이었다. MH 방법에서 5번 문항은 통계적으로 유의하지 않았지만, 비교적 남녀집단 간 동질성 확보가 어려운 것으로 볼 수 있는 수치(3.417)가 도출되었다. 4번 문항은 χ^2_{MH} 값이 10.663으로 매우 현격하게 남학생에게 유리한 문항이었다. 체육교과 차별기능문항 관련 연구인 Lee & Yang(2016)과 Lee et al.(2010)의 연구에서 차별기능문항으로 추출된 문항의 최대 χ^2_{MH} 값이 각각 4.560과 4.549인 것을 보면 4번 문항의 차별적 기능은 매우 현저한 것이다. 우도비 분석 방법에서도 4번 문항과 5번 문항의 G^2 의 값은 각각 11.8과 7.9로 차별기능문항으로 판별되었다. 우도비 검정 방법을 통해 차별기능문항을 판별한 No et al.(2010)의 연구결과에서는 문항들의 최대 G^2 값은 7.2이었다. 이는 본 연구의 4번 문항이 우도비 분석 방법을 통한 판별에서도 차별적 기능을 크게 하고 있는 문항이라는 것을 보여준다.

문항반응이론의 문항특성곡선 확인을 통한 차별기능문항의 판별에서 4번과 5번 문항은 모두 균일적인 차별기능을 수행하고 있었는데, MH 방법과 같이 4번 문항은 남학생에게 유리하고, 5번 문항은 여학생에게 유리하였다. 선행연구들(Lee et al., 2010; Lee & Yang; 2016)에서도 체육교과 차별기능문항들은 균일적인 차별기능을 수행하고 있었는데, 이는 4번과 5번 문항의 경우도 동일하였다. 그리고 문항반응이론의 우도비 검정 방법을 통해서도 4번과 5번 문항이 차별기능문항으로 판별되었다. 우도비 검정 방법은 차별기능문항을 판별하는 최근의 선행연구들(Lee et al., 2010; Lee & Yang; 2016)에서는 시도되지 않은 방법이었다.

차별기능문항의 추출과 관련하여 MH 방법은 4번 문항만이 통계적으로 유의하게 차별기능문항으로 판별하였지만, 문항반응이론의 문항특성곡선과 우도비 검정 방법은 좀 더 엄격하게 5번 문항까지도 차별기능문항으로 판별하였다. 즉, 차별기능문항을 추출할 때 문항반응이론을 적용하는 것이 고전검사이론보다 폭넓게 차별기능문항을 탐색하였다.

셋째, 질적 분석을 통해 최종적으로 차별기능문항으로 추출된 문항은 4번 문항이었다. 통계적인 기법을 통해 차별적으로 기능하는 문항군을 추출하는 방법은 두 집단의 피험자들을 대응시켜 능력을 비교하기 때문에 평균에 의한 집단 간 차이 검정에 의한 분석 결과보다 해석이 더욱 타당하지만(Lee, 2004), 차별기능문항 추출과 관련해서는 심도있는 내용 분석이 요구된다. 왜냐하면 통계적 기법을 통해 판별된 차별기능문항의 선정은 문항의 편파성을 보여주는 충분조건이 아니라 필요조건이기 때문이다(Angoff, 1993).

질적 분석을 통한 최종적인 차별기능문항의 추출에서 5번 문항은 제외되고 4번 문항이 차별기능문항으로 추출되었다. 그리고 학교스포츠클럽 활동의 차별적인 운영, 즉 학교체육교육과정의 차별적인 운영이 차별기능문항의 발생 원인으로 유력하게 지적되었다. 이러한 결과는 궁극적으로 차별적인 단위학교 체육교육과정의 수정을 요구한다. 행복중학교가 운영한 차별적인 학교체육교육과정 운영의 방법처럼, 학교스포츠클럽운영을 성별에 따라 운영하는 방법은 전국의 많은 학교들에서 시행되고 있다. 행복중학교도 학교스포츠클럽에서 일반교과교사들이 여학생들을 지도하면서 다른 학교들처럼 학생들을 방치하는 문제가 발생하였다. 그리고 남학생들은 전문스포츠 강사와 체육교사가 축구를 비롯한 구기를 지도하였다. 이러한 상황은 다른 학교들처럼 학교스포츠클럽 운영에 있어 지도인력 부족의 문제로 발생한다(Kim et al., 2013; Seo et al., 2014). 이런 상황에서 행복중학교처럼 모든 학교에서 체육시간에 차별적으로 운영되었던 종목을 대상으로 교수학습을 진행하고 지필시험을 치르지는 않을 수 있으므로 행복중학교의 사례는 소수일 수 있다. 그리고 학교 현장에서 학교스포츠클럽 운영상에 발생되고 있는 인력 부족의 문제도 그 해결이 쉽지 않은 상황이다. 그럼에도 불구하고 학교스포츠클럽 운영에 있어 남녀학생들 모두에게 균등한 교육기회를 제공해야 한다는 원칙은 현실적인 문제로 포기해서는 안 되는 교육적 가치이다. 특히 체육교과에서는 전통적인 성차에 대한 고정관념과 사회문화적인 사회화 과정의 차이로 남녀학생들이 신체활동 참여와 인식에 차이가 나는 것은 주지의 사실이다. 따라서 남녀 학생들에게 공평한 교수학습기회를 제공하는 것은 차별기능문항의 발생 가능성

을 줄이는 측면에서도 매우 중요하다.

물론 남녀학생들에게서 나타나는 사회문화적인 배경의 차이로 발생하는 차별기능문항의 발생은 짧은 기간의 교수학습만으로 극복할 수 있는 쉬운 문제가 아니다. 그러나 단위학교 체육교육과정을 운영할 때 주의를 기울인다면 차별기능문항이 발생을 최소화할 수 있을 것이다. 구체적으로 체육수업상황과 학교스포츠클럽 활동상황에서 대안을 찾을 수 있는데, 체육수업상황에서는 수업과정에 초점을 두는 방식으로, 여학생들을 위해 스포츠 종목의 다양한 상황에 대한 예시를 구체적으로 제시하는 것과 여학생들이 남학생들과의 혼성팀에서 리더 역할을 하도록 하여, 수업에 적극적인 참여자가 되도록 유도하여 여학생들과 체육교사의 의사소통의 기회를 확대하는 것이다. 이는 정부시책인 여학생체육활성화(Ministry of Education, 2016)와 맥을 같이 한다. 그러나 체육시간에 여학생들이 상대적으로 남학생들보다 체육교사와 의사소통 기회를 충분히 갖지 못하는 것은 지필시험에서 차별기능문항을 유발하는 원인이 될 수 있다. Lee et al.(2010)과 Lee & Yang(2016)에서 차별기능문항이 발생한 원인은 모두 체육교사가 운동능력이 좋은 남학생 위주로 의사소통을 하는 것이었다. 이는 Francis(2000)가 지적하였듯이, 체육교과 성차별적인 수업과정의 대표적인 사례는 체육교사들이 체육수업에서 여학생보다 운동능력이 좋은 남학생들에게 더 많이 피드백을 주는 경향이 있다는 주장과 맥을 같이한다.

학교스포츠클럽 운영과 관련해서는 종목에 따라 남녀학생들을 분리하여 운영하는 것이 아니라 혼성으로 운영하는 것을 적극적으로 고려해야 한다. 예를 들어 학교스포츠클럽의 활동 종목을 남학생은 축구·농구, 여학생은 피구로 배정하는 것은 차별기능문항을 유발하는 상황적 요인이 될 수 있기 때문인데, 이러한 차별적 종목운영은 Lee et al.(2010)의 연구에서도 문제로 지적되었다. 경기도용인교육지원청에서는 2013년부터 관내 학교스포츠클럽대회 축구종목을 남녀혼성으로 운영하였다. 경기도교육청에서는 차별적인 체육교육과정 운영을 극복하고 동시에 여학생체육활성화를 목적으로 2016년부터 교육감배 학교스포츠클럽대회에서 초등학교 6학년, 중학교 3학년, 고등학교 2학년들을 대상으로 혼성축구대회를 운영한다. 이 대회에 나올 수 있는 혼성팀들은

한 학급을 팀으로 운영하여 해당학교에서 학급대항 리그 전 대회에서 우승하고 다시 지역 교육청대회에서 우승한 학급이다. 구체적인 경기방식은 상대팀이 서로 전반전은 여학생들끼리, 후반전은 남학생들끼리 경기를 하는 것이다. 경기도교육청의 이러한 정책은 여학생체육을 활성화할 뿐만 아니라, 단위학교 체육교육과정 운영에서 성별에 따른 차별을 최소화하도록 유도하고 있다.

정규 학교체육교육과정에서 남녀 불평등의 발생은 사회적인 선입관과 이에 대한 무비판적 순응 또는 현실적인 한계로 인한 교사의 의지 부족과 무관심이 그 원인이 될 수 있다. 그러나 본 연구의 경우처럼 학교 현장에서 성차별적인 체육교육과정 운영 사례가 발견되더라도, 현장 교사들이 다음연도 단위학교 체육교육과정을 계획·운영할 때 양성평등을 실현할 수 있도록 하는 구체적인 운영 매뉴얼은 없는 상황이다. 이런 상황은 본 연구에서 성차별적인 교육과정 운영으로 지적된 단위학교 학교스포츠클럽 운영에서도 유사하게 나타난다. 일반적인 학교스포츠클럽 운영 매뉴얼의 부재문제는 연구자들(Kim et al., 2013; Seo et al., 2014)을 통해 지속적으로 지적되었다. 그러나 2013년부터 시작된 교육부 산하 학교스포츠클럽 리그운영 지원센터를 중심으로 최근에 다양한 종목과 단위학교 및 지역단위에서 학교스포츠클럽을 운영하기 위한 다양한 정보가 지원되고 있다. 그러나 이곳에서 지원되는 학교스포츠클럽 운영 매뉴얼에서도 양성평등 관점에서 학교스포츠클럽을 운영하기 위한 구체적인 내용은 언급되지 않고 있다. 추후에 이런 운영 매뉴얼이 개발될 때에는 미국의 Title IX처럼 남녀학생들이 공평하게 체육활동에 참여하고 즐길 수 있도록 하는 가이드라인이 포함되어야 할 것이다.

지금까지 논의된 내용을 바탕으로 다음과 같이 후속 연구를 제안하고자 한다.

첫째, 본 연구에서는 차별기능문항 추출 방법으로 선행연구들(Lee et al., 2010; Lee & Yang, 2016)에서 적용하지 않았던 우도비 검정 방법을 사용하여 검정 방법의 폭을 확장하였다. 후속연구에서는 차별기능문항의 다양한 속성들을 고려한 추출 기법들을 체계적으로 적용하는 것이 필요하다. 이를 위하여 Dorans & Potenza(1994)가 제안한 대응변수의 속성(관찰점수, 잠재변수)과 문항 수행을 결정하는 방법(모수적 방법,

비모수적 방법)에 따른 기법들, 그리고 균일적 차별기능문항과 비균일적 차별기능문항을 판별할 수 있는 다양한 방법들을 적용할 수 있다.

둘째, 본 연구는 한 개의 단위학교에서도 일부 학생들을 연구대상으로 하였다. 따라서 연구결과 해석을 일반화하는 데는 주의가 필요하다. 추후에는 차별적인 학교 체육교육과정 운영 속에서 여러 급의 다양한 학생들을 대상으로 한 후속 연구가 필요하다.

셋째, 본 연구에서 차별기능문항으로 추출된 문항은 내용 타당도가 낮고, 인지행동 유목이 적용수준이면서 특정동작 상황을 표현하는 문항이었다. 이러한 결과는 차별기능문항군 연구가 필요할 수 있음을 짐작하게 하는 대목이다. 따라서 앞으로는 차별기능문항군 연구와 그 원인을 탐색하는 후속 연구가 필요하다.

결론

본 연구의 연구결과와 논의를 종합하여 결론을 내리면 다음과 같다.

첫째, 문항들의 일반적인 문항양호도는 적정수준이었으나 일부 문항들에서는 남녀별 차이가 나타났으며 적용수준의 지필문항(4번 문항)은 여학생들에게 보다 어려운 문항이 될 수 있음을 알 수 있었다.

둘째, 차별기능탐색과 관련하여 문항반응이론을 활용한 차별기능탐색 방법(문항특성곡선, 우도비 검정)이 고전검사이론을 적용한 방법(MH방법)보다 폭넓게 차별기능문항 탐색을 하였다.

셋째, 최종적으로 차별문항으로 추출된 4번 문항은 단위학교 체육교육과정이 성차별적으로 운영될 때, 지필 검사에서 차별기능문항이 발생할 수 있다는 것을 보여주었다.

지금까지 차별기능문항 추출 연구들은 해당 검사에서 차별기능문항이 발생했을 때, 질적 분석을 통해 그 원인을 추론하고 추후에 있을 교수학습 계획과 운영에 유용한 정보를 제시하는 형태를 유지하였다. 그러나 본 연구는 역으로 연구자들이 차별적인 단위학교 체육교육과정 운영 사례를 발견하고 체육지필검사에서 차별기능문항이 추출되는지를 탐색함으로써, 차별적인 단위학교 체육

교육과정 운영이 지필검사의 공정성을 해칠 수 있다는 것을 보여주었다는데 의의가 있다. 물론 본 연구에서 적용한 문항반응이론은 그 사용에 있어 엄격한 기준(Lord, 1968; Hulin et. al., 1982)이 요구되지만, 본 연구와 같이 단위학교 수준의 자료를 분석하는 선행 연구들(Lee et al, 2010; Lee & Yang, 2016)에서도 문항반응이론을 적용한 결과가 연구자들에게 유용한 정보를 줄 수 있음을 보여주었으며, 이는 본 연구의 결과 해석의 유용성을 뒷받침한다. 물론 본 연구에서는 차별 기능문항 판별에 고전검사이론의 MH 방법도 적용하였는데, 분석결과 4번 문항은 차별기능이 매우 크게 나타났다. 이는 문항반응이론을 적용한 분석결과와 일치하였으며, 차별기능문항 판별에 있어 문항반응이론을 적용한 본 연구 결과의 신뢰성을 뒷받침한다고 볼 수 있다.

참고문헌

- Ahn, D. K. & Eom, K. H. (2014). School Sports and School Activity Satisfaction-Empirical Analyses of Students in Residential International High School. *The Institute of Social Research*, 7(2), 31-51.
- Anderson, L. W. & Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Rahts, J. & Wittroch, M. C. (2001). *A Taxonomy for Learning, Teaching, and Assessment: A Revision of Blooms's Taxonomy of Educational Objectives*. Boston: Allyn & Bacon.
- Agoff, W. H. (1993). Perspectives on differential item functioning methodology. In P. W. Holland & H. Wamer(Eds.), *Differential item functioning*(pp. 3-13). Hillsdale, NJ: Lawrence Erlbaum Association.
- Brooker, R., Kirk, D., Braiuka, S. & Brangrove, A. (2000). Implementing a game sense approach to teaching junior high school basketball in a naturalistic setting. *European Physical Education Review*, 6(1), 7-26.
- Cha, E. J. & Kim, Y. J. (2014). Analysis of the Differences in Personalities of Adolescent Participating in School Sports Clubs. *The Korean Journal of Physical Education*, 53(1), 25-34.
- Cho, H. M. & Lim, S. Y. (2012). Transference of Gender Ideologies, Its Re-production, and Teachers' Dilemma in Physical Education in the Elementary School. *The Korean Journal of the Elementary Physical Education*, 18(2), 1-14.
- Cha, S. J., Lee, T. K., Park, C. H. & Lee, H. J. (2014). Effect of Sportscasting on Cognitive Learning in Physical Education. *Journal of Learner-Centered Curriculum and Instruction*, 14(5), 125-144.
- Choi, C. H. & Park, J. H. (2014). Changes in Item Characteristics of Test Delivery Modes based on Linguistic Comprehension. *The Korean Journal of Physical Education*, 53(3), 793-808.
- Dorans, N. J. & Potenza, M. T. (1994). *Equity assessment for polytomously scored items: A taxonomy of procedures for assessing differential item functioning*. Princeton, NJ: Educational Testing Service.
- Douglas, J., Roussos, L. & Stout, W. (1996). Item bundle DIF hypothesis testing: Identifying suspect bundle and assessing their DIF. *Journal of Education Measurement*, 33(4), 465-484.
- Ebel, R. L. & Frisbie, D. A. (1991). *Essential of Educational Measurement*(5th ed.). Engelwood Cliff, NJ: Prentice-Hall.
- Francis, B. (2000). *Boys, girls and achievement: Addressing the classroom issues*. London: Routledge.
- Green, B. F. Bock, R. D. Humphreys, L. G. & Reckase, M. D. (1984). Technical Guidelines For Assessing Computerized Adaptive Tests. *Journal of Educational Measurement*, 21(4), 347-360.
- Jeong, K. S., Ahn, Y. G. & Shin, K. C. (2008). Exploring the realities and causes of the sexism styles expressed in elementary school teachers' language during the physical education classes. *Korean Journal of Sport Pedagogy*, 15(3), 105-121.
- Joo, J. M. (2015). *Legal and Institutional Proposals to Realize Gender Equity in School Sports: Application of Title IX of the U.S.* *The Journal of Sports and Law*, 18(1), 65-93.
- Kim, I. S. & Lee, H. J. (2004). Descriptive Analysis of Tests Items for Evaluating Cognitive Domain in 8th Grade Physical Education. *The Korean Journal of Physical Education*, 43(4), 251-260.
- Kim, K. C., Kim, T.C. & Heo, H. M. (2013). A Study on the Current State of and Improvement Measures for School Sports Club Activities in Creative Activities. *The Korean*

- Journal of Physical Education*, 52(4), 61-74.
- Kim, M. Y. & Cho, M. H. (2006). The Experience Sport Education and Gender Equality by Female Students. *Korean Journal of Sports Pedagogy*, 13(3), 145-161.
- Kim, S. Y. (2001). *Exploration of Differential Item Functioning*. Seoul: Kyoyook Book.
- Kim, Y. H. & You, J. A. (2004). Analysis of Sex Equity and Stereotype in K-12 Physical Education Textbooks Based on the 7th National Curriculum. *Journal of Korean Association of Physical Education and Sport for Girls and Woman*, 18(2), 11-22.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hulin, C. L., Lissak, R. I. & Drasgow, F. (1982). Recovery of two- and three-parameter logistic item characteristics curves: a Monte Carlo Study. *Applied Psychological Measurement*, 6, 249-260.
- Hwang, S. H. & Kim, H. S. (2012). Effects of the After School Sports Club Activity on the School Life Satisfaction in the Middle School Students. *Journal of The Korea Contents Association*, 12(12), 771-778.
- Lee, D. Y., Kim, S. W. & Gil, I. J. (2014). Exploration of Gender-Related Differential Bundle Function on Career Aptitude Test for Design(CATD). *Journal of Educational Evaluation*, 27(4), 945-964.
- Lee, G. W. & Won, M. A. (2013). The Influence of Physical Self-Concept, Ego-Resilience and School Happiness of School Sports Club Activity Participation Middle School Student. *Korean Society for The Study of Physical Education*, 18(1), 147-161.
- Lee, K. H. (2004). *Investigation of Gender Differences in College Scholastic Ability Test(CSAT) Mathematics Section by using Differential Bundle Functioning Method*. In Prehypertension Master Dissertation, Ewha Womans University.
- Lee, M. K., Sohn, W. S. & No, U. K.(2007). Analysis of Open Question and Evaluation Frame: PISA 2000, PISA 2003, PISA 2006. *KICE Research Report ORM 2007-25*. Seoul: KICE.
- Lee, M. Y. (2014). The Role of Female Leader in Elite Sports Department of the Creative Economy Policy. *2014 Korean Physical Education Association for Girls and Women Spring-Conference Proceedings*, 69-80.
- Lee, T. K., Park, I. Y. & Lee, H. J. (2010). Goodness of Multimedia Test Program for Assessing Basketball Knowledge. *The Korean Journal of Measurement and Evaluation in Physical Education and Sport Science*, 12(2), 13-29.
- Lee, T. K. & Yang, H. W. (2016). Examination of Goodness of School-based Paper and Pencil Test Applying Classical Test Theory and Three-Parameter Logistic Model in Physical Education: Facilitating Female Student's Participation in Sport. *The Korean Journal of Educational Methodology Studies*, 28(1), 1-27.
- Lord, F. M. (1968). An analysis of the verbal scholastic aptitude test using bimbaum's three-parameter logistic model. *Educational and Psychological Measurement*, 26, 989-1020.
- No, U. K., Kim, J. H. & Kim, S. J. (2010). Detecting the Differential Item Functioning by Gender on the Problem Solving Literacy of PISA 2003. *The Korean Association for Educational Methodology Studies*, 22(4), 165-194.
- Mallenbergh, G. (1989). Item bias item response theory. *International Journal of Educational Research*, 13(2), 127-143.
- Metzler, M. W. (2005). *Instructional models for physical education*(2nd ed.). Boston: Ally & Bacon.
- Ministry of Education (2016). *2016 Promotion Plan of Activation for School Physical Education*. Seoul: Ministry of Education.
- Park, J. Y. (2011). Analysis of question type and contents of middle school physical education examination. *Korena Journal of Sport Pedagogy*, 18(1), 33-50.
- Roussos, L. A. & Stout, W. F. (1996). A multidimensionality-based DIF analysis paradigm. *Applied Psychological Measurement*, 20, 353-371.
- Seo, J. Y., Kim, K. C., Yu, C. W., Kim, T. C. & Heo, H. M. (2014). A Performance Analysis of School Sports Club Activity. *KICE Research Report RRC 2014-11*. Seoul: KICE.
- Seong, T. J. (1994). Detecting the Differently Functioned Items by Gender on the 1st Tryout of the 1994's College Scholastic Ability Test. *Journal of Educational Evaluation*, 7(2), 87-101.

- Seong, T. J. (2002). *Modern educational evaluation*. Seoul: HAKJISA corp.
- Seong, T. J. (2009). *Application and Understanding of Item Response Theory*. Seoul: KYOYOOK Book.
- Sohn, W. S. (2012). Application of Differential Bundle Functioning Techniques for Understanding Sources of Differential Functioning. *Secondary Institute of Education, 60(4)*, 917-935.
- Yeo, J. K. & Nam, S. W. (2009). A Development of Scale on Sources of Reverse Discrimination of Gender in Coeducational Physical Education Class. *Korean Journal of Sociology of Sport, 22(3)*, 91-103.
- Youn, H. S. (2013). A narrative inquiry on what a middle school girl student who was once 'Wangtta' at school gained while she worked as a manager of the school baseball club. *Korean Journal of Sport Pedagogy, 20(2)*, 159-176.

문항반응이론의 차별기능문항 분석을 활용한 단위학교 차별적 교육과정 탐색

이태구(연세대학교) · 손지영(수주중학교) · 양희원(연세대학교)

본 연구의 목적은 단위학교에서 실천되고 있는 교육과정 내 스포츠클럽활동이 성별에 따라 차별적인 교육과정으로 운영되었을 때, 학기말에 실행된 지필평가에 출제된 문항들이 차별기능문항으로 판별되는지 분석·확인하는 것이다. 연구대상은 경기도 행복중학교(가명) 2학년 학생 276명(남:135, 여:141)을 대상으로, 1학기 기말고사에서 학생들이 응답한 데이터를 이용하여 집단 간 차별적으로 기능하는 문항을 분석하고 질적 분석을 하였다. 행복중학교 2학년 학생들은 1학년 입학 후부터 2학년 1학기까지 매주 스포츠클럽활동을 실시하였는데, 남학생들은 축구를 비롯한 구기운동을, 여학생들은 피구를 실천하였다. 그리고 2학년 1학기에 축구를 체육시간을 통해 학습하고, 2학년 1학기 기말고사에 축구관련 다섯 문항이 출제되었다. 분석결과, 고전검사를 적용한 방법(MH 방법)과 문항반응이론을 적용한 방법(문항특선곡선 비교방법, 우도비 검정 방법) 모두에서 4번과 5번 문항이 차별기능문항으로 판별되었으며, 심도 있는 질적 분석을 통해 최종적으로 4번 문항이 여학생에게 불리한 차별기능문항으로 추출되었다. 내용타당도가 낮고 인지행동 유목이 적용수준인 4번 문항이 차별기능을 수행하는 것으로 추출된 본 연구의 결과는, 체육지필검사에서 적용수준인 문항이 남학생보다 운동종목의 경험이 적은 여학생에게 차별적으로 기능할 수 있음을 보여준다. 이와 같은 결과는 앞으로 단위학교에서 체육교육과정을 계획하고 운영할 때, 남녀에게 교육과정이 공정하게 운영될 수 있도록 세심한 주의가 필요함을 시사하고 있다.

주요어: 문항반응이론, 차별기능문항, 체육교육과정, 학교스포츠클럽