

리듬운동학문분야 적용을 위한 구조방정식모형의 이해*

김응준**

한국체육대학교

A research on the understanding of structural equation modeling for rhythmic exercise study

KIM, Eung-Joon

Korea National Sport University

Abstract

The purposes of this study are to explain the concept of factor analysis used frequently in rhythmic exercise study through focusing on the misuse of principal component analysis and to introduce both basic concepts and theories in order to help research beginners from the rhythmic exercise study. Document analysis was adapted for a research method. As a result, this research firstly found that principal component analysis was not a common factor analysis. Secondly, this study presented the basic concepts and definitions for the application of the structural equation modeling. Thirdly, six points which are the key points in the structural equation modeling were confirmed. Finally, the study suggested the examples of the structural equation modeling.

Key words : common factor model, confirmatory factor analysis, structural equation model, SEM

I . 서론

리듬운동학문 분야에서 과학적 탐구의 목적은 연구자가 관심을 가지고 있는 현상에 대한 기술, 설명, 예측, 통제에 있다. 연구자가 리듬운동을 이해하고 설명

하는 것으로 만족하고자 한다면 리듬운동을 구성하고 있는 변수들 간의 원인-효과관계, 즉 인과관계에 대한 법칙을 확보하는 것이 필수적으로 요구되지 않는다. 그러나 리듬운동이라는 현상에 대한 예측이나 통제가 목적이라면 리듬운동과 관련된 변수들간의 원인-효과

* 이 논문은 2015년 한국리듬운동학회 추계학술발표내용을 보완 정리한 논문임.

** ejkim@knsu.ac.kr

관계에 관한 법칙을 확보하여야 한다. 다시 말하면 선행조건 ㉠이 원인일 때, 그에 대한 효과로 ㉡이 발생한다는 법칙이 확보 되었을때에 비로소 리듬운동 현상을 예측할 수 있고 리듬운동 현상을 통제할 수 있다는 말이다.

실험설계방법에서 두 집단의 특성을 동일하게 한다는 방법으로 무선헌당이 요구되어지나, 교육현장이라는 학교에서는 학급편성을 무시하고 무작위로 실험집단과 통제집단에 배정하기는 어려운 실정이다. 실험에 요구되는 선행조건들을 만족시키기 어려운 것이 사회과학의 현실이다. 또한 실험이 완벽하게 수행되어도 연구설계나 실험에 동원된 변수의 수가 많아지면 변수간의 인과관계에 관한 추론을 이끌어 내기가 어렵다. 변수의 수가 많아지면 어떤 변수가 종속변수에 어느 정도 영향력을 미치는가는 수학적으로 확인이 되지만, 실제적으로 어떤 의미를 갖고 해석이 가능한가는 또 다른 문제이다.

최근 10여년 동안 체육학문은 연구방법 면에서 비약적인 발전이 되어왔다. 체육학문 본연의 모습으로 발전해온 영역도 많지만, 다른 타 학문분야의 기법을 적용하여 이론 업적도 이루 헤아릴 수 없을 정도이다. 수많은 기법들 중 영향력이나 사용된 횟수로 구조방정식모형이 으뜸일 것이다. 본 연구의 주제이기도 한 구조방정식모형(structural equation modeling)은 완벽하게 통제되는 실험연구가 어렵거나 가능하지 않은 연구상황에서 변수간의 인과관계를 검증가능하게 해주는 통계적 기법이다. 변수간의 인과관계라는 말은 인과관계가 있는 변수들의 구조라고 할 수도 있다.

구조방정식모형은 생성과 발전과정이 여러 학문의 발전에 힘입은바 크다. 구조방정식모형의 근간이라고 할 수 있는 잠재변수모형(latent variable model), 측정모형(measurement model), 경로도(path diagram)은 계량경제학에서의 회귀분석, 심리측정학에서의 요인분석, 그리고 생물측정학에서의 경로분석 토대 위에서 생성되었다(이기중, 2000). 최근 구조방정식모형에 대한 서

적들이 수없이 많이 출판되어있으며, 계속 출판되고 있는 실정이다. 본 연구에서 구조방정식의 생성과 역사적 발전을 언급할 필요가 없다고 판단하고 이는 다른 교재를 읽어보시라고 추천한다.

리듬운동학문분야에 구조방정식을 적용하기 위한 기초이론으로 회귀분석, 탐색적요인분석, 확인적요인분석, 경로분석을 필수적으로 이해하여야 적용할 수 있다. 본 연구의 목적은 첫째, 탐색적 요인분석에서 주성분분석과 공통요인분석을 개념적으로 비교하여 주성분분석이 요인분석이 아니라는 개념을 제시하고, 둘째, 구조방정식모형의 기초로 ① 구조방정식모형의 주요용어, ② 구조방정식모형을 설정할 때 잊지 말아야 할 사실들, ③ 구조방정식 모형의 예를 제시하는 것이다.

II. 탐색적 요인분석

요인분석에서 공통요인모형(common factor model)의 기본적인 개념은 측정변수 사이의 관계를 소수의 잠재변수로 요약하여 설명하는 것을 의미한다. 즉, 측정변수 사이의 관련성을 설명하는 이론구조(theoretical construct)를 개발 혹은 검증하는 것이 공통요인모형의 주된 목적이다(Fabrigar, Wegener, MacCallum, & Strahan, 1999). 공통요인모형을 구성하는 방법론적 개념은 측정변수 간의 상관관계를 소수의 잠재변수를 통해 설명한다는 점, 그리고 이를 통해 실재하는 요인을 통계적으로 검증한다는 점, 요인 추정의 반복가능성(재현가능성)을 판단할 수 있는 정보를 제공한다는 점, 이러한 일련의 과정을 통해 이론의 개발 및 발전에 기여한다는 장점을 지닌다(Brown, 2006; DeVellis, 2012). EFA 모형은 다양하게 존재하지만, 권장되는 추정방법은 주축분해법(Principal Axis Factoring: PAF)과 최대우도법(Maximum Likelihood: ML)으로 분류된다. 두 방법

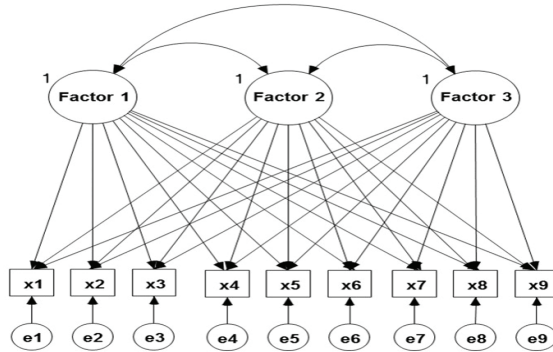


그림 1. 탐색적 요인분석 모형

에 대한 선택기준은 논란이 있지만, 두 방법이 다른 EFA 추정방법보다 가장 정확한 추정을 한다는 것은 관련 연구자들의 공통된 합의이다(Costello & Osborne, 2005; Fabrigar et al., 1999; Hayton et al., 2004).

EFA 모형을 통계적 도식으로 나타내면 <그림 1>과 같다. <그림 1>에서 $x_1 \sim x_9$ 는 측정변수, $factor1 \sim factor3$ 는 공통요인(즉, 잠재변수, 요인), $e_1 \sim e_9$ 는 고유분산(unique variance, error variance)을 의미한다. 즉, 공통요인분석에서 독립변수는 잠재변수가 되고, 종속변수는 측정변수가 되지만, 독립변수가 숨어 있는 잠재변수이므로, 분석 과정에서 이를 추정하게 된다(이순목, 2000; Fabrigar et al., 1999). 따라서 공통요인모형의 기본적인 개념은 잠재변수가 항상 독립변수라는 점이다. 잠재변수가 이미 존재하고 있고, 존재하고 있는 잠재변수가 현실(측정변수)을 설명한다는 점에서 행동·사회과학의 기본 철학을 반영하는 현실적 모형이라 볼 수 있다(Brown, 2006).

상기와 같이 EFA는 공통요인모형을 반영하면서, 특히 척도개발에 있어서 요인 수(잠재변수의 수)를 결정하는 과학적 검증방법이라 볼 수 있다. 그러나 EFA는 요인에 대한 주관적 해석, 객관적이 아닌 주관적인 요인 수 결정, 요인분석의 이해 부족, 이론적 근거 없는 요인회전 등으로 가장 오용되는 방법이기도 하다(Fabrigar et al., 1999). 실제로 행동·사회과학 연구자들이 가장 많이 EFA로 적용하는 주성분 분석(Principal

Component Analysis: PCA)은 단순히 자료 축소의 방법이지 EFA가 아니다(Brown, 2006; Costello & Osborne, 2005; DeVellis, 2012; Fabrigar et al., 1999; Hayton et al., 2004). PCA가 EFA가 될 수 없는 이유는 다음과 같다.

먼저 PCA의 목적은 소수의 주성분으로 측정변수의 분산 설명을 최대화 하는 데 있다. 즉, 주성분의 개념은 측정변수들의 분산의 총합점수라 볼 수 있기 때문에, 이론적 배경을 고려하지 못하고, 자료 의존적이기 때문에 일반화 및 요인의 재현성이 제한된다(Fabrigar et al., 1999). <그림 2>에 제시한 PCA의 통계적 도식에서 $PC1$ 을 설명하는 수학적 관계는 $PC1 = b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9$ 으로 표현할 수 있다. 위의 식에서 $b_1 \sim b_9$ 는 측정변수가 주성분 $PC1$ 을 설명하는 회귀계수이지만, EFA와는 달리 독립변수는 측정변수가 되고, 종속변수는 각 성분이 된다. 또한 PCA에서는 첫 번째 주성분($PC1$)을 먼저 추출하고, 잔여상관행렬에서 두 번째 주성분($PC2$)을 추출하게 된다. 즉, PCA는 측정변수들의 분산의 총합점수라는 점, 요인이 독립변수가 아닌 종속변수라는 점, 요인 추출 시 가중치가 부여된다는 점에서 다음과 같은 문제점을 지닌다. 첫째, $PC1$ 을 먼저 추출하고 잔여상관행렬에서 $PC2$ 를 가중치를 주어서 추출하기 때문에, 주성분들이 추출되는 순서에 따라 분산의 크기가 감소한다는 점이다. 따라서 두 번째 이후 추출되는 주성분은 분산의 설명력이 아닌

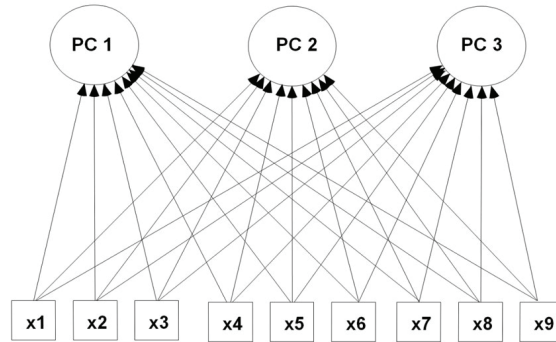


그림 2. 주성분 분석 모형

잔여량의 개념이라 볼 수 있다(Costello & Osborne, 2005). 둘째, 동일한 변인들로부터 추출되는 주성분들은 서로 독립적이기 때문에, 주성분 간의 관련성이 없게 된다. 따라서 주성분 분석 후에 회전(rotation)을 실시하는 것은 개념적으로 수행할 필요가 없게 된다(이순목, 2000). 셋째, 고유분산을 고려하지 않고, 고유분산이 포함된 상태에서 분석을 진행하기 때문에, 오차가 포함된 주성분을 추출하게 된다(Fabrigar et al., 1999). 넷째, 주성분 분석에서 추출된 주성분은 단순히 변인이 합해져서 축약된 것에 불과하므로, 잠재변수나 요인을 고려하지 않는다는 점이다(Fabrigar et al., 1999). 따라서 주성분분석으로 척도개발을 진행한 결과물과 요인분석에 대한 증거를 제시하지 않은 결과물은 타당하지 않다.

Ⅲ. 구조방정식 모형의 기초이론

구조방정식모형의 기초이론에서 언급되는 모형도와 그림은 2006년 한국상당심리학회 동계학술연구교제인 <상당심리연구의 최근 동향과 연구방법론>에서 인용되었습니다.

1. 구조방정식 모형의 주요 용어

구조방정식모형은 요인분석과 경로분석의 결합으로 정의될 수 있다. 구조 방정식 모형이 널리 쓰이는 이유는 첫째, 측정오차를 통제가능하다. 둘째, 매개변수의 사용이 용이하다. 셋째, 이론모형에 대한 통계적 평가가 가능하다는 것이다.

1) 잠재변수와 측정변수

첫번째로 알아야 하는 용어는 잠재변수(latent variable)와 측정변수(measured variable)이다. 아래 <그림 3>는 잠재변수와 측정변수의 관계이다.

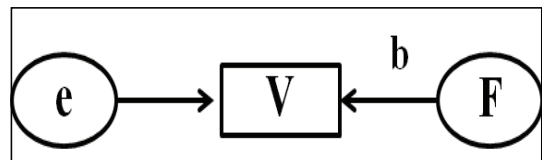


그림 3. 잠재변수, 측정변수의 관계

V는 측정변수, e는 측정오차, F는 잠재변수이다. 잠재변수와 측정변수의 관계에서 $V = bF + e$, 표준화된 $b^2 =$ 측정변수의 신뢰도를 의미한다. b는 요인분석에서 요인계수와 같다.

2) 내생변수와 외생변수

두 번째로 알아야 하는 용어는 내생변수(endogenous variable)와 외생변수(exogenous variable)이다. 외생변수는 다른 변수에 영향을 주는 변수이며, 내생변수는 다른 변수로부터 영향을 받는 변수이다. 아래 <그림 4>에서 변수 1은 외생변수이고, 변수2, 변수3은 내생변수이다.

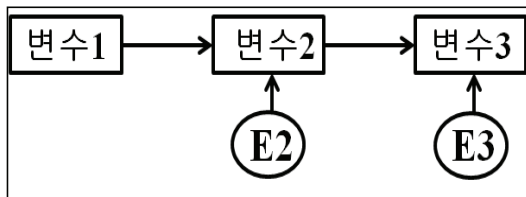


그림 4. 내생변수와 외생변수

3) 오차

오차는 모든 내생변수에 대하여 설정하여 준다. <그림 4>에서 E2, E3은 변수2와 변수3이 내생변수이므로 오차를 가지게 된다. 왜냐하면 내생변수는 다른 변수에 의하여 설명이 되므로 항상 오차를 수반하는 것이다.

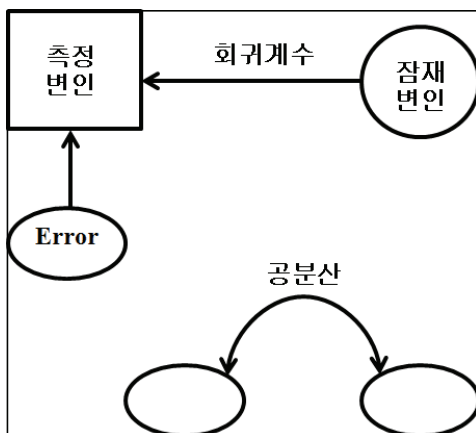


그림 5. 경로모형에서의 기호

<그림 5>에서 잠재변인이라는 것은 요인분석에서 요인 또는 Factor 라고 한다.

4) 미지수(parameter)

미지수(parameter)는 추정하고자 하는 값으로 정의된다. 미지수는 자유미지수와 고정 미지수로 나누어진다.

① 자유미지수

자유미지수란 회귀계수, 변량, 공변량을 의미한다. 알아두어야 할 것은 내생변수의 변량과 공변량은 자유미지수가 될 수 없다. 그러나 외생변수의 공변량은 자유미지수가 될 수 있으며, 외생변수의 변량은 반드시 미지수가 되어야 한다.

② 고정미지수

고정미지수라는 것은 미지수이긴 하지만 계수를 일정한 값으로 고정시키는 것이다. 첫째, 오차의 계수는 항상 1.0으로 고정시킨다. 즉 $y = a + bx + e$ 으로 확인된다. 둘째, 잠재변수에 척도를 부여하기 위하여 고정된 값이 필요하다.

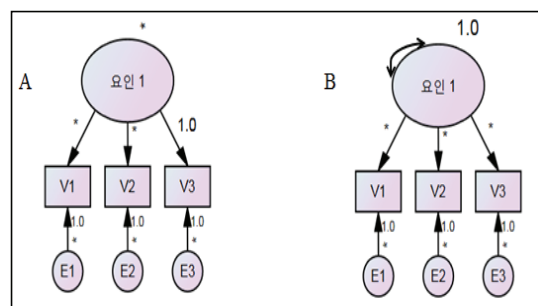


그림 6. 잠재변수에 대한 척도의 결정

과거 전통적인 회귀분석에서는 독립변수와 종속변수의 척도가 다름에도 측정변수를 사용하여 분석하였다. 그러나 구조 방정식 모형에서는 <그림 6>에서 보

는 바와 같이 잠재변수를 사용하여 회귀분석을 처리한다. 즉 척도비교를 위하여 고정된 값 1.0을 부여하여 상대적인 추정을 하는 것이다. 셋째, 고정미지수는 특정 미지수의 추정치(경로계수 또는 요인계수)에 대한 가설검증에서 사용된다.

5) 자유도

구조 방정식 모형에서 자유도란 모형이 간단한 정도라고 정의할 수 있다. 자유도는 변수의 수가 정해져 있을 때, 미지수의 수가 줄어들수록 자유도는 커진다. 자유도의 개념은 모형의 적합도를 평가하는데 매우 중요하다. <그림 7>은 측정변수의 수를 p 라고 하고 자유미지수의 수를 q 라고 하였을 때 자유도를 구하는 방법을 도식화 한 것이다.

● 최대 가능한 미지수의 수 : $\frac{p(p+1)}{2}$

$p=4$ 이면, $\frac{p(p+1)}{2} = \frac{4(5)}{2} = 10$

p : 측정변인의 수
 q : 자유미지수의 수

● 자유도 : $\frac{p(p+1)}{2} - q$

그림 7. 자유도 계산공식

6) 측정모형과 구조모형

측정모형(measurement model)은 잠재변수와 측정(표시)변수 사이의 관계이며, 구조모형은 각 변수의 구조적관계이다.

측정모형은 요인분석모형이라고도 하며 <그림 8>에서 점선으로 표시된 부분이다. 구조모형은 관계모형이라고도 하며 <그림 8>에서 실선으로 표시된 부분이다.

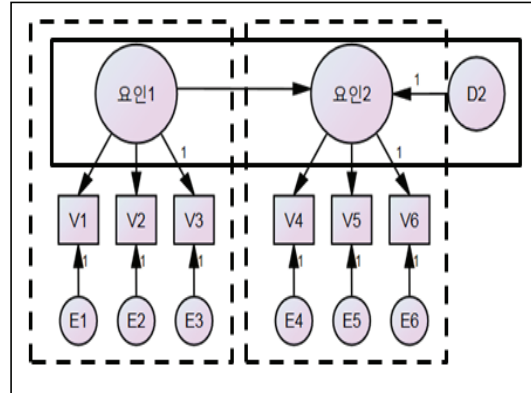


그림 8. 측정모형과 구조모형

2. 구조방정식 모형설정 시 잊지 말아야 할 사항들

본 연구에서 구조 방정식 모형을 리듬운동학문에 적용하는데 있어서 가장 중요한 것으로 리듬운동 현상에 구조 방정식 모형을 설정할 때 반드시 지켜야 할 사항들을 기술하고자 한다.

첫째, 외생변수에 대해서만 공변량을 자유미지수로 설정한다.

둘째, 모든 내생변수에 대해 오차변수를 설정한다.

셋째, 외생변수에 대해서는 오차변수를 설정하지 않는다.

넷째, 모든 외생변수(오차변수 포함)의 변량은 자유미지수로 설정한다.

다섯째, 잠재변수에 척도를 부여하기 위해 하나의 요인계수를 1.0으로 고정하거나 잠재변수의 변량을 1.0으로 고정한다.

여섯째, 오차변수의 계수를 1.0으로 고정한다. (홍세희, 2006)

3. 구조방정식 모형의 예

일반적으로 구조 방정식 모형의 예로는 회귀모형, 경로모형, 구조 방정식 모형을 들 수 있다.

1) 회귀모형

일반적으로 구조 방정식 모형의 예로 회귀모형을 $y=a+bx+e$ 라고 하자. 이때 일반적 회귀분석에서는 절편인 a 를 추정한다. 그러나 구조 방정식 모형에서는 절편 a 를 0으로 고정하고 분석한다. 이러한 회귀모형의 문제점은 측정변수의 신뢰도가 완벽하지 않으므로 추정되는 회귀계수가 부정확하다. 또한 회귀모형에서는 매개모형에 대한 평가가 이뤄지지 못한다.

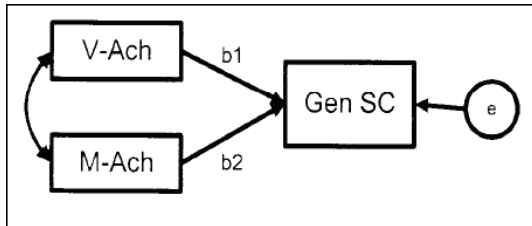


그림 9. 학업에 대한 자아개념

<그림 9>에서 Gen SC은 일반적 자아개념, V-Ach는 언어 성취점수, M-Ach는 수학 성취점수이다.

2) 경로모형

경로모형은 여러 개의 회귀모형이 하나의 틀 안에 놓여 있는 것이다. 측정변수만을 가지고 회귀분석을 하는 것으로, 측정변수의 신뢰도가 완벽하지 않으므로 추정되는 회귀계수 또는 경로계수가 부정확하다.

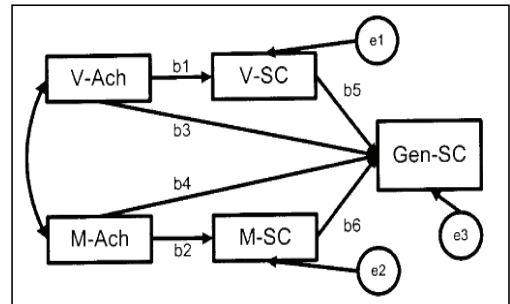


그림 10. 경로모형

<그림 10>에서 V-SC는 언어 자아개념, M-SC는 수학 자아개념이다.

3) 구조 방정식 모형

다음 <그림 11>은 구조 방정식 모형의 예시이다.

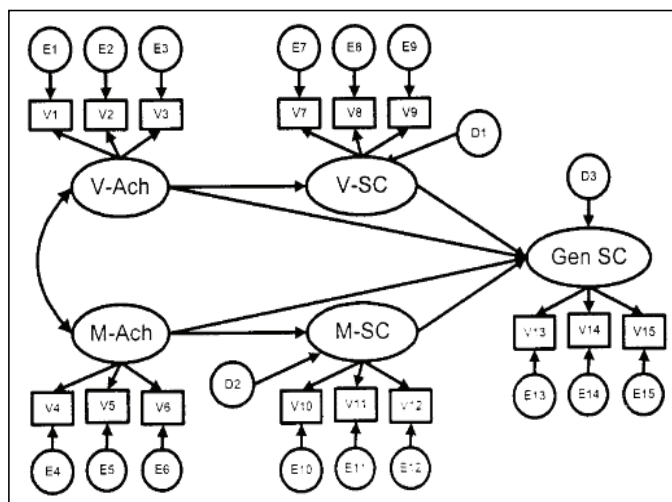


그림 11. 구조 방정식 모형

이외에도 아래와 같이 다양한 모형들이 존재하며, 구조 방정식 모형에서 분석이 가능하다(홍세희, 2006).

■ 다양한 모형들

- 1. 다 집단 분석 (Multi-group analysis)
- 2. 잠재평균분석 (Latent Mean Analysis)
- 3. 매개모형
 - 매개효과에 대한 가설검증
 - Bootstrap 방법을 적용한 매개효과 검증
- 4. 상호작용모형
- 5. 쌍빙모형
- 6. 자동화된 방법을 이용한 모형개발
- 7. 잠재성장 모형

그림 12. 구조방정식의 다양한 모형들

IV. 결론 및 제언

이 연구는 리듬운동학문 분야에서 일반적으로 요인 분석으로 요인추출 방법으로 주성분분석을 실시하여 요인분석이라고 여기는 것이 잘못되었다는 점을 선행 연구를 통하여 알리고, 리듬운동학문 분야에 구조 방정식 모형을 적용시키는데 있어 기초이론을 제공하는 데 연구의 목적이 있었다.

연구의 결론은 다음과 같다.

첫째, 주성분분석은 공통요인분석이 아닌 합점수를 사용하여 요인의 수를 줄이는 목적에만 사용되는 방법으로 공통요인분석이라 볼 수 없다.

둘째, 리듬운동학문분야에 구조 방정식 모형을 적용하기 위해서 기본적인 용어를 제시하였고, 반드시 기억해야하는 사항 6가지를 아래와 같이 제시하였다.

① 외생변수에 대해서만 공변량을 자유미지수로 설

정한다.

② 모든 내생변수에 대해 오차변수를 설정한다.

③ 외생변수에 대해서는 오차변수를 설정하지 않는다.

④ 모든 외생변수(오차변수 포함)의 변량은 자유미지수로 설정한다.

⑤ 잠재변수에 척도를 부여하기 위해 하나의 요인 계수를 1.0으로 고정하거나 잠재변수의 변량을 1.0으로 고정한다.

⑥ 오차변수의 계수를 1.0으로 고정한다.

셋째, 구조방정식 모형의 예시로 회귀모형, 경로모형, 구조방정식 모형을 예시로 제시하였다.

이와 같은 결론으로 다음과 같은 제언을 한다. 첫째, 리듬운동학문 분야에서 구조 방정식 모형을 적용하기 위해서는 우선적으로 변수를 적게 가지고 있는 척도를 중심으로 연구를 진행하는 것이 좋겠다. 왜냐하면, 리듬운동연구가 많은 않은 이유로 사례수가 적어도 변수 갯수가 적은 즉 자유미지수가 적은 모형은 적합하기 때문이다.

둘째, 최근 체육학에서 구조 방정식 모형을 적용한 논문의 경향이 복잡하고 어려운 모형만을 적용하여 투고하고 게재되고 있는 실정입니다만 리듬운동학회 지에서는 앞서 연구결과에서 제시한 회귀모형과 경로 모형도 충분한 의미있는 결과를 제시할 수 있으니 이러한 모형도 분석하여 연구하기를 추천합니다.

참고문헌

- 이기종(2000). 구조방정식모형. 서울: 교육과학사.
- 이순목(2000). 요인분석의 기초. 서울: 교육과학사.
- 홍세희(2006). 구조방정식 모형의 이론과 응용, 연세대학교, 사회복지학과.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological

- research. *Psychological Methods*, 4(3), 272-299.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. NY: Guilford Press.
- DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation*, 10(7), 1-9.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, 7(2), 191-205.

논문투고일 : 2015. 10. 31.

심사일 : 2015. 12. 15.

심사완료일 : 2015. 12. 30.