

AUDIT-K 척도의 요인구조 및 측정불변성 검증: 내포(nested) 모형 비교를 위한 RMSEA_D 지수의 활용을 중심으로

박 병 선¹ | 이 선 영^{1*}

¹ 국립강릉원주대학교

* 교신저자: 이선영 (sylee@gwnu.ac.kr)

초 록

본 연구는 한국복지패널 자료를 활용하여 AUDIT-K의 요인구조 및 성별과 연령대 집단별 측정불변성, 종단적 측정불변성에 대한 검증을 통해 척도의 타당성과 활용성을 확인하기 위한 목적으로 수행되었다. 이 과정에서 최근 주목받고 있는 RMSEA_D를 활용하여 기존의 내포 모형의 비교 방식과 어떠한 결과의 차이가 있는지를 검증하였다. 분석에 활용된 자료는 4차년도 자료를 기준으로 16차년도 자료까지 사용하였다. 본 연구의 주요 결과는 다음과 같다. 첫째, AUDIT-K는 2요인 및 3요인 구조 모두 적합한 것으로 나타났다. 둘째, 성별 및 연령대 집단의 측정불변성은 약한불변성(Weak) 모형까지 유지되는 것으로 나타났다. 셋째, 1년 간격의 5년간 종단적 측정불변성은 엄격한불변성(Strict) 모형까지 유지되었고, 4년 간격의 13년간의 종단적 측정불변성은 강한불변성(Strong) 모형까지 유지되는 것으로 나타났다. 넷째, 내포 모형 비교를 위한 Δ RMSEA와 RMSEA_D를 비교하여 살펴보았을 때, 일부 모형의 비교에서 결과의 차이가 있음을 확인하여 내포 모형 비교 기준의 적용에 있어 주의가 필요함을 제시하였다. 마지막으로 본 연구의 함의와 한계를 제시하였다.

주요 용어: AUDIT-K, 음주 문제, 한국복지패널, 요인구조, 측정불변성, RMSEA_D

알기 쉬운 요약

이 연구는 왜 했을까? 본 연구에서는 알코올 문제와 관련한 수준을 측정하기 위한 도구로서 많이 활용되고 있는 AUDIT-K 척도의 측정이 성별과 연령대별 집단에 따라 그리고 시간의 흐름에 따라 일관성 있도록 이루어지는지를 검증하는 것으로 목적으로 하였다. 이와 동시에 구조방정식 모형의 비교에서 활용될 수 있는 RMSEA_D 지수의 적용 가능성에 대해 살펴보았다.

새롭게 밝혀진 내용은? AUDIT-K 척도를 통해 측정된 음주 문제 수준 점수를 남녀집단과 65세를 기준으로 하는 연령대별 집단에 비교할 수 있음을 확인하였고, 단기 혹은 장기적으로 반복측정된 점수에 대하여 시간의 흐름에 따라 비교하는 것도 가능함을 확인하였다. 또한, 구조방정식 모형의 비교에 있어 기존의 지수를 활용하는 것보다 RMSEA_D 지수의 활용이 조금 더 정확할 수 있음을 확인하였다.

앞으로 무엇을 해야 하나? 향후 AUDIT-K 척도를 활용한 음주 문제 수준을 측정하고 집단과 시점의 비교가 가능하기 때문에 다양한 연구에서 활용될 수 있을 것이다. 또한, 구조방정식 모형의 비교에 있어 RMSEA_D 지수의 적용을 통한 모형의 비교가 활발히 이루어질 필요가 있다.

본 논문은 2023년 한국복지패널 학술대회에서 발표된 원고를 수정·보완하였음.

■ 투고일: 2023. 10. 05.

■ 수정일: 2023. 12. 15.

■ 게재확정일: 2023. 12. 19.

1. 서론

알코올사용장애는 반복적인 음주로 인하여 사회적 또는 직업상의 문제가 발생함에도 불구하고 음주를 지속하는 알코올 남용과 내성과 금단증상이 있는 상태로 자기통제가 불가능한 알코올 의존을 포함하는 정신장애이다(국립정신건강센터, 2021). 이러한 알코올 사용장애는 전 세계적으로도 질병에 따른 사회적 부담과 관련한 가장 중요한 위험요소 중 하나이기 때문에 알코올의 사용을 정확하게 측정하고 잠재적인 알코올 관련 문제를 조기에 선별하고 식별하는 것은 공공보건의 측면에서 상당히 중요하다고 할 수 있다(Skogen et al., 2019).

이를 위해 1989년 세계보건기구(WHO)에서는 개인의 음주량 및 빈도, 음주로 인한 폐해 등을 다면적으로 평가하여 알코올 사용의 문제를 선별하기 위하여 AUDIT(Alcohol Use Disorders Identification Test)라는 척도를 개발하여 보급하였다(Babor et al., 2001). 이 척도는 알코올사용장애의 조기 선별뿐 아니라 유해음주 및 위험음주 등을 선별할 수 있도록 10개의 문항(3개의 하위요인)으로 구성되어 있으며, 다수의 국가에서 활용하고 있어 국제적으로도 비교가 가능한 척도로 알려져 있다(Erford et al., 2021). 최근 AUDIT 척도의 활용 및 타당화에 관한 연구들에서 원칙도가 제시한 3요인의 구조가 아닌 다른 요인구조를 제안하는 연구가 증가하고 있으며, 다양한 집단에 관한 측정불변성에 관한 논의도 다수 제시되는 등 척도의 타당화 및 활용성을 높이기 위한 노력들이 많이 이루어지고 있는 실정이다(Erford et al., 2021; Peng et al., 2012; Moehring et al., 2018; Skogen et al., 2019; Thorisdottir et al., 2020; Tulliao et al., 2016). 그러나 국내의 경우에는 AUDIT 척도에 대한 한국판 버전으로 AUDIT-K에 대한 신뢰도 및 타당도를 검증한 이병욱 외(2000)의 연구를 제외하고는 척도의 타당화에 관한 체계적인 논의가 부족한 것이 사실이다. 한국복지패널에서는 4차년도부터 AUDIT-K 척도를 활용하여 음주관련 사항을 지속적으로 측정하여 자료를 제공하고 있으며, 이 자료를 활용하여 횡단적, 종단적 측면에서의 음주 문제와 관련된 다양한 연구들이 국내에서 진행되고 있다. 하지만, 한국복지패널에서 사용되고 있는 AUDIT-K 척도의 경우, 요인구조에 대한 검증 및 측정불변성과 관련한 체계적인 논의가 이루어지지 못했다. 구체적으로, 특정한 척도를 활용하여 집단 간 및 시간의 흐름에 따른 비교를 위해서는 집단 간 또는 시간의 흐름에 따른 측정이 일관되게 이루어지

고 있다는 측정불변성에 관한 경험적 증거가 있어야 함에도 불구하고 국내에서 사용되는 AUDIT-K 척도의 경우 이러한 체계적 검증이 이루어지지 않은 것으로 파악된다(Leitgöb et al., 2022). 이러한 측면에서 한국복지패널 자료를 활용한 AUDIT-K 척도의 요인구조 및 횡단적 및 종단적 측정불변성에 관한 검증을 바탕으로 AUDIT-K의 국내 적용에 관한 타당성 및 활용성에 관한 경험적 근거의 마련이 필요하다고 할 수 있다.

한편, 구조방정식을 수행하는 과정에서 내포 모형을 비교하는 방법과 관련하여 카이제곱 차이 검증, 근사 적합도 지수의 비교 등 다양한 논의들과 더불어 발전하고 있는데, 최근 RMSEA의 차이값을 중심으로 모형을 비교하는 방법에 대한 한계를 지적하며 χ^2 차이 검증 등을 고려한 RMSEA의 계산을 통한 모형 비교를 시도하는 연구가 증가하고 있다(Brace, 2020; Lasker, J., 2022; Orjiakor et al., 2023; Savalei et al., 2023). 초기 구조방정식 모형의 비교를 위해 사용되던 χ^2 차이 검증 표본의 크기에 민감하다는 한계로 인하여 RMSEA와 CFI와 같은 근사적합도(approximate fit) 지수를 활용한 모형의 비교가 증가하고 있다(Chen, 2007; Cheung & Rensvold, 2002; Savalei et al., 2023). 그러나 RMSEA와 CFI의 차이값을 중심으로 한 내포 모형의 비교 또한 일정한 한계가 있음을 제시하며 χ^2 차이 검증의 값과 자유도의 차이 등을 고려한 RMSEA, 즉 $RMSEA_D$ 를 활용한 모형의 비교를 제안하였다(Brace, 2020). $RMSEA_D$ 를 활용할 경우, 기존 RMSEA와 동일한 기준을 적용할 수 있으며, CI(Confidence Interval)의 계산이 가능하다는 것 등의 장점이 있어 χ^2 및 RMSEA의 차이를 중심으로 하는 분석의 한계를 극복할 수 있다고 하였다(Savalei et al., 2023).

이상에서 살펴본 바와 같이 본 연구의 목적은 크게 두 가지로 제시할 수 있다. 첫째, 한국복지패널에서 사용되고 있는 AUDIT-K 척도의 요인구조, 성별 및 연령대별 집단에 따른 측정불변성, 단기 및 장기적 측면의 종단적 측정불변성의 검증을 통해 척도의 타당성과 활용성을 경험적으로 확인하고자 한다. 둘째, 구조방정식 모형 비교 및 측정불변성 검증과 같이 내포 모형에 대한 비교를 위해 주로 사용되고 있는 RMSEA 차이값과 최근에 제시된 $RMSEA_D$ 값을 활용한 모형 비교의 차이점과 유용성 등 모형 비교를 위해 사용되는 두 지수의 차이에 관한 경험적 확인을 목적으로 한다.

II. 선행연구 고찰

1. AUDIT-K 척도

개인의 음주와 관련한 음주량, 빈도, 음주로 인한 폐해 등에 관해 평가를 할 수 있는 도구인 AUDIT는 알코올사용장애의 조기 선별뿐 아니라 유해 음주 및 위험 음주 등을 선별할 수 있는 도구이다(이분희 외, 2014; Babor et al., 2001). AUDIT는 총 3개의 하위요인과 10개의 문항으로 구성되어 있는데, 1~3번 문항은 음주의 양과 빈도 등을 평가하며, 4~6번 문항은 알코올 의존과 관련된 증상을 평가하고, 7~10문항은 음주와 관련하여 발생하는 폐해에 대해서 평가하도록 구성되어 있다(Babor et al., 2001). 특히, AUDIT는 초기 개발 당시 6개국에 걸친 일차 의료환경을 고려한 비교 연구를 통해 개발되었으며, 국제 질병 분류(ICD-10)의 기준에 따라 알코올 사용의 스크리닝을 위한 국가 간 표준화 평가도구로서 활용되고 있다(Erford et al., 2021). 더욱이 최근까지 40개 이상의 언어로 번역되어 사용되고 있으며, 청소년기부터 후기 성인기까지의 다양한 연령 범주뿐 아니라 문화적 다양성, 간결성, 높은 정확도 등을 가진 아주 유용한 척도라고 할 수 있다(Erford et al., 2021).

이러한 AUDIT 척도의 요인구조 및 문항에 관한 선행연구를 살펴보면, 기존의 3요인의 10문항 척도(Babor et al., 2001)와 더불어 원척도를 좀 더 간결하게 줄인 AUDIT-C(3문항)와 AUDIT-3(3번 문항 단독)의 사용을 제안한 연구(Bush et al., 1998), AUDIT-PC(5문항)를 제안한 연구(Gomez et al., 2005), AUDIT-4(4문항)를 제안한 연구(Gual et al., 2002), AUDIT-8(8문항)을 제안한 연구(Paulus et al., 2023) 등 그 문항의 수를 달리하여 측정에 활용되고 있을 뿐 아니라, 원척도 10문항의 요인 구조와 관련해서도 여러 연구들에서 1요인 모형, 2요인 모형, 3요인 모형으로 구분하여 타당화를 시도한 결과 1요인(Skogen et al., 2019) 및 2요인(Erford et al., 2021; Peng et al., 2012; Moehring et al., 2018; Thorisdottir et al., 2020; von der Pahlen et al., 2008)과 3요인(Tuliao et al., 2016) 등 각각의 연구에서 서로 다른 모형이 지지되는 등 일관된 결과를 보여주지 못하는 것으로 나타났다.

국내에서는 이병욱 외(2000)의 연구에서 척도의 신뢰도와 타당도를 처음으로 검증한 것으로 보이며, 이후 일부의 연구에서 AUDIT-K에 대한 최적 절단값(조근호 외, 2009), 위험음

주에 대한 하위척도 비교(이분희, 2014) 등의 검증을 통한 척도 연구가 진행된 바 있다. 비록 이병욱 외(2000)의 연구에서 AUDIT-K의 신뢰도 및 타당도를 검증하였다고는 하나, 원척도(Babor et al., 2001)의 9번과 10번 문항에 대한 측정을 원척도의 3점 척도가 아니라 다른 문항과 동일하게 5점 척도로 구성하여 검증을 실시하였기 때문에 원척도와 정확하게 일치하지 않는다. 더욱이 앞서 제시한 국외의 다양한 축약형 척도 개발 및 요인구조 검증 등과 관련된 선행연구들과 같이 AUDIT-K의 요인구조 등에 관한 체계적인 분석이 이루어졌다고 보기에는 어려움이 있다. 이에 본 연구에서는 한국복지패널 자료를 중심으로 선행연구에서 제시하고 있는 AUDIT-K의 요인구조에 대한 분석을 통해 가장 적합한 요인구조 모형을 검증하고자 한다.

2. 척도의 측정불변성

측정불변성(Measurement Invariance, MI)은 측정동등성(Measurement Equivalence) 혹은 측정비교가능성(Measurement Comparability) 등의 용어로 혼용되어 사용되는데, 특정한 척도를 통해 해당 개념을 측정함에 있어 문화, 지리적 지역, 국가, 시점, 언어, 데이터 수집 방법 등의 의미있는 비교 대상에 대해 동일한 속성의 측정값을 보여주는 것과 관련된 것이다(Leitgöb et al., 2022). 즉, 동일한 척도를 활용하여 특정한 집단 간의 횡단 및 종단적으로 측정된 점수를 집단 간 혹은 시점 간 비교를 위해서는 각 집단 및 시점에서 측정된 값의 속성(구조와 의미 등)이 동일하게 나타나고 있음이 전제되어야 함을 의미한다. 구체적으로 집단 간 비교를 위한 측정불변성이 확보되지 않았다는 것은 해당 척도의 의미와 구조가 비교하고자 하는 각 집단에 따라 동일하게 나타나지 않기 때문에 실제적인 차이를 비교한다기 보다는 ‘젓가락과 포크를 비교하게 되는 오류를 범하게 된다(배성우 외, 2020; Chen, 2008; Horn & McArdle, 1992). 또한, 종단적 측정불변성이 확보되지 않았다는 것은 해당 척도의 의미와 구조가 시간의 흐름에 따라 동일하게 나타나지 않을 수 있음을 의미하고(Liu et al., 2017), 이는 결국 해당 척도로 측정된 값과 의미가 시간의 흐름에 따른 진정한 변화(true change)를 반영하는 것인지를 명확하게 파악하기 어렵게 된다는 것을 의미한다(Bertola et al., 2020; Widaman et al., 2010). 이러한 측정불변성을 검증하는 절차는 다양하게 제시되고 있으나 일반적

으로 측정불변성의 각 수준에 따른 제약되는 내용 및 비교 유형, 평가와 관련한 사항을 간략히 정리하면 다음의 <표 1>과 같다.

한편, AUDIT 척도의 측정불변성을 검증하는 연구로 국외에서는 몇몇 연구들이 확인되고 있으나(Erford et al., 2021;

Peng et al., 2012; Moehring et al., 2018; Skogen et al., 2019; Thorisdottir et al., 2020; Tuliao et al., 2016), 국내에서는 현재까지 분석이 거의 이루어지지 못한 것으로 보인다. 국외에서 분석된 측정불변성의 결과를 간략히 정리하면 다음의 <표 2>와 같다. 각각의 연구에서 대상자는 서로 다르지만

표 1. 측정불변성의 수준

수준 (Invariance level)	의미하는 바 (What it implies)	허용되는 그룹 간 비교 유형 (Type of comparison across groups allowed)	측정불변성 수준의 평가 (How the invariance level may be assessed)
형태불변성 (Configural invariance)	그룹 간 동일한 항목과 동일한 구조 설정	기준 모형	다중집단 확인적 요인분석(MGCFEA)의 모형 적합성 기준에 따른 평가
약한(부분) 불변성 ([Full or partial] Metric(Weak) invariance)	각 문항에 대한 그룹 간 동일한 요인부하량 설정 (부분 메트릭불변성을 위해 동일한 요인적재량을 가진 항목이 최소한 2개 이상)	표준화되지 않은 연관성(공분산, 관심 있는 다른 이론적 구성과의 표준화되지 않은 회귀 계수)	형태불변성 모형에 비해 적합도가 크게 떨어지지 않음
강한(부분) 불변성 ([Full or partial] Scalar(Strong) invariance)	각 문항에 대한 그룹 간 동일한 요인부하량 및 절편 설정 (부분 스칼라불변성을 위해 동일한 요인 적재량 및 절편을 가진 항목이 최소한 2개 이상)	표준화되지 않은 연관성 및 잠재평균	약한(부분)불변성 모형에 비해 적합도가 크게 떨어지지 않음
엄격한불변성 (Strict invariance)	각 문항에 대한 그룹 간 동일한 요인부하량, 절편 및 오차분산 설정	표준화되지 않은 연관성 및 잠재평균	강한(부분)불변성 모형에 비해 적합도가 크게 떨어지지 않음

자료: Leitgöb et al.(2022)의 <표 1> 및 참조

표 2. AUDIT 척도의 측정불변성 선행연구

연구	대상	사례 수	통계프로그램 및 추정방식	측정불변성 검증 결과	
Peng et al.(2012)	15개 국가 일반인 조사	27,478	Mplus / WLSMV estimator	Canada 성별(2집단)	Scalar invariance 유지
				Sweden 성별(2집단)	Scalar invariance 유지
				Uganda 성별(2집단)	Scalar invariance 유지
Tuliao et al. (2016)	필리핀 대학생 미국 대학생	255 1,259	Mplus 6 / MLR estimator	인종 (필리핀과 미국 2집단)	Metric invariance 유지
Moehring et al.(2018)	EARLINT network data (6개 개별 연구의 자료로 일반인 및 환자군 등 포함)	28,345	Mplus 7.4 / WLSMV estimator	성별(2집단)	Strict invariance 유지
Skogen et al.(2019)	노르웨이 국민	4,318	R / semTools and lavaan packages	성별(2집단)	Scalar invariance 유지
				연령(65세 기준 2집단)	Scalar invariance 유지
				교육(4집단)	Scalar invariance 유지
Thorisdottir et al.(2020)	캐나다 퇴역 군인	1,669	Amos 22 / ML estimation	PTSD 증상 유무(2집단)	invariance 유지 안 됨
Erford et al.(2021)	대학생	4,756	Mplus 8.3 / WLSMV estimator	인종 (라틴계와 비라틴계 2집단)	Scalar invariance 유지
				학력 (학부와 대학원 2집단)	Configural invariance 유지

성별(Peng et al., 2012; Moehring et al., 2018; Skogen et al., 2019), 연령(Skogen et al., 2019), 인종(Erford et al., 2021; Tulliao et al., 2016) 등 다양한 측면에서의 측정불변성이 유지되고 있음을 보여주었다.

이상과 같이 국외의 경우에는 AUDIT 척도에 대한 측정불변성에 관한 결과들이 척도의 타당도 및 사용에 관한 경험적 정보를 제공해 주고 있다. 하지만, 국내의 경우에는 한국복지패널에서 AUDIT-K를 사용하여 음주 문제를 오랫동안 측정하여 대규모의 자료를 제공하고는 있지만 척도의 요인구조에 관한 타당화에 대한 검증이 적절하게 이루어지지 못한 것으로 보인다. 더욱이 최근 패널 종단 자료를 활용한 집단 간 비교 및 종단 연구들이 다양하게 이루어지고 있는 상황이지만, 척도의 측정불변성에 대한 검증이 체계적으로 이루어져 있지 않은 상황이다. 이러한 측면에서 본 연구에서는 한국복지패널 자료의 AUDIT-K 척도에 대하여 성별집단과 연령대집단을 중심으로 하는 측정불변성 및 단기간 및 장기간에 대한 종단적 측정불변성에 대한 다양한 검증을 통해 척도의 타당성을 경험적으로 확인하고자 한다.

3. 내포 모형 비교를 위한 RMSEA_D 지수의 활용

일반적으로 측정불변성 검증과 같은 내포(nested) 모형에 대한 비교를 위하여 χ^2 의 차이를 활용하는 것이 많이 알려져 있으나, χ^2 차이 검증 방법이 표본의 크기에 민감하여 정확한 차이에 대한 검증이 어려움이 있다는 것이 알려지며, 구조방정식 모형에 대한 모형적합도 지수의 차이값을 중심으로 비교하는 방법이 제시되었다(Chen, 2007). 특히, 잠재변수를 활용하여 모형을 비교할 경우 가장 많이 사용되는 적합도 지수인 RMSEA를 중심으로 두 모형 간 RMSEA의 차이값(이하 Δ RMSEA)을 활용한 모형의 비교가 많이 이루어지고 있다(Ropovik, 2015). 예를 들어, Model A와 Model B에 해당하는 RMSEA값을 각각 RMSEA_A와 RMSEA_B라고 할 때, Δ RMSEA=RMSEA_B-RMSEA_A라고 할 수 있으며, 두 모형의 비교 기준으로 Δ RMSEA>.010(N≤300) 또는 .015(N>300)(표본 크기에 따라 다르게 적용함)이 일반적으로 적용되고 있다(Chen, 2007). 그러나 최근 Savalei et al.,(2023)의 연구에서 내포 모형의 비교에서 활용되는 Δ RMSEA의 값에 대한 몇 가지의 단점을 밝히며 RMSEA_D의 사용을 제안하고 있다. 해당 연구에서 제시한 Δ RMSEA의 단점으로 첫째, 기존 RMSEA 지

수의 컷오프 기준을 적용할 수 없으며, 다만, 별도의 컷오프 기준을 제안한 연구는 있다(Chen, 2007). 둘째, CI값을 구하는 것이 어려우며(Lai, 2020), 작은 표본의 컷오프에 의존하는 연구의 경우 사례 수의 변화를 과도하게 해석할 수 있다. 셋째, 가장 중요한 부분으로 내포 모형을 비교함에 있어 $\Delta\chi^2$ (두 모형 간 χ^2 값의 차이)값과 Δdf (두 모형 간 자유도 차이)가 유사하게 변화될 때 초기 자유도가 크면 Δ RMSEA가 작아지는 등 민감도가 부족한 경우가 있다고 제시하며, 이러한 한계를 어느 정도 극복할 수 있는 지수로서 RMSEA_D 사용을 제안하였다(Savalei et al., 2023, pp.10-15).

내포 모형의 비교를 위해 제안된 RMSEA_D를 구체적으로 살펴보면 다음과 같다. 앞선 예와 같이, Model A의 주요 값을 χ^2_A , df_A 라고 하고 Model B의 주요 값에 대해 χ^2_B , df_B 라고 할 때, RMSEA_D의 계산식은 아래와 같다(Savalei et al., 2023; Rpubs, <https://rpubs.com/JLLJ/SBFNMC> 참조).

$$D = \chi_B^2 - \chi_A^2$$

$$df_D = df_B - df_A$$

$$\text{single-group models} \quad RMSEA_D = \sqrt{\frac{D - df_D}{df_D(N - 1)}}$$

$$\text{multiple-group models} \quad RMSEA_{D, MG} = \sqrt{\frac{D - df_D}{df_D(N - N_g)/N_g}}$$

이렇게 계산된 RMSEA_D는 Model B에서 얻게 되는 추가 자유도에 따른 Model B의 제약조건에 의해 나타나는 추가적인 부적합의 정도로 해석될 수 있으며, 기존에 사용되는 RMSEA의 적합도 기준을 그대로 활용할 수 있다. 이 과정에서 $D < df_D$ 일 경우 RMSEA_D는 0의 값을 가지게 된다(Savalei et al., 2023).

실제 Savalei et al.(2023)의 연구에서는 과거에 분석된 연구의 결과를 토대로 측정불변성의 분석을 위한 모형의 비교를 위하여 Δ RMSEA를 활용하였을 때와 RMSEA_D를 활용할 때 최종적인 결론을 내림에 있어 일부 차이를 보일 수 있음을 제시하고 있다(Savalei et al., 2023, pp.15-19). 이에 본 연구에서는 내포 모형의 비교에 있어 주로 사용되었던 Δ RMSEA와 최근 주목받고 있는 RMSEA_D의 비교를 바탕으로 AUDIT-K 척도의 측정불변성 모형의 선택에 있어서 어떠한 차이가 있는지를 경험적으로 살펴보고자 하였다.

III. 연구 방법

1. 분석 자료

본 연구에서 사용된 자료는 한국복지패널 자료로 2006년 을 시작으로 매년 패널에 대한 조사가 진행되고 있으며, 2023 년 현재 17차 자료까지 공개되어 있다. 본 연구에서는 AUDIT-K 척도를 처음 사용하고 있는 2009년의 4차년도 자 료를 기준으로 척도에 응답한 6,639명의 자료를 활용하였다. 또한, 종단적 측정불변성의 검증을 위하여 4차년도를 기준으 로 1년 단위 5개년도(4, 5, 6, 7, 8차 data)와 4년 단위 5개년 도(4, 7, 10, 13, 16차 data)의 분석을 위하여 각 분석 연도에 해당하는 AUDIT-K 척도에 대해 모두 응답한 대상자 각 2,828명과 1,426명의 자료를 사용하였다. 각 분석에 사용된 연구 대상자의 구체적인 인구사회학적 특성은 <표 3>에 제시 하였다.

2. AUDIT-K 척도

한국복지패널에서 사용 중인 AUDIT-K 척도는 세계보건기 구에서 제공하는 AUDIT(Alcohol Use Disorder Identification Test) 척도(Babor et al., 2001)를 최은진 외(2006)가 사용한 것으로 문제 음주와 알코올 의존에 관한 사항에 관한 10개의 문항으로 구성되어 있다. 척도를 통해 지난 1년간 개인이 경 험한 음주의 빈도와 양, 알코올 의존증세, 음주와 관련된 문제 의 세 영역으로 나누어 측정하고 있는데, 해로운 음주행동 영 역의 3문항(음주의 빈도, 음주량, 고위험 음주의 빈도), 알코 올 의존 영역의 3문항(음주에 대한 통제력 상실, 음주에 대한 증대된 동기, 해장술), 위험한 음주 영역 4문항(음주 후 후회 감, 취중의 일을 기억 못함, 음주 관련 상해, 음주 관련 문제 경험)으로 구성되어 있다(최은진 외, 2006). 각 문항별로 응답 의 내용에서 차이가 있는데, 1번 문항(0점=전혀 마시지 않는 다, 1점=월 1회 이하, 2점=월 2~4회, 3점=주 2~3회, 4점=주 4회이상), 2번 문항(0점=1~2잔 정도, 1점=3~4잔 정도, 2점 =5~6잔 정도, 3점=7~9잔 정도, 4점=10잔 이상), 3~8번 문항 (0점=전혀 없음, 1점=몇 달에 한 번, 2점=한 달에 1~2번, 3점 =주에 1~2번, 4점=거의 매일)은 5점 척도로 측정하고 있으며, 9~10번 문항(0점=전혀 없음, 2점=과거에는 있었지만 지난 한 해 동안 없었음, 4점=지난 1년 동안 있었음)은 3점 척도로

측정하도록 되어 있다. 본 연구에서 사용된 연차별 척도의 내 적일관성 분석 결과, 4차년도의 Cronbach's $\alpha=.892$, 5차 .788, 6차 .798, 7차 .802, 8차 .780, 10차 .778, 13차 .766, 16차 .768로 나타났다.

3. 분석 방법

가. 요인구조 분석

본 연구에서는 초기 AUDIT-K 척도의 요인구조에 대하여 확인적 요인분석을 실시하여, 분석하고자 하는 자료와 가장 잘 부합하는 요인모형을 확인하였다. 선행연구를 통해 알려져 있는 요인구조는 아래와 같이 1요인, 2요인, 3요인으로 구분 할 수 있다(Erford et al., 2021; Moehring et al., 2018; Skogen, et al., 2019; Tuliao et al., 2016; Thorisdottir et al., 2020; von der Pahlen et al., 2008).

- ① 1 요인 모형은 10개 문항이 단일요인으로 구성되어 있 음을 가정한다.
- ② 2 요인 모형은 요인 1(1, 2, 3번 문항)과 요인 2(4, 5, 6, 7, 8, 9, 10번 문항)의 2 요인 구조로 구성되어 있음 을 가정한다.
- ③ 3 요인 모형은 요인 1(1, 2, 3번 문항)과 요인 2(4, 5, 6번 문항), 요인 3(7, 8, 9, 10번 문항)의 3개 요인으로 구성되어 있음을 가정한다.

나. 측정불변성 검증

AUDIT-K 척도의 측정불변성은 4차년도 자료를 기준으로 성별집단 및 연령대집단으로 구분하여 살펴보고자 하였다. 이 는 측정불변성의 검증과 관련한 선행연구에서 집단을 구분하 는 기준으로 가장 많이 활용되고 있는 기준으로서 성별 및 연 령대를 활용하였다. 특히, 연령대 집단의 경우 음주를 할 수 있는 연령을 기준으로 집단을 구분할 때 가장 많이 활용되고 있는 일반적인 기준인 성인과 노인으로 구분하고자 하였으며, 선행연구에서 AUDIT 척도의 측정불변성 검증에서도 활용되 고 있는 65세를 기준으로 성인과 노인 집단을 구분하였다 (Skogen et al., 2019). 구체적으로 성별집단의 경우, 남자집 단 및 여자집단으로 구분하였고, 연령대 집단의 경우 일반적 으로 노인을 구분하는 65세를 기준으로 65세 미만과 65세 이

상으로 집단을 구분하였다.

측정불변성의 분석을 위하여 앞선 요인구조의 분석 결과를 바탕으로 가장 적합하게 나타나고 있는 요인구조 모형에 대하여 기저모형의 분석을 통해 각 집단별 요인구조 모형의 적합성을 검토한 후, 다집단 확인적 요인분석을 통해 형태불변성(configural), 약한불변성(Weak), 강한불변성(Strong), 엄격한 불변성(Strict)의 단계로 각 모형의 분석을 실시하였다(Byrne & Stewart, 2006; Lasker, 2022; Leitgöb et al., 2022; Millsap & Yun-Tein, 2004; Schmitt & Kuljanin, 2008; Svetina et al., 2020; Vandenberg & Lance, 2000).

다. 종단적 측정불변성 검증

종단적 측정불변성은 AUDIT-K 척도가 처음으로 사용된 4차년도를 기준으로 5개년도의 자료에 대한 분석을 실시하였는데, 1년 간격의 5개년도(4, 5, 6, 7, 8차 data)와 4년 간격의 5개년도(4, 7, 10, 13, 16차 data)로 구분하여 단기적 및 장기적 측면에서의 종단적 측정불변성에 대한 검증을 각각 실시하였다.

종단적 측정불변성의 분석의 경우에도 앞선 요인구조 분석 결과 가장 적합하게 나타난 요인구조 모형에 대하여 분석하고자 하는 5개 시점의 하위요인 및 측정 문항을 모두 투입한 연구모형에 대하여 형태불변성(configural), 약한불변성(Weak), 강한불변성(Strong), 엄격한불변성(Strict)의 단계로 각 모형의 분석을 실시하였다(Liu et al., 2017; Mackinnon et al., 2022; Millsap, 2010; Millsap & Cham, 2012; Newsom, 2015; Pitts et al., 1996; Widaman et al., 2010; Zhang & Li, 2022).

라. 자료의 분석 및 모형 비교

본 연구에서는 연구 대상자의 인구사회학적 특성 및 주요 변수의 특성에 관한 분석을 위하여 SPSS 24.0을 사용하였고, 주요 요인구조 모형 및 측정불변성의 검증을 위하여 Mplus 8.4 프로그램의 MLR(Maximum Likelihood Estimation with Robust Standard Errors) 추정 방식을 사용하였으며(Hricovec et al., 2023; Muthén & Muthén, 2019; Tullio et al., 2016), RMSEA_D의 분석 및 CI값의 분석을 위해서 R-project Version 4.3.1과 lavaan packages를 활용하였다.

AUDIT-K 척도의 요인구조에 대한 모형의 비교를 위하여 AIC(Akaike Information Criterion)와 BIC(Bayesian Information Criterion)을 사용하여 비교하였으며, 두 지수의 값이 낮을수록 더 적합하다고 볼 수 있다(Burnham & Anderson, 2004). 다음으로 AUDIT-K 척도의 측정불변성 및 종단적 측정불변성 모형의 비교를 위하여 현재 많이 사용되고 있는 RMSEA(Δ RMSEA) 값의 차이와 더불어 RMSEA_D를 활용하여 각 모형의 차이를 비교함으로써 내포 모형의 비교에 더 적합한 것으로 제시되고 있는 RMSEA_D에 대한 활용성에 대해 확인하고자 하였다. 본 연구의 분석에 사용된 사례 수가 크다는 것으로 반영하여 Δ RMSEA > .015를 기준으로 모형의 적합 여부를 판단하였다(Chen, 2007). 또한, RMSEA_D의 경우 기존 RMSEA 적합도 지수의 기준인 0.05 이하일 때 좋은 적합도(close fit), 0.05~0.08은 괜찮은 적합도(fair fit), 0.08~0.10은 좋지 않은 적합도(mediocre fit), 0.10 이상은 심각하게 좋지 않은 적합도(poor fit)로 판단할 수 있으며, RMSEA 및 RMSEA_D의 CI는 90%의 수준에서 검정하였다(Browne & Cudeck, 1992; Kline, 2023; Savalei et al., 2023).

IV. 연구 결과

1. 인구사회학적 특성

본 연구의 분석에 사용된 연구 대상자의 인구사회학적 특성은 다음의 <표 3>과 같다.

2. AUDIT-K의 확인적 요인분석 결과

선행연구에서 제시된 AUDIT-K 척도의 요인구조 모형을 바탕으로 확인적 요인분석을 실시한 결과를 다음의 <표 4>에 제시하였다. 1요인 모형의 경우 모형적합도가 수용가능한 수준에 있지 않은 것으로 나타났으나, 2요인 모형과 3요인 모형은 모형적합도가 수용가능한 수준에 있는 것으로 나타났다. 그리고 AIC와 BIC의 값 또한 1요인 모형에 비해 2요인 모형과 3요인 모형의 값이 낮은 것으로 나타났다. 따라서 최종 모형으로 2요인 모형과 3요인 모형을 선택할 수 있는데, AIC와 BIC의 값에서 차이가 있기는 하지만 본 연구에서는 선행연구에서도 최종 모형으로 다수 채택된 2요인 모형을 최종 모형으로

표 3. 연구 대상자의 인구사회학적 특성

(N, %)

항목	구분	요인구조 및 측정불변성 사례 N=6,639	종단적 측정불변성 사례	
			4-8차 N=2,828	4-16차 N=1,426
성별	남성	3,930(59.2)	2,052(72.6)	1,093(76.6)
	여성	2,709(40.8)	776(27.4)	333(23.4)
연령	전체	M=46.60세(SD=16.35세) 최소 17세, 최대 98세	M=46.24세(SD=14.66세) 최소 18세, 최대 87세	M=44.16세(SD=12.62세) 최소 18세, 최대 84세
	65세 미만	5,417(81.6)	2,401(84.9)	1,304(91.4)
	65세 이상	1,222(18.4)	427(15.1)	122(8.6)
결혼 상태	유배우	4,448(67.0)	2,051(72.5)	1,071(75.1)
	사별	449(6.8)	110(3.9)	30(2.1)
	이혼	281(4.2)	144(5.1)	69(4.8)
	별거	60(0.9)	32(1.1)	15(1.1)
	미혼	1,394(21.0)	489(17.3)	241(16.9)
	기타	5(0.1)	2(0.1)	-
학력	초등학교 졸업 이하	1,268(19.1)	421(14.9)	155(10.9)
	중학교 졸업	669(10.1)	312(11.0)	144(10.1)
	고등학교 졸업	2,305(34.7)	1,074(38.0)	576(40.4)
	대학교 졸업 이상	2,397(36.1)	1,021(36.1)	551(38.6)

표 4. AUDIT-K 척도의 요인모형에 대한 모형적합도

척도	모형	χ^2	df	CFI	TLI	SRMR	RMSEA (90% CI)	AIC	BIC
AUDIT-K (N=6,639)	1 factor	2504.467	35	.787	.726	.067	.103[.100, .107]	136015.765	136219.787
	2 factor	430.658	34	.966	.955	.033	.042[.038, .046]	131631.147	131841.969
	3 factor	389.703	32	.969	.957	.031	.041[.037, .045]	131540.337	131764.761

로 선택하였다(Erford et al., 2021; Moehring et al., 2018; Thorisdottir et al., 2020; von der Pahlen et al., 2008).²⁾

3. 기저모형 검토

측정불변성에 관한 검증을 위하여 기저모형의 적합도에 관한 분석을 실시한 결과가 다음의 <표 5>에 제시되었다. 기저 모형은 앞서 가장 적합한 요인모형으로 선택된 2요인 모형을 기준으로 하였다. 성별을 기준으로 남성과 여성 집단 모두와 연령대를 기준으로 65세 미만과 65세 이상 집단에 대하여 각

각 확인적 요인분석한 결과, 모든 집단에서 요인구조모형의 적합도가 수용가능한 수준으로 나타나 측정불변성 분석이 가능함을 확인하였다.

다음으로 종단적 측정불변성 검증을 위하여 기저모형에 대한 분석을 실시하였는데, 구체적으로 단기적 종단측정불변성을 검증하는 1년 간격의 5개년 자료(4, 5, 6, 7, 8차)와 4년 간격의 5개년 자료(4, 7, 10, 13, 16차)에 대하여 2요인 모형에 대한 확인적 요인분석을 실시하였다. <표 6>에 제시된 바와 같이, 각 시점에 해당하는 요인구조 모형에 대한 모형적합도 지수는 모두 수용가능한 수준으로 나타나 종단적 측정불변성

2) 상대적으로 더 복잡한 3요인 모형의 전체 적합도(df=32)가 이보다 간명한 2요인 모형(df=34)의 적합도보다 더 높은 것은 예상 가능한 일이며, 복잡한 모형이 갖는 적합도 상의 이점이 AIC의 산출 과정에서 자유 모수가 많음으로 인해 받는 불이익을 상쇄시킬 만큼 크거나, BIC에서 사례 수를 가중치로 한 불이익을 상쇄시킬 만큼 크다고 볼 수 없다고 판단하였기 때문에 AIC와 BIC값이 약간 높음에도 불구하고 2요인 모형을 선택하였다(Kline, 2023). 이는 AIC 및 BIC에 기초하여 모형을 선택하는 경우 표집 오차가 많이 개입하게 되는데, BIC의 경우 사례 수가 커짐에 따라 표집오차가 증가하기 때문에 다양한 사례 수에 대해 모형의 적합성이 다르게 나타날 수 있어 주의를 요한다는 점을 고려하였다(Preacher & Merkle, 2012).

표 5. AUDIT-K 척도의 횡단적 측정불변성의 기저모형 검토

구분	Model fit index							
	χ^2	df	CFI	TLI	SRMR	RMSEA (90% CI)		
AUDIT-K (N=6,639)	성별	남성 (n=3,930)	424.855	34	.949	.932	.041	.054[.050, .059]
		여성 (n=2,709)	93.342	34	.976	.968	.028	.025[.019, .032]
	연령대	65세 미만 (n=5,417)	370.431	34	.967	.957	.033	.043[.039, .047]
		65세 이상 (n=1,222)	58.184	.34	.984	.979	.030	.024[.013, .034]

표 6. AUDIT-K 척도의 종단적 측정불변성에 대한 기저모형 검토

구분	Model fit index							
	χ^2	df	CFI	TLI	SRMR	RMSEA (90% CI)		
AUDIT-K	4~8차 (5 wave) N=2,828	4차	310.727	34	.948	.931	.041	.054[.048, .059]
		5차	285.039	34	.939	.919	.044	.051[.046, .057]
		6차	264.023	34	.952	.937	.041	.049[.044, .054]
		7차	270.161	34	.947	.930	.042	.050[.044, .055]
		8차	249.789	34	.948	.931	.045	.047[.042, .053]
	4~16차 (5 wave) N=1,426	4차	219.710	34	.938	.918	.045	.062[.054, .070]
		7차	155.509	34	.945	.927	.044	.050[.042, .058]
		10차	174.621	34	.941	.922	.046	.054[.046, .062]
		13차	113.197	34	.962	.949	.040	.040[.032, .049]
		16차	119.175	34	.956	.942	.048	.042[.034, .050]

에 대한 분석이 가능함을 확인하였다.

4. 측정불변성 검증

성별(남성, 여성)과 연령대(65세 미만, 65세 이상) 집단을 기준으로 측정불변성 검증을 실시한 결과가 다음의 <표 7>에 제시되어 있다. 구체적으로 성별 집단의 측정불변성의 분석결

과, $\Delta RMSEA > .015$ 를 기준으로 할 경우, 형태불변성과 약한 불변성 모형 간 $\Delta RMSEA = .000$ 이고, 약한불변성과 강한불변성 모형 간 $\Delta RMSEA = .026$ 으로 나타나 약한불변성(Weak)까지 유지되는 것으로 나타났다. 그리고 RMSEA_D와 CI값을 기준으로 할 경우 약한불변성과 강한불변성 모형 간 RMSEA_D = .155와 CI의 상한 값이 꽤 많은 적합도(fair fit)의 수준인 .08을 넘어서는 것으로 나타나 좋은 적합도를 보인다고

표 7. AUDIT-K 척도의 횡단적 측정불변성 검증 및 모형 비교(N=6,639)

Invariance model	Model fit index					Model comparison					
	χ^2	df	CFI	SRMR	RMSEA	T _D	df _D	$\Delta RMSEA$	RMSEA _D	90% CI for RMSEA _D	
성별	Configural	400.878	68	.960	.036	.038	-	-	-	-	
	Weak	440.328	76	.956	.047	.038	39.45	8	0.000	0.034	[.026, .042]
	Strong	1257.791	86	.859	.148	.064	817.463	10	0.026	0.155	[.148, .162]
	Strict	2986.306	96	.652	.291	.095	1728.515	10	0.031	0.227	[.220, .234]
연령대	Configural	378.103	68	.970	.033	.037	-	-	-	-	
	Weak	392.875	76	.970	.039	.035	14.772	8	-0.002	0.016	[.006, .030]
	Strong	736.860	86	.938	.060	.048	343.985	10	0.013	0.100	[.093, .107]
	Strict	890.899	96	.924	.070	.050	154.039	10	0.002	0.066	[.059, .073]

볼 수 없기 때문에 약한불변성(Weak)까지 유지되는 것으로 해석할 수 있어 두 결과가 동일한 것으로 나타났다.

다음으로 연령대 집단의 경우, 형태불변성과 약한불변성 모형 간 $\Delta RMSEA = -.002$ 이고, 약한불변성과 강한불변성 모형 간 $\Delta RMSEA = .013$ 으로 나타났고, 강한불변성과 엄격한불변성 모형 간 $\Delta RMSEA = .002$ 로 나타나 엄격한불변성(Strict)까지 유지되는 것으로 나타났으나, $RMSEA_D$ 의 경우 약한불변성과 강한불변성 모형 간 $RMSEA_D$ 가 .100이었고 CI의 상한 값이 .08을 넘어서고 있어서 약한불변성(Weak invariance) 모형까지 유지되는 것으로 나타나 $\Delta RMSEA$ 를 중심으로 한 비교와 차이를 보이는 것으로 확인되었다. 본 연구에서는 $\Delta RMSEA$ 값을 기준으로 비교를 하기보다는 기존의 한계를 어느 정도 극복하고 있는 $RMSEA_D$ 를 중심으로 결과를 해석하고자 하며, 이 기준에 따르면, 최종적으로 성별집단 및 연령대집단의 측정불변성의 경우 모두 약한불변성(Weak) 모형까지 유지되는 것으로 볼 수 있다.

5. 종단적 측정불변성 검증

종단적 측정불변성을 검증하기 위하여 각 모형에 대한 분석을 위해서 기준이 되는 아무런 제약이 없이 형태가 같다고 보는 형태불변성(Configural invariance) 모형을 시작으로, 요인부하량에 대한 제약을 설정하는 약한불변성(Weak invariance), 절편에 대한 제약을 설정하는 강한불변성(Strong invariance), 오차분산에 대한 제약을 설정하는 엄격한불변성(Strict invariance) 모형 순서로 각각의 제약을 설정하는 모형에 대한 분석을 순차적으로 실시하였다. 본 연구에서는 단기

적 측면의 종단적 측정불변성 검증을 위하여 4차~8차까지 1년 간격의 5개 wave(총 5년)를 중심으로 한 모형(LMI-model 1)과 장기적 측면의 종단적 측정불변성 검증을 위하여 4차~16차까지 4년 간격의 5개 wave(총 13년)를 중심으로 한 모형(LMI-model 2)으로 구분하여 분석을 실시하였다. 이러한 종단적 측정불변성의 분석에 사용된 자료는 AUDIT-K 질문에 대해 분석 대상이 되는 5개 wave 모두 응답한 대상으로 LMI-model 1에서 2,828 case가 사용되었고, LMI-model 2에서 1,426 case가 사용되었으며, 구체적인 대상자 각각의 인구사회학적 특성은 <표 4>에서 확인할 수 있다.

주요 분석 결과는 다음의 <표 8>과 같다. 먼저 4차년도에서 8차년도까지 1년 간격의 자료에 대한 종단적 측정불변성 LMI-model 1을 검증한 결과, $\Delta RMSEA$ 이 각각 .000, .001, .002으로 $\Delta RMSEA > .015$ 를 기준을 넘어서지 않아 엄격한불변성(Strict) 모형까지 유지되는 것으로 나타났고, $RMSEA_D$ 와 CI값을 기준으로 할 경우 $RMSEA_D$ 가 .044, .052, .066이고 CI의 상한값이 괜찮은 적합도(fair fit)의 수준인 .08을 넘어서지 않은 것으로 나타나 좋은 적합도를 보인다고 볼 수 있어 엄격한불변성(Strict) 모형까지 유지되는 동일한 결과를 보이는 것으로 나타났다.

다음으로 4차년도에서 16차년도까지 4년 간격의 자료에 대한 종단적 측정불변성 LMI-model 2를 검증한 결과, $\Delta RMSEA$ 이 각각 .001, .001, .012로 엄격한불변성(Strict) 모형까지 유지되는 것으로 나타났으나, $RMSEA_D$ 와 CI값을 기준으로 할 경우 강한불변성과 엄격한불변성 모형 간 $RMSEA_D = .152$ 이고 CI의 상한 값이 .08을 넘어서고 있어 강한불변성(Strong) 모형까지 유지되는 것으로 나타나 $\Delta RMSEA$ 를 중심

표 8. AUDIT-K 척도의 종단적 측정불변성 검증 및 모형 비교

Invariance model		Model fit index					Model comparison				
		χ^2	df	CFI	SRMR	RMSEA	T_D	df_D	$\Delta RMSEA$	$RMSEA_D$	90% CI for $RMSEA_D$
LMI-model 1 4-8차 (5 wave) N=2,828	Configural	2646.585	1030	.956	.056	.024	-	-	-	-	-
	Weak	2743.505	1062	.955	.059	.024	96.92	32	0.000	0.027	[.022, .031]
	Strong	3016.922	1094	.948	.060	.025	273.417	32	0.001	0.052	[.047, .056]
	Strict	3556.767	1134	.935	.064	.027	539.845	40	0.002	0.066	[.063, .070]
LMI-model 2 4-16차 (5 wave) N=1,426	Configural	2014.417	1030	.948	.057	.026	-	-	-	-	-
	Weak	2136.012	1062	.943	.061	.027	121.595	32	0.001	0.044	[.038, .051]
	Strong	2340.884	1094	.934	.062	.028	204.872	32	0.001	0.062	[.055, .068]
	Strict	3701.087	1134	.865	.078	.040	1360.203	40	0.012	0.152	[.147, .158]

으로 한 비교와 차이가 있는 것으로 확인되었다.

이상의 결과를 종합하면, AUDIT-K척도의 경우 단기적으로는 5년부터 장기적으로 13년에 걸친 종단적 측정불변성이 엄격한불변성 및 강한불변성 모형까지 유지되는 것으로 나타났기 때문에 본 척도를 사용하여 단기 혹은 장기적인 반복측정에 사용 가능할뿐 아니라 측정 시점별 비교가 가능한 척도임을 경험적으로 확인하였다. 다만, 모형의 비교에 있어 Δ RMSEA와 RMSEA_D를 비교한 결과에서 일부 차이가 있는 것이 확인되어, 내포 모형(nested model)의 비교에 있어 비교의 기준 적용에 대한 좀 더 세밀한 주의가 필요하다고 하겠다.

V. 논의 및 결론

본 연구에서는 한국복지패널 자료를 활용하여 AUDIT-K 척도의 요인구조, 성별 및 연령대별 집단에 따른 측정불변성 및 종단적 측정불변성의 검증을 통해 척도의 타당성과 활용성을 검토하고자 하였으며, 이 과정에서 모형의 비교를 위하여 최근 주목받고 있는 RMSEA_D를 활용하여 기존의 모형 비교 방식과의 어떠한 차이가 있는지를 검증하였다.

주요 연구 결과와 함의를 정리하면 다음과 같다. 첫째, 한국복지패널 4차년도 자료를 중심으로 살펴본 AUDIT-K 척도의 요인 구조는 원척도에서 제시하는 3요인의 구조와 2요인 구조 모두 적합한 것으로 나타나 음주 행동과 위험한 음주 등 음주 문제에 관한 분석에서 두 요인 모형 모두 활용될 수 있음을 확인하였다. 그러나 본 연구에서는 앞서 제시한 바와 같이 3요인 모형(Babor et al., 2001)보다는 조금 더 간명하고 다수의 연구에서 더 적합한 것으로 제시된 2요인 모형(Erford et al., 2021; Peng et al., 2012; Moehring et al., 2018; Thorisdottir et al., 2020; von der Pahlen et al., 2008)을 최종 모형으로 선택하여 이후 측정불변성에 대한 분석을 실시하였다.

둘째, AUDIT-K 척도의 성별과 연령대 집단 간 측정불변성 검증 결과, 성별 및 연령대 집단 모두 약한불변성(Weak) 모형까지 유지되는 것으로 나타났다. 이러한 결과는 성별집단 및 연령집단에 대한 강한불변성 모형까지 유지되는 것으로 나타난 선행연구의 결과와는 일부 차이가 있었다(Peng et al., 2012; Moehring et al., 2018; Skogen et al., 2019). 일반적으로 잠재변수 사이의 공분산과 표준화되지 않은 회귀계수를 비

교할 수 있기 위해서는 약한불변성 모형까지는 유지되어야 하기 때문에 AUDIT-K 척도의 경우 측정불변성이 유지된다고 할 수 있다. 다만, 잠재 평균을 비교하기 위해서는 강한불변성(Strong)까지 유지되어 한다는 것이 일반적인 기준(Leitgöb et al., 2022)이기 때문에 AUDIT-K 척도를 활용하여 성별과 65세를 기준으로 하는 연령대 집단 간 잠재 평균을 비교하기에는 일정한 한계가 있다고 할 수 있다.

셋째, 종단적 측정불변성 검증 결과, 1년 간격의 자료에 대한 종단적 측정불변성 LMI-model 1에서는 엄격한불변성(Strict) 모형까지 유지되는 것으로 나타났으며, 4년 간격의 자료에 대한 종단적 측정불변성 LMI-model 2는 강한불변성(Strong) 모형까지 유지되는 것으로 나타났다. 현재까지 종단적 측정불변성에 관한 선행연구가 이루어지지 않아 직접적인 비교를 할 수는 없지만, AUDIT-K가 단기 혹은 장기적인 반복 측정에 사용 가능한 척도임을 경험적으로 확인하였다. 이러한 결과는 최근 패널 자료의 확대에 의한 종단 연구의 증가를 고려할 때, AUDIT-K를 통해 측정된 값의 의미와 구조가 시간이 흐름에 따라 변하지 않음을 보여주기 때문에 종단 연구에서 충분히 활용될 수 있음을 시사한다(Liu et al., 2017).

넷째, 내포 모형 비교를 위한 Δ RMSEA와 RMSEA_D의 차이를 중심으로 살펴봤을 때, 성별집단은 동일하게 나타났으나 연령대 집단의 경우 결과에서 차이가 있었다. 즉, 연령대 집단의 측정불변성 분석에서 Δ RMSEA를 통해 비교할 경우 엄격한 불변성까지 유지되었으나 RMSEA_D를 적용할 경우 약한불변성 모형까지 유지되는 것으로 나타났다. 또한, LMI-model 2의 검증 결과에서도 Δ RMSEA를 통해 비교할 경우 엄격한 불변성까지 유지되었으나 RMSEA_D를 적용할 경우 강한불변성 모형까지 유지되는 것으로 나타나 결과에 있어 차이가 있었다. 이러한 분석 결과는 어떠한 모형 비교의 기준을 활용하느냐에 따라 최종적인 결과가 다르게 해석될 수 있음을 경험적으로 보여준 것이라고 할 수 있다. 즉, 앞서 제시한 바와 같이 초기 자유도가 클 경우 Δ RMSEA가 작아지는 등의 한계가 있을 수 있음을 인정한다면, 이러한 한계를 일부 보완해줄 수 있는 RMSEA_D를 활용한 모형의 비교 결과를 받아들이는 것이 합리적이라고 할 수 있겠다(Brace, 2020; Savalei et al., 2023). 다만, 내포 모형의 비교에 있어 RMSEA_D가 Δ RMSEA보다 한계가 적고 유용하게 활용될 수 있지만, 분명한 것은 RMSEA_D 또한 한계가 없는 것은 아니며, 기존 RMSEA 지수가 갖는 한계를 그대로 계승하고 있는 것들이 있기 때문에 완전

하다고 볼 수는 없다(Savalei et al., 2023, pp.25-28). 따라서 모형의 적합성을 비교하는 많은 지수들이 처음 만들어진 이후 여러 연구자들에 의해 지속적으로 개선이 이루어져 오고 있기 때문에 새롭게 확장되고 정의되는 적합도 지수를 적절히 활용하여 구조방정식 모형의 비교를 실시하는 것이 바람직하다고 할 수 있겠다.

정리하면, 본 연구를 통해 AUDIT-K 척도의 요인구조는 선행연구에서 다루고 있는 2요인 및 3요인구조가 모두 자료와 부합하고 있었으며, 성별 및 연령대별 측정불변성에 대해서는 약한불변성 모형까지 나타나고 있기 때문에 집단 간 비교는 가능하지만 집단 간 잠재평균을 비교함에 있어서는 주의가 필요하다 할 수 있다. 종단적 측정불변성의 경우에는 강한측정불변성 이상이 유지되었기 때문에 종단적 측정 및 비교 분석에 충분히 사용될 수 있다고 할 수 있다. 그리고 향후 구조방정식의 내포모형에 대한 비교를 실시할 경우, 연구자들은 기존에 사용되고 있는 Δ RMSEA와 새롭게 제시된 RMSEA_D 활용의 유용성에 대한 이해를 바탕으로 가장 적합한 지수가 활용될 수 있도록 할 필요가 있다.

마지막으로 본 연구의 한계는 다음과 같다. 첫째, 본 연구의 분석에 활용된 한국복지패널 자료가 대표성을 가지는 우수한 패널 자료라고 하더라도 모든 인구집단에 일반화하기에는 일부 제약이 있을 수 있다. 둘째, 본 연구에서는 장기적인 측면에서의 측정불변성을 모형의 복잡성을 고려하여 4년 단위로 측정된 값을 중심으로 분석하였으나, 향후 분석에서는 측정시점을 모두 고려한 종단적 측정불변성에 대한 검증이 이루어질 때 더 정확한 결과를 얻을 수 있을 것으로 기대된다. 셋째,

본 연구의 종단적 측정불변성을 위하여 분석된 사례 중 시간이 지남에 따라 탈락하게 되는 사례들이 있었는데, 본 연구의 척도가 음주 문제를 다룬다는 특성을 고려한다면 탈락한 사례들의 음주 특성 및 관련 요인들에 관한 추가적인 분석이 필요하다고 할 수 있다. 넷째, 측정불변성을 검증하기 위한 집단 구분 기준으로 선행연구에서 다루어지고 있는 인종, 학력 등과 같은 기준뿐 아니라 다양한 생애주기를 기준으로 연령 집단을 구분하여 측정불변성에 대한 추가적 검증이 필요하다. 특히, 본 연구에서는 65세를 기준으로 연령대를 구분하고 있지만, 음주 문제의 특성을 고려한다면 추후 연구에서는 65세보다는 낮은 연령을 기준으로 연령대 집단을 구분하여 분석을 실시할 필요도 있다. 다섯째, 본 연구에서는 내포모형을 비교하기 위한 지수로서 많이 사용되는 CFI 지수를 확인하지는 않았는데, 그 이유는 CFI 또한 RMSEA와 유사한 한계가 있다는 점을 고려하였다(Savalei et al., 2023, pp.29-31). 추후 CFI의 정확한 활용에 관한 경험적 연구들이 확대될 경우 구조방정식 모형의 비교에 있어 추가적으로 고려해 볼 필요도 있다.

박병선은 경북대학교에서 정신건강사회복지 전공으로 박사학위를 받았으며, 강릉원주대학교에서 교수로 재직 중이다. 주요 관심 분야는 정신건강 문제, 척도의 타당화이며, 현재 척도의 측정불변성 등을 연구하고 있다.
(E-mail: bspark@gwnu.ac.kr)

이선영은 일본 도시샤대학교에서 사회복지정책 전공으로 박사학위를 받았으며, 강릉원주대학교 부교수로 재직 중이다. 주요 관심 분야는 돌봄, 정신건강 문제이며, 현재 척도의 측정불변성 등을 연구하고 있다.
(E-mail: sylee@gwnu.ac.kr)

참고문헌

- 국립정신건강센터. (2021). 2021년 정신건강실태 조사 보고서. 보건복지부 국립정신건강센터.
- 배성우, 김이영, 도문학, 김혜지, 박병선. (2020). 축약형 우울척도 (CES-D-10) 의 요인구조 및 측정불변성 검증: Andersen form과 Boston form을 중심으로. *정신건강과 사회복지*, 48(1), pp.33-55.
- 이병욱, 이충현, 이필구, 최문종, 남궁기. (2000). 한국어판 알코올 사용장애 진단 검사(AUDIT: Alcohol Use Disorders Identification Test)의 개발. 신뢰도 및 타당도 검사. *중독정신의학*, 4(2), pp.83-92.
- 이분희, 이소희, 이준석, 황세희, 지홍, 장경호 등. (2014). 남성 일반인에서 위험음주에 따른 한국어판 Alcohol Use Disorders Identification Test(AUDIT-K)의 하위척도 비교. *중독정신의학*, 18(2), pp.80-86.
- 조근호, 채숙희, 박애란, 이해국, 신임희, 민성호. (2009). 위험 음주자의 선별을 위한 한국어판 Alcohol Use Disorders Identification Test(AUDIT-K)의 최적 절단값. *중독정신의학*, 13(1), pp.34-40.
- 최은진, 강은정, 김나연, 김경남, 박미형, 이난희. (2006). 국민건강영양조사 제3기(2005): 성인보건교육의 실태. 보건복지부 한국보건사회연구원.
- Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., Monteiro, M. G., & World Health Organization. (2001). *AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care*(No. WHO/MSD/MSB/01.6 a). World Health Organization.
- Bertola, L., Bensenor, I. M., Gross, A. L., Caramelli, P., Barreto, S. M., Moreno, A. B. et al. (2020). Longitudinal measurement invariance of neuropsychological tests in a diverse sample from the ELSA-Brasil study. *Brazilian Journal of Psychiatry*, 43, pp.254-261. <http://dx.doi.org/10.1590/1516-4446-2020-0978>.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, 33(2), pp.261-304. doi: 10.1177/0049124104268644.
- Brace, J. C. (2020). *Relaxed methods for evaluating measurement invariance within a multiple-group confirmatory factor analytic framework*. PhD Thesis. University of British Columbia.
- Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods and Research*, 21(2), pp.230-258. <https://doi.org/10.1177/0049124192021002005>.
- Byrne, B. M., & Stewart, S. M. (2006). The macs approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(2), pp.287-321.
- Bush, K., Kivlahan, D. R., McDonnell, M. B., Fihn, S. D., & Bradley, K. A. (1998). The AUDIT alcohol consumption questions(AUDIT-C). An effective brief screening test for problem drinking. *Archives of Internal Medicine*, 158(16), pp.1789-1795. <https://doi.org/10.1001/archinte.158.16.1789>.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), pp.464-504. doi: 10.1080/10705510701301834.
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparison in cross-cultural research. *Journal of Personality and Social Psychology*, 95(5), pp.1005-1018.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), pp.233-255. doi: 10.1207/S15328007SEM0902_5.
- Erford, B. T., Sriken, J., Sherman, M. F., Hibbs, J. S., Smith, H. L., Kipper-Smith, A., & Niarhos, F. (2021). Psychometric analysis, internal structure, and measurement invariance of the Alcohol Use Disorders Identification Test (AUDIT) scores from a large university sample. *Measurement and Evaluation in Counseling and Development*, 54(3), pp.188-205.
- Gomez, A., Conde, A., Santana, J. M., & Jorran, A. (2005). Diagnostic usefulness of brief versions of Alcohol Use Disorders Identification Test (AUDIT) for detecting hazardous drinkers in primary care settings. *Journal of Studies on Alcohol*, 66, pp.305-308. <https://doi.org/10.15288/jsa.2005.66.305>.
- Gual, A., Segura, L., Contel, M., Heather, N., & Colom, J. (2002). AUDIT-3 and AUDIT-4: Effectiveness of two short forms

- of the Alcohol Use Disorders Identification Test. *Alcohol and Alcoholism*, 37(6), pp.591-596. <https://doi.org/10.1093/alcalc/37.6.591>.
- Hricovec, M. M., Su, C. C., Bart, T. A., Marsh, K. F., Alsup, C. K., & Cicero, D. C. (2023). Measurement invariance of the Personality Inventory for the DSM-5 across U.S. East Asian, Southeast Asian, and White participants. *Personality Disorders: Theory, Research, and Treatment. Advance online publication*, pp.1-10. <https://doi.org/10.1037/per0000647>.
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, 18(3-4), pp.117-144.
- RPubs. A Better Version of RMSEA for Nested Model Comparisons. <https://rpubs.com/JLLJ/SBFNMC>에서 2023. 6. 30. 인출
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.
- Lai, K. (2020). Confidence interval for RMSEA or CFI difference between nonnested models. *Structural Equation Modeling*, 27(1), pp.16-32. doi: 10.1080/10705511.2019.1631704.
- Lasker, J. (2022). What is Not Tested is Not Debunked: A Comment about Measurement Invariance. *PsyArXiv Preprints*. doi: 10.31234/osf.io/r7e6f.
- Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K. et al. (2022). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research*, 102805.
- Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological methods*, 22(3), pp.486-506.
- Mackinnon, S., Curtis, R., & O'Connor, R. (2022). A tutorial in longitudinal measurement invariance and cross-lagged panel models using lavaan. *Meta-Psychology*, 6, pp.1-20.
- Millsap, R. E. (2010). Testing measurement invariance using item response theory in longitudinal data: An introduction. *Child Development Perspectives*, 4(1), pp.5-9.
- Millsap, R. E., & Cham, H. (2012). Investigating factorial invariance in longitudinal data(109-127). In Laursen, B. P., Little, T. D., & Card, N. A. (Eds), *Handbook of developmental research methods*. New York: Guilford.
- Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research*, 39(3), pp.479-515.
- Moehring, A., Krause, K., Guertler, D., Bischof, G., Hapke, U., Freyer-Adam, J. et al. (2018). Measurement invariance of the alcohol use disorders identification test: Establishing its factor structure in different settings and across gender. *Drug and alcohol dependence*, 189, pp.55-61.
- Muthén, L. K., & Muthén, B. O. (2019). *Mplus. Version 8.4 [Computer program]*. Los Angeles, CA: Muthén and Muthén.
- Newsom, J. T. (2015). *Longitudinal Structural Equation Modeling: A Comprehensive Introduction*. New York: Routledge.
- Orjiakor, C. T., Sellbom, M., Keeley, J. W., & Bagby, R. M. (2023). Measurement invariance of the Personality Inventory for the DSM-5 (PID-5) for Nigerian and White American university students. *Psychological Assessment*, 35(8), pp.715-720. <https://doi.org/10.1037/pas0001251>.
- Paulus, D. J., Rogers, A. H., Capron, D. W., & Zvolensky, M. J. (2023). Maximizing the use of the Alcohol Use Disorders Identification Test (AUDIT) as a two-step screening tool. *Addictive behaviors*, 137, p.107521.
- Peng, C. Z., Wilsnack, R. W., Kristjanson, A. F., Benson, P., & Wilsnack, S. C. (2012). Gender differences in the factor structure of the Alcohol Use Disorders Identification Test in multinational general population surveys. *Drug and alcohol dependence*, 124(1-2), pp.50-56.
- Pitts, S. C., West, S. G., & Tein, J. (1996). Longitudinal measurement models in evaluation research examining stability and change. *Evaluation and Program Planning*, 19(4), pp.333-350.
- Preacher, K. J., & Merkle, E. C. (2012). The problem of model selection uncertainty in structural equation modeling. *Psychological methods*, 17(1), pp.1-14.
- Ropovik, I. (2015). A cautionary note on testing latent variable models. *Frontiers in psychology*, 6, 1715.
- Savalei, V., Brace, J. C., & Fouladi, R. T. (2023). We need to change how we compute RMSEA for nested model comparisons in structural equation modeling, *Psychological Methods. Advance online publication*. <https://doi.org/10.>

1037/met0000537.

- Schmitt, N., & Kuljanin, G., (2008). Measurement invariance: Review of practice and implications. *Human Resource Management Review, 18*(4), pp.210-222.
- Skogen, J. C., Thørrisen, M. M., Olsen, E., Hesse, M., & Aas, R. W. (2019). Evidence for essential unidimensionality of AUDIT and measurement invariance across gender, age and education. Results from the WIRUS study. *Drug and alcohol dependence, 202*, pp.87-92.
- Svetina, D., Rutkowski, L., & Rutkowski, D. (2020). Multiple-group invariance with categorical outcomes using updated guidelines: an illustration using M plus and the lavaan/semtools packages. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(1), pp.111-130.
- Thorisdottir, A. S., Mason, J. E., Vig, K., & Asmundson, G. J. (2020). Factor structure and measurement invariance of the Alcohol Use Disorders Identification Test (AUDIT) in a sample of military veterans with and without PTSD. *Substance use & misuse, 55*(8), pp.1370-1377.
- Tuliao, A. P., Landoy, B. V. N., & McChargue, D. E. (2016). Factor structure and invariance test of the alcohol use disorder identification test (AUDIT): comparison and further validation in a US and Philippines college student sample. *Journal of ethnicity in substance abuse, 15*(2), pp.127-143.
- Vandenberg, R. J. & Lance, C. E. (2000), A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), pp.4-70.
- von der Pahlen, B., Santtila, P., Witting, K., Varjonen, M., Jern, P., Johansson, A., & Sandnabba, N. K. (2008). Factor structure of the Alcohol Use Disorders Identification Test (AUDIT) for men and women in different age groups. *Journal of Studies on Alcohol and Drugs, 69*, pp.616-621.
- Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within longitudinal structural equation models: Measuring the same construct across time. *Child Development Perspectives, 4*(1), pp.10-18.
- Zhang, L., & Li, Z. (2022). Factor structure and longitudinal measurement invariance of the K6 among a national representative elder sample of China. *BMC Public Health, 22*(1), p.1789.

Validating the Factor Structure and Measurement Invariance of the AUDIT-K Scale:

Focusing on the Use of the RMSEA_D Index for Nested Model Comparison

Park, Byung-Sun¹ | Lee, Sun-Young¹

¹ Gangneung-Wonju National University

Abstract

This study was conducted to verify the validity and utility of the AUDIT-K by examining its factor structure and measurement invariances across gender and age groups, and longitudinal measurement invariance using data from the Korean Welfare Panel. In this process, RMSEA_D, which has recently attracted attention, was used to verify any differences with the existing nested model comparison method. The data used in the analysis ranged from the 4th year to the 16th year. The main results of this study are as follows. First, AUDIT-K was found to be suitable for both two-factor and three-factor structures. Second, measurement invariance across gender and age groups was found to be maintained up to the weak invariance model. Third, longitudinal measurement invariance over 5 years at 1-year intervals was maintained up to the strict invariance model, and longitudinal measurement invariance over 13 years at 4-year intervals was found to be maintained up to the strong invariance model. Fourth, when comparing Δ RMSEA and RMSEA_D for nested model comparison, it was confirmed that there were differences in results when comparing some models, suggesting that caution is needed in applying nested model comparison standards. Finally, the implications and limitations of this study were presented.

Keywords: AUDIT-K, Drinking Problem, Korean Welfare Panel, Factor Structure, Measurement Invariance, RMSEA_D