미래사회
Journal of Future Society
2026. Vol. 17, No. 3, 23～38.
https://doi.org/10.22987/jifso.2026.17.3.23

스마트폰 과의존 척도의 측정구조와 점수 해석 범위:
2022년 스마트폰 과의존 실태조사 자기보고 자료에 대한 후속 측정학적 증거 검토

진 수 붕*

국민대학교

본 연구는 2016년 개편 이후 국가조사에서 지속적으로 활용되어 온 스마트폰 과의존 자기보고 10문항 척도의 측정구조와 점수 해석 범위를 후속 검토하였다. 2022년 「스마트폰 과의존 실태조사」 비식별 마이크로데이터 중 자기보고 10문항 완전응답자 23,097명을 분석하고, 4점 서열형 문항에 평균 및 분산 조정 가중최소제곱법(weighted least squares with mean and variance adjustment, WLSMV)을 적용하여 단일요인, 2요인, 조절실패·현저성·문제적 결과의 상관 3요인 모형을 비교하였다. 문항별 수정 문항–총점 상관은 .635～.715, 표준화 부하량은 .770～.887이었다. 3요인 모형이 상대적으로 가장 충분한 자료지지를 받았으나 robust 비교적합지수(comparative fit index, CFI)=.952, Tucker-Lewis 지수(Tucker-Lewis index, TLI)= .933, 근사오차평균제곱근(root mean square error of approximation, RMSEA)=.106이었고 요인상관은 .742～.880, 현저성–문제적 결과 이질특성–동질특성 비율(heterotrait-monotrait ratio, HTMT)= .887이었다. 전체 10문항의 Cronbach α=.910, ordinal α=.937로 내적 일관성은 높았지만 엄격한 단일차원성이나 진단적 타당성을 의미하지는 않았다. 따라서 총점은 응답 전반을 요약하는 기술적 종합지표이며, 높은 요인상관과 HTMT에 근거할 때 세 하위영역을 서로 독립된 점수로 강하게 해석하기보다 내용 프로파일로 활용하는 편이 더 신중하다. 연령정보가 없어 생애주기별 측정동일성은 검증하지 못했다.

주요어: 스마트폰 과의존, 측정구조, 점수 해석, WLSMV, 국가조사

* 주저자: 진수붕/국민대학교 일반대학원 사회학과 박사과정/서울시 성북구 정릉로 77
/Tel: 02-910-4470/E-mail: 385537328@qq.com

Ⅰ. 서론

스마트폰 과의존을 어떻게 측정하고 해석할 것인가는 단순히 사용시간이 길고 짧은지를 판단하는 문제와 구별될 필요가 있다. 스마트폰은 학습, 업무, 정보검색, 관계 유지와 여가가 동시에 이루어지는 다기능 매체이기 때문에 동일한 이용시간도 이용 목적과 통제 가능성, 기능적 결과에 따라 의미가 달라질 수 있다. 실제로 자기보고 이용시간과 기기 로그 사이에는 반복적인 불일치가 보고되었고(Ellis, 2019; Parry et al., 2021), 높은 이용량 자체를 문제적 사용이나 기능손상과 동일시해서는 안 된다는 지적이 제기되어 왔다(Kardefelt-Winther, 2014; Panova & Carbonell, 2018). 따라서 과의존 측정은 이용량뿐 아니라 조절의 어려움, 주의와 사고의 점유, 부정적 결과와 같은 경험적 특성을 함께 고려해야 한다.

한국에서는 2016년 한국정보화진흥원이 기존 인터넷(K-척도)과 스마트폰(S-척도)을 통합·정비하여 스마트폰 과의존 척도를 개편하였고, 조절실패, 현저성, 문제적 결과의 세 영역과 총 10문항 체계를 제시하였다(권선중, 엄나래, 2016). 당시 개편은 유아동·청소년·성인·고령층의 생애주기별 적용대상을 구분하고, 국가조사에서 위험수준을 일관되게 기술할 수 있는 통합 척도체계를 마련하는 데 목적이 있었다. 이후 이 척도는 국가승인통계인 「스마트폰 과의존 실태조사」에서 지속적으로 활용되어 총점에 기초한 위험수준 분류와 세 내용영역의 설명을 함께 제공해 왔다(과학기술정보통신부·한국지능정보사회진흥원, 2023). 즉 후속연구의 쟁점은 척도의 존재 가능성 자체보다, 이미 제도화되어 반복 사용되는 총점과 하위영역 점수에 각각 어느 정도의 해석 강도를 부여할 수 있는가에 있다.

이 지점에서 측정구조와 점수 사용의 구분이 중요하다. 상관된 3요인 모형이 단일요인 모형보다 더 나은 적합을 보인다고 해서 총점이 자동으로 부적절해지는 것은 아니다. 반대로 세 영역이 이론적으로 구분된다고 해서 각 하위점수가 일반 과의존 수준과 독립적인 고유 정보를 충분히 제공한다고 단정할 수도 없다. 요인 간 상관이 높고 문항들에 강한 일반요인이 존재하는 경우, 문항의 내용조직은 다차원적으로 유지되면서도 실제 점수의 신뢰분산은 일반요인에 크게 지배될 수 있다(Reise, 2012; Rodriguez et al., 2016). 따라서 '어떤 요인구조가 상대적으로 더 잘 맞는가'와 '총점 및 하위점수가 각각 무엇을 의미하는가'는 서로 연결되지만 구별되는 측정학적 질문이다.

2016년 개편 이후의 연구는 이 척도를 위험군 분류, 요인구조 확인, 하위영역별 실질변수 활용 등 서로 다른 방식으로 사용해 왔다. 박정혜(2019)는 국가조사 자료에서 공식 위험군 분류를 활용했고, 이현엽 외(2023)는 병사 표본에서 조절실패·현저성·문제적 결과의 3요인 구조를 확인했으며, 이문수 외(2024)는 변화한 디지털 환경을 반영한 청소년용 척도-R에서도 유사한 상위 구성틀을 유지하였다. 유승엽(2026)은 세 영역을 각각 설명변수로 사용하였다. 이 축적은 세 내용영역의 활용 가능성을 보여주지만, 제도화된 10문항 척도에서 구조적 구분, 문항·점수의 일관성, 영역 간 판별가능성이 실제 총점과 하위점수의 사용 범위를 어디까지 뒷받침하는지는 별도로 정리되지 않았다. 본 연구의 기여는 새로운 요인 수를 제안하는 데 있지 않다. 동일한 평균 및 분산 조정 가중최소제곱법(weighted least squares with mean and variance adjustment, WLSMV) 주분석 체계 안에서 문항 수준 정보, 경쟁구조, 절대적합, 요인 간 관련성, 이질특성-동질특성 비율(heterotrait-monotrait ratio, HTMT)와 신뢰도를 함께 검토하고, 또 그 결과를 총점의 기술적 요약·영역 프로파일·독립적 하위점수라는 서로 다른 점수 사용의 해석 기준과 연결하는 데 있다.

이에 본 연구는 2022년 「스마트폰 과의존 실태조사」 자기보고 마이크로데이터를 이용하여 두 가지

연구문제에 답한다. 첫째, 단일요인, 일부 영역을 결합한 2요인, 조절실패·현저성·문제적 결과의 상관 3요인 모형 중 어떤 구조가 동일한 WLSMV 추정틀에서 상대적으로 더 충분한 자료지지를 받는가? 둘째, 최종 모형의 robust 절대적합, 요인상관, HTMT와 총점·하위영역의 신뢰도를 함께 고려할 때 총점과 하위영역 점수는 각각 어느 범위까지 해석할 수 있는가? 국소잔차, 성별 층화와 주관적 과다사용 인식은 두 연구문제의 결론을 확장하기 위한 별도 가설검정이 아니라 해석의 경계와 결과의 안정성을 확인하는 보조자료로만 사용한다. 핵심 판단은 WLSMV 주분석에 한정한다.

Ⅱ. 이론적 배경

1. 스마트폰 과의존 척도의 개편과 국가조사 활용

2016년 척도 개편은 기존의 인터넷 및 스마트폰 중독 척도를 통합·정비하고, 스마트폰 과의존을 현저성, 조절실패, 문제적 결과로 설명하는 체계로 전환한 작업이었다(권선중, 엄나래, 2016). 현저성은 스마트폰 관련 생활패턴과 생각이 다른 활동보다 두드러지는 상태, 조절실패는 개인의 목표에 비해 스마트폰 이용을 자율적으로 조절하기 어려운 상태, 문제적 결과는 신체적·심리적·사회적 부정적 결과에도 불구하고 이용이 지속되는 상태를 가리킨다. 개편은 당시 국가조사의 적용범위가 유아동에서 고령층까지 확대되는 상황에서 생애주기별 측정체계를 정비하는 목적도 포함하였으며, 유아동 관찰자용과 청소년·성인·고령층용 척도를 구분해 제시하였다. 현행 국가조사는 청소년·성인 등 자기보고 대상에서 세 영역을 10문항으로 측정하고 총점을 과의존 위험수준 기술의 기본 지표로 사용한다(과학기술정보통신부·한국지능정보사회진흥원, 2023).

원 개발보고서의 검증 초점은 '척도를 어떻게 구성하고 위험수준 분류에 사용할 것인가'에 있었다. 보고서는 기존 K-척도와 S-척도의 개념 및 문항을 재검토하고, 생애주기별 적용대상에 맞추어 문항을 재구성한 뒤 경험자료 분석을 통해 최종 문항체계와 영역 구성을 정리하였다. 청소년·성인·고령층 자기보고형에서는 조절실패 3문항, 현저성 3문항, 문제적 결과 4문항의 10문항 구조가 채택되었고, 총점에 근거한 일반사용자군·잠재적위험군·고위험군 분류체계가 국가조사와 현장 활용에 연결되었다(권선중, 엄나래, 2016). 즉 원 개발연구는 문항구성, 영역구조, 내적 일관성과 위험군 판정에 사용할 수 있는 점수체계의 형성을 뒷받침했다. 다만 당시 보고서의 목적은 척도 형성과 분류체계의 확립에 있었으므로, 제도화 이후 총점과 세 하위영역을 병행 사용할 때 구조적 적합도, 영역 간 구별 가능성, 점수 신뢰도를 종합하여 각 점수에 어느 정도의 해석을 부여할 수 있는지는 별도의 문제로 남는다. 이 차이 때문에 원 개발연구의 타당화와 본 연구의 재검토는 중복되지 않는다. 전자가 '척도 형성 및 분류체계'를 확립했다면, 본 연구는 동일한 WLSMV 주분석 체계에서 구조적 적합도·요인 간 관련성·판별타당성·신뢰도 증거를 함께 검토하여 실제 점수 사용의 해석범위를 재점검한다.

이러한 구조는 실무적으로 두 종류의 정보를 동시에 제공한다. 하나는 전체 위험수준을 압축한 총점이고, 다른 하나는 조절실패·현저성·문제적 결과라는 내용별 프로파일이다. 그러나 두 종류의 점수가 함께 제공된다는 사실만으로 양쪽이 동일한 수준의 독립적 측정정보를 가진다고 볼 수는 없다. 총점과 하위영역은 서로 다른 추론을 요구하며, 특히 하위영역 간 상관이 높을 경우 하위점수의 차이가 일반 과의존 수준을 넘어서는 고유한 차이를 얼마나 반영하는지 별도로 확인해야 한다.

따라서 본 연구의 재검토는 원 개발연구의 결론을 반복하는 작업과 구별된다. 원 개발연구가 척도의 형성과 위험군 분류체계를 확립했다면, 본 연구는 제도화 이후 실제 점수 사용의 근거를 다시 점검한

다. 구체적으로 (a) 사전 정의된 세 영역의 상대적 구조가 현재 자료에서 어떻게 재현되는지, (b) 각 문항과 점수의 일관성이 어느 수준인지, (c) 영역 간 관련성과 판별가능성을 함께 고려할 때 총점과 하위영역을 각각 어디까지 해석할 수 있는지를 단계적으로 연결한다. 핵심은 '세 영역으로 구성되어 있다'는 구조적 진술과 '세 하위점수를 서로 독립된 개인차 점수로 사용할 수 있다'는 점수사용 진술을 분명하게 구분하여 제시하는 데 목적이 있다.

2. 선행연구의 사용 방식과 남은 측정학적 공백

2016년 개편 척도는 국가조사에서 위험수준을 산출하는 도구로 정착한 뒤, 후속 연구에서는 주로 과의존 수준을 설명변수 또는 결과변수로 활용하는 방식으로 확산되었다. 예를 들어 박정혜(2019)는 2017년 스마트폰 과의존 실태조사의 과의존 위험군 중학생 499명을 대상으로 스마트폰 사용행태와 자기인식을 분석하였다. 이 연구는 공식 척도의 분류 결과가 후속 실증연구에서 직접 활용되고 있음을 보여주지만, 연구목적 자체는 10문항의 요인구조나 하위점수의 독립성을 검증하는 데 있지 않았다.

척도 자체의 구조를 직접 검토한 연구도 존재한다. 이현엽 외(2023)는 1,651명의 병사 자료에서 확인적 요인분석을 실시하여 조절실패·현저성·문제적 결과의 3요인 구조가 원척도의 구성과 일치한다고 보고하였다. 이는 세 내용영역이 특정 집단에서도 재현될 수 있다는 중요한 근거다. 다만 해당 연구의 중심 목적은 병사 집단에서 척도의 요인구조를 확인한 뒤 자기통제, 회복탄력성, 리더십 등의 보호요인을 분석하는 것이었으며, 일반요인을 통제한 후 각 하위점수에 남는 고유한 신뢰분산까지 분리해 평가하지는 않았다.

최근의 척도 개선 연구는 또 다른 맥락을 제공한다. 이문수 외(2024)는 청소년 1,000명을 포함한 개발·타당화 절차를 통해 20문항의 스마트폰 과의존 척도-R을 제안하고, 조절력 상실·현저성·문제적 결과의 3개 상위 구성요인을 유지하였다. 한편 유승엽(2026)은 대학생 연구에서 조절실패, 현저성, 문제적 결과를 각각 설명변수로 사용하였다. 즉 후속 연구에서는 세 영역을 구조적 구성요인으로 확인하거나 실질적 하위변수로 활용하는 관행이 계속되고 있다. 바로 그렇기 때문에 현행 국가조사의 10문항 척도에서 세 하위점수가 총점과 구별되는 독립적 점수로 해석될 만큼 충분한 고유 정보를 갖는지는 별도의 질문으로 남는다.

이러한 선행연구를 종합하면 기존 연구가 해결한 문제와 본 연구가 다루는 후속 문제를 구분할 수 있다. 원 개발연구는 통합 척도와 위험군 분류체계를 형성했고, 후속 요인연구는 세 내용영역의 재현 가능성을 제시했으며, 활용 연구는 총점·위험군 또는 하위영역을 실증분석의 변수로 사용해 왔다. 그러나 이 축적만으로는 세 하위영역이 일반 과의존 수준을 넘어 독립적인 개인차 정보를 충분히 제공하는지, 그리고 총점의 높은 내적 일관성을 어느 수준의 점수 해석까지 확장할 수 있는지 알 수 없다. 본 연구는 이 문제를 문항 수준 정보와 WLSMV 구조적 적합도, 잠재요인 간 관련성, HTMT, 총점·하위영역 신뢰도를 함께 검토함으로써 다룬다. 따라서 본 연구의 독창성은 새로운 척도나 이론을 제안하는 데 있지 않고, 이미 제도화된 국가척도에서 '내용영역의 구분', '총점의 기술적 사용', '하위점수의 독립적 사용'을 구별하여 실제 연구와 실무에서 각 점수 사용의 해석 범위를 구체화한다는 데 있다.

3. 요인구조와 점수 해석은 동일한 질문이 아니다

확인적 요인분석에서 상관 3요인 모형은 각 문항이 사전에 정의된 영역에 주로 적재되고 세 잠재요인은 상호상관을 갖는다고 가정한다. 이러한 모형이 단일요인 모형보다 우수하면 세 내용영역을 구분하

는 설명이 상대적으로 더 적합하다고 볼 수 있다. 그러나 요인상관이 매우 높거나 판별타당성 지표가 경계수준을 넘는다면, 세 요인을 서로 독립적인 개인차 점수로 해석하는 데는 주의가 필요하다. HTMT와 같은 지표는 이 문제를 보완적으로 확인하는 방법이다(Henseler et al., 2015).

점수 해석에서는 구조적 구분 외에 신뢰도와 판별가능성을 함께 볼 필요가 있다. 전체 10문항의 내적 일관성이 높다는 것은 문항응답을 하나의 합산값으로 요약했을 때 그 값이 문항 간 공분산을 일관되게 반영한다는 뜻이지, 그 합산값이 하나의 순수한 잠재특성을 정확히 측정하거나 임상적·정책적 판정을 정당화한다는 뜻은 아니다. 반대로 각 하위영역의 α가 높더라도 영역 간 잠재상관과 HTMT가 매우 높으면, 그 하위점수가 다른 영역과 구별되는 독립적 개인차를 충분히 담는다고 보기 어렵다. 따라서 본 연구의 주 판단은 WLSMV 최종 모형의 robust 적합도, 요인상관과 HTMT, 그리고 총점·하위영역 신뢰도를 함께 읽되 각 지표가 답할 수 있는 질문의 범위를 넘어서지 않는 방식으로 구성하였다.

본 연구의 핵심 판단은 동일한 서열형 추정틀에서 산출한 결과들로 구성하였다. 구체적으로 WLSMV 경쟁모형의 상대적 비교와 최종모형의 robust 절대적합을 구분하고, 요인상관과 HTMT를 신뢰도 결과와 함께 검토하였다. 이때 각 지표는 서로 다른 측정학적 질문에 답하는 정보로 명확히 구분하여 사용하였다. 모형의 상대적 우세가 하위점수의 독립성을 직접 입증하는 것으로 보지 않았고, 높은 신뢰도 역시 엄격한 단일차원성이나 진단적 타당성의 근거로 확대하지 않았다. 따라서 연구문제 1과 2의 결론은 WLSMV 주분석과 그에 직접 연결되는 신뢰도·판별가능성 정보의 범위에서만 제시하였다.

4. 국가조사 척도의 재검토가 필요한 이유

국가조사에서 반복적으로 사용되는 척도는 동일한 점수가 조사보고, 위험군 분류, 예방교육, 상담 및 후속 학술연구로 이어지기 때문에 '각 점수를 어디까지 해석할 것인가'가 특히 중요하다. 총점과 세 하위영역이 동시에 제공되는 경우, 총점은 전체 위험수준을 압축하는 기능을 하고 하위영역은 어려움의 내용을 구분하는 기능을 할 수 있다. 그러나 하위점수를 서로 독립된 위험등급처럼 사용하는 것과 내용 프로파일로 사용하는 것은 요구되는 측정학적 근거가 다르다. 본 연구는 이 차이를 명시함으로써 국가조사 척도의 점수 사용을 세 수준—총점 요약, 영역 프로파일, 독립적 하위점수—으로 나누어 각 수준에 필요한 해석 조건을 구분한다.

자료가 허용하는 분석범위도 사전에 구분하였다. 확보한 비식별 마이크로데이터에는 임상진단, 기기로그, 개인별 연령·교육·소득·직업·지역 및 개인가중치가 포함되어 있지 않다. 따라서 본 연구는 자기보고 10문항의 문항 수준 작동, 측정구조, 점수 신뢰정보와 동일 설문 내 보조변수와의 횡단적 관련성에 초점을 둔다. 임상적 준거타당도, 수신자 조작 특성(receiver operating characteristic, ROC)/곡선하 면적(area under the curve, AUC) 기반 절단점 재설정, 모집단 유병률, 연령별 측정동일성 또는 인과효과는 현재 자료로 답할 수 없는 별도의 질문이다. 이러한 구분은 분석의 결손을 숨기기 위한 것이 아니라, 이미 개발된 척도에서 현재 자료가 확인할 수 있는 범위를 명확히 하기 위한 연구설계상의 한계 설정이다.

Ⅲ. 연구방법

1. 자료와 분석대상

본 연구는 과학기술정보통신부의 국가승인통계

제120019호인 「2022년 스마트폰 과의존 실태조사」 비식별 마이크로데이터를 이용하였다. 조사 보고서와 조사표는 한국지능정보사회진흥원 ‘스마트폰 과의존 실태조사’ 게시판에서 공개되어 있으며(https://www.nia.or.kr/site/nia_kor/ex/bbs/List.do?cbIdx=65914), 마이크로데이터 이용 경로는 ICT 통계포털에서 통계청 MDIS 또는 별도 신청 절차로 안내되고 있다(https://www.itstat.go.kr; 접속일 2026. 9. 22). 본 연구에서 2022년 자료를 사용한 이유는 연구 설계 당시 문항별 응답구조와 결측패턴을 확인할 수 있는 비식별 마이크로데이터를 확보한 차수였으며, 후속 자료의 공개·확보에는 시차가 있었기 때문이다. 연구 목적은 2026년 현재의 최신 과의존 수준이나 시계열 변화를 추정하는 것이 아니라, 동일 척도체계가 적용된 한 조사차수에서 10문항의 측정구조와 점수 해석 범위를 검토하는 데 있다. 확보한 마이크로데이터는 전체 24,638명으로 구성되어 있었다(과학기술정보통신부·한국지능정보사회진흥원, 2023). 자기보고용 과의존 10문항은 만 10세 이상에게 적용되고 만 3~9세에는 보호자용 별도 척도가 적용된다. 이에 따라 23,097명이 자기보고용 10문항에 모두 응답하였고 1,541명은 해당 10문항에 응답하지 않았다. 1~9개 문항만 응답한 부분응답 사례는 없었으므로 자기보고 척도가 적용된 23,097명을 최종 분석대상으로 하였다.

주분석 표본은 남성 11,259명(48.7%), 여성 11,838명(51.3%)이었다. 주관적 과다사용 인식의 응답은 1점 1,679명, 2점 12,628명, 3점 8,294명, 4점 496명이었다. 확보한 비식별 마이크로데이터에는 성별 외에 개인별 연령, 교육, 소득, 직업, 지역 및 개인가중치가 포함되어 있지 않았기 때문에 해당 변수에 대한 보정이나 모집단 가중추정은 수행하지 않았다.

<표 1> 분석표본의 구성

구분	N	비율/처리
비식별 마이크로데이터 전체	24,638	100.0%
자기보고 10문항 완전응답	23,097	93.7%; 최종 분석표본
자기보고 10문항 비응답	1,541	6.3%; 공식 조사설계상 만 3~9세는 보호자용 척도 적용
남성	11,259	최종 분석표본의 48.7%
여성	11,838	최종 분석표본의 51.3%

주: 자기보고용 10문항은 만 10세 이상에게 적용된다. 만 3~9세에는 보호자용 별도 척도가 적용되며, 자기보고 10문항의 부분응답 사례는 없었다.

2. 측정변수

스마트폰 과의존 자기보고 척도는 조절실패(self-control failure, CF) 3문항(CF1~CF3), 현저성(salience, SA) 3문항(SA1~SA3), 문제적 결과(problematic consequences, PR) 4문항(PR1~PR4)의 총 10문항으로 구성되며, 각 문항은 1점(전혀 그렇지 않다)에서 4점(매우 그렇다)까지 응답한다. 주관적 과다사용 인식은 “하루 평균 스마트폰을 이용하는 시간이 과도하다고 생각하는가”에 대한 4점 단일문항이다. 과다사용 인식은 객관적 이용시간이나 임상적 진단이 아니며, 본 연구에서는 외부 준거라기보다 동일 설문 내 보조적 수렴정보로만 사용하였다. 문항 수준 점검에서는 각 문항 평균, 해당 문항을 제외한 나머지 9문항 합과의 수정 문항–총점 상관, 최종 WLSMV 모형의 표준화 부하량을 함께 확인하였다.

<표 2> 과의존 10문항의 기술통계, 수정 문항-총점 상관과 표준화 부하량

영역	코드	문항 요약	M	수정 문항–총점 r	λ
조절실패	CF1	이용시간을 줄이려 할 때마다 실패한다	2.050	.679	.868

영역	코드	문항 요약	M	수정 문항–총점 r	λ
조절실패	CF2	이용시간을 조절하기 어렵다	2.087	.668	.853
조절실패	CF3	적절한 이용시간을 지키기 어렵다	2.042	.651	.824
현저성	SA1	옆에 있으면 다른 일에 집중하기 어렵다	1.923	.683	.808
현저성	SA2	스마트폰 생각이 머리에서 떠나지 않는다	1.833	.681	.819
현저성	SA3	이용하고 싶은 충동을 강하게 느낀다	1.833	.681	.810
문제적 결과	PR1	이용 때문에 건강 문제가 생긴 적이 있다	1.822	.635	.770
문제적 결과	PR2	이용 때문에 가족과 심하게 다툰 적이 있다	1.766	.670	.822
문제적 결과	PR3	사회적 관계에서 심한 갈등을 경험했다	1.731	.687	.862
문제적 결과	PR4	업무·학업 수행에 어려움이 있다	1.773	.715	.887

주: N=23,097. 수정 문항–총점 r은 해당 문항을 제외한 나머지 9문항 합과의 Pearson 상관이며, λ는 WLSMV 상관 3요인 모형(M3)의 표준화 요인부하량이다.

3. 분석전략

분석은 두 연구문제에 맞추어 세 단계로 구성하였다. 첫째, 문항 평균·수정 문항–총점 상관·표준화 부하량과 전체 및 하위영역 신뢰도를 확인하여 기본적인 문항·점수 수준 정보를 점검하였다. 둘째(연구문제 1), 사전에 정의한 세 측정구조를 동일한 WLSMV로 비교하고 상대적으로 가장 나은 모형의 robust 절대적합을 별도로 평가하였다. 셋째(연구문제 2), 최종모형의 요인상관과 HTMT를 신뢰도와 함께 검토하여 총점·영역 프로파일·독립 하위점수의 해석 범위를 구분하였다. 국소잔차, 성별 층화와 주관적 과다사용 인식은 핵심 결론을 추가하는 별도의 분석이 아니라 해석의 경계와 결과의 안정성을 확인하는 보조자료로만 요약하였다. 주분석의 결론은 동일한 WLSMV 추정틀과 직접 연결되는 정보에 한정하였다.

1) 주 측정모형 비교

첫 번째 연구문제를 위해 4점 서열형 문항의 특성을 고려하여 세 경쟁모형을 WLSMV로 추정하였다(Rosseel, 2012; DiStefano & Morgan, 2014). M1은 10문항을 하나의 요인으로 보는 단일요인 모형, M2는 조절실패와 '현저성+문제적 결과'를 구분한 2요인 모형, M3는 조절실패·현저성·문제적 결과를 각각 구분한 상관 3요인 모형이다. 모형 간 상대비교에서는 동일한 WLSMV 추정틀의 조정 χ^2와 표준화 잔차평균제곱근(standardized root mean square residual, SRMR)을 제시하고, 중첩관계가 명확한 M2와 M3의 차이는 평균·분산조정 차이검정으로 확인하였다(Satorra, 2000). 반면 최종 M3 자체의 절대적합 판단에는 robust 비교적합지수(comparative fit index, CFI)·Tucker-Lewis 지수(Tucker-Lewis index, TLI)·근사오차평균제곱근(root mean square error of approximation, RMSEA)만 사용하였다(Hu & Bentler, 1999). 이처럼 '모형 간 상대비교'와 '최종모형의 절대적합'을 서로 다른 단계로 분리하고, 동일한 모형에 대해 산출방식이 다른 RMSEA를 한 표에 병렬 제시하지 않음으로써 지표 해석의 혼동을 최소화하였다. 최종 점수 해석에서는 여기에 요인상관과 HTMT를 함께 고려하였다(Henseler et al., 2015).

본 연구는 이미 개발되어 반복 사용 중인 2016년 척도의 사전 정의된 구조를 후속 검토하는 연구이므로, 새로운 요인 수를 탐색하기 위한 탐색적 요인분석(exploratory factor analysis, EFA)를 핵심 절차로 두지 않았다. 대신 원 척도의 이론적 구조에서 직접 파생된 경쟁모형을 동일한 WLSMV로 비교하

고, 문항 수준에서는 평균·수정 문항–총점 상관·표준화 부하량을 확인하였으며, 점수 수준에서는 Cronbach α와 ordinal α, 영역 간 구별 가능성에서는 요인상관과 HTMT를 함께 검토하였다. 이 절차는 문항의 기본적 일관성, 구조적 적합성, 점수 신뢰도와 판별가능성을 연구문제에 직접 연결하기 위한 것이다. 임상진단이나 독립 기능손상 기준이 없으므로 ROC/AUC와 절단점 재설정은 수행하지 않았고, 동일한 이유로 본 연구는 공식 위험분류의 재타당화를 주장하지 않는다. 따라서 본 연구의 범위는 '완전한 척도 재개발'이 아니라 '제도화된 10문항의 구조와 점수 사용에 대한 후속 측정학적 증거 검토'이다.

2) 총점과 하위영역의 점수 해석 진단

두 번째 연구문제는 주분석과 동일한 분석체계 안에서 판단하였다. 전체 10문항과 세 하위영역의 Cronbach α 및 polychoric 상관에 기초한 ordinal α를 산출하여 내적 일관성을 확인한 뒤, 최종 WLSMV 3요인 모형의 robust CFI·TLI·RMSEA, 잠재요인 상관, HTMT를 함께 검토하였다. 신뢰도는 '문항응답 합산값의 일관성'에 관한 정보로만 사용하였다. 따라서 전체 10문항의 α가 높더라도 엄격한 단일차원성, 단일 잠재특성의 정밀 측정, 개인 진단이나 위험군 분류의 정당화로 확대하지 않았다. 하위영역의 자체 신뢰도가 높더라도 높은 요인상관과 HTMT는 영역 간 판별가능성이 제한적일 수 있음을 간접적으로 시사하므로, 각 하위점수를 서로 독립된 개인차 점수로 강하게 해석하기에는 근거가 충분하지 않은 것으로 판단하였다. 따라서 점수 사용을 세 수준으로 구분하였다.

3) 보조 민감도 분석의 역할

보조분석은 세 가지로 제한하였다. M3의 국소잔차는 특정 문항쌍의 내용중첩 가능성을 점검하고, 성별 층화는 동일한 3요인 모형의 기본 적합 양상이 두 성별 표본에서 대체로 유사한지를 SRMR로 확인하며, 주관적 과다사용 인식과 세 영역의 관련성은 동일 설문 내 수렴정보를 확인한다. 이러한 결과들은 정식 측정동일성, 독립 외부준거 타당도 또는 인과효과를 입증하는 자료로 사용하지 않았으며, 보조분석 자체에서 추가적인 핵심 결론을 도출하지 않았다.

4. 연구윤리와 AI 활용 고지

본 연구는 비식별 기존 마이크로데이터를 이용한 2차 분석이며, 연구대상자와의 직접 접촉이나 개인식별정보 처리는 이루어지지 않았다. 원자료는 재배포하지 않고 집계된 분석결과만 제시하였다. 생성형 인공지능(artificial intelligence, AI)는 문헌 정리, 원고 구조 점검, 분석 코드 작성·검산 보조 및 문장 교정에 활용하였으며, 변수 매핑, 통계분석 선택, 결과 확인과 해석, 참고문헌 검토 및 최종 원고의 책임은 저자에게 있다.

Ⅳ. 연구결과

1. 기술통계와 신뢰도

원자료의 평균과 표준편차(standard deviation, SD)를 재확인한 결과, 주분석 표본의 과다사용 인식 평균은 2.33(SD=.64), 조절실패 2.06(SD=.68), 현저성 1.86(SD=.64), 문제적 결과 1.77(SD=.60), 10문항 전체평균 1.89(SD=.56)이었다. 표 2의 문항별 평균은 1.731~2.087, 수정 문항–총점 상관은 .635~.715, 최종 M3의 표준화 요인부하량은 .770~.887 범위였다. 따라서 수정 문항–총점 상관과 부하량만을 기준으로 볼 때 특정 문항 하나가 현저하게 고립되는 패턴은 관찰되지 않았다. 다만 이러한 지표

는 문항삭제 시 α 변화, 상·하위집단 변별도, 임상 준거와의 분류정확도를 대체하지 않으므로 문항의 모든 심리측정 특성이 확인되었다고 해석하지 않았다. Cronbach α는 조절실패 .827, 현저성 .789, 문제적 결과 .847, 전체 10문항 .910이었고 ordinal α는 각각 .882, .849, .899, .937이었다. 하위영역 간 Spearman 상관은 조절실패–현저성 .678, 조절실패–문제적 결과 .623, 현저성–문제적 결과 .712였으며, 과다사용 인식과 하위영역 평균의 Spearman 상관은 각각 .485, .437, .373이었다.

<표 3> 주요 변수의 기술통계와 신뢰도

변수	M	SD	왜도	Cronbach α	ordinal α
과다사용 인식	2.33	.64	.069	—	—
조절실패	2.06	.68	.307	.827	.882
현저성	1.86	.64	.641	.789	.849
문제적 결과	1.77	.60	.602	.847	.899
10문항 전체평균	1.89	.56	.486	.910	.937

주: N=23,097. 하위영역 간 Spearman ρ는 CF–SA=.678, CF–PR=.623, SA–PR=.712였다. 과다사용 인식과 CF, SA, PR의 Spearman ρ는 각각 .485, .437, .373이었다.

2. 주 측정모형의 상대적 비교

첫 번째 연구문제에 대한 WLSMV 경쟁모형 비교에서 단일요인 M1의 조정 χ^2는 14,026.72(df=35), SRMR=.064였고, 현저성과 문제적 결과를 결합한 M2는 조정 χ^2=6,855.48(df=34), SRMR=.044로 기술적으로 개선되었다. 조절실패·현저성·문제적 결과를 구분한 상관 3요인 M3는 조정 χ^2=3,288.04(df=32), SRMR=.030으로 세 모형 중 가장 작은 값을 보였으며, M2 대비 평균·분산조정 차이검정도 유의하였다($\Delta\chi^2$=1,736.23, Δdf=2, p<.001). 따라서 사전에 정한 경쟁구조 가운데 M3가 상대적으로 가장 충분한 자료지지를 받았다고 판단하였다. 이 판단은 '3요인 구조가 절대적으로 잘 맞는다'는 의미가 아니라, 검토한 세 사전모형 사이의 상대적 선택을 뜻한다.

최종 M3의 절대적합은 모형 간 비교와 분리하여 robust 지표로 평가하였다. robust CFI=.952, TLI=.933, RMSEA=.106으로, 특히 RMSEA는 절대적합에 분명한 한계가 있음을 보여주었다. 잠재요인 상관은 조절실패–현저성 .848, 조절실패–문제적 결과 .742, 현저성–문제적 결과 .880으로 높았고, HTMT도 각각 .847, .750, .887이었다. 특히 현저성–문제적 결과 HTMT=.887은 보수적인 .85 참고기준을 넘었다. 따라서 구조적 결론은 '검토한 경쟁모형 가운데 3요인 구성이 상대적으로 가장 적절하지만, 최종모형의 절대적합과 세 영역의 판별가능성에는 유의한 한계가 남는다'는 수준으로 제한하였다. 표 4에는 모형 간 비교에 필요한 정보만 제시하고, 최종 M3의 robust CFI/TLI/RMSEA는 본문에서 절대적합 판단에만 사용하였다.

<표 4> WLSMV 경쟁 측정모형의 상대비교

모형	구조	조정 χ^2(df)	SRMR	상대비교 해석
M1	단일요인	14,026.72 (35)	.064	기준모형
M2	CF/ (SA+PR)	6,855.48 (34)	.044	M1보다 기술적으로 개선
M3	CF/ SA/PR	3,288.04 (32)	.030	M2 대비 $\Delta\chi^2$=1,736.23, Δdf=2, p<.001; 상대적으로 최선

주: 조정 χ^2와 SRMR은 동일한 WLSMV 추정틀에서 세 사전모형의 상대적 비교를 위해 제시하였다. 최종 M3의 절대적합은 본문에서 별도로 보고한 robust CFI=.952, TLI=.933, RMSEA=.106으로 판단하였다. M3 대 M2 평균·분산조정 차이검정: $\Delta\chi^2$=1,736.23, Δdf=2, p<.001. 따라서 표의 모형비교 정보와 본문의 robust 절대적합 정보는 서로 다른 판단단계를 담당하며, 동일 지표의 상충된 판정으로 해석하지 않는다.

3. 총점과 하위영역의 점수 해석 범위

두 번째 연구문제는 WLSMV 최종모형의 구조 정보와 신뢰도 결과를 함께 고려하여 판단하였다. 전체 10문항의 Cronbach α=.910, ordinal α=.937로 문항응답 합산값의 내적 일관성은 높았다. 그러나 이 값은 '10개 응답이 서로 일관되게 함께 변한다'는 점을 보여줄 뿐, 합산값이 하나의 잠재특성을 엄격하게 측정한다거나 개인 진단·위험군 분류를 정당화한다는 근거는 아니다. 실제 분석에서 M3의 robust RMSEA=.106과 높은 요인상관은 단일차원적 해석을 제한한다. 하위영역의 Cronbach α는 .789~.847, ordinal α는 .849~.899로 양호했지만, 잠재요인 상관 .742~.880과 HTMT 최대 .887은 영역 간 판별가능성이 제한될 수 있음을 보여주는 간접적 근거로 신중하게 해석하였다.

따라서 본 분석자료와 자기보고 10문항의 범위에서 방어 가능한 점수 사용은 다음과 같다. 10문항의 합 또는 평균은 응답 전반을 요약하는 기술적 종합지표로 보고할 수 있으나, 이를 '단일 잠재특성 점수', '개인 진단점수', '공식 위험군 절단점의 재타당화 결과', '연령집단 간 직접 비교점수'로 해석해서는 안 된다. 조절실패·현저성·문제적 결과 점수는 응답자가 어떤 유형의 어려움을 상대적으로 더 많이 보고했는지 설명하는 내용 프로파일로 활용할 수 있다. 다만 높은 요인상관과 HTMT는 하위영역의 독립성이 제한적일 가능성을 간접적으로 보여주므로, 서로 독립된 위험등급이나 별개의 진단점수로 강하게 해석하기에는 현재의 근거가 충분하지 않다. 이는 독립성을 직접 검증한 결과가 아니라 요인상관과 HTMT에 근거한 간접적 판단이다.

4. 보조 민감도 분석의 요약

세 보조분석은 <표 5>에 압축해 제시하였다. 최대 국소잔차는 |.091|(SA1−CF3)이었고, 성별 분리 적합의 SRMR은 남성 .031, 여성 .034였으며, 주관적 과다사용 인식과 조절실패·현저성·문제적 결과의 잠재상관은 각각 .601, .549, .460이었다. 이 결과는 각각 국소적 내용중첩 가능성, 성별 분리 SRMR 수준의 유사성, 동일 설문 내 수렴정보를 보여주는 범위에서 해석하였다. 사후 잔차상관으로 주 모형 적합도를 개선하지 않았고, 성별 결과를 정식 측정동일성으로, 과다사용 인식을 독립적 임상·객관 준거로 해석하지 않았다.

<표 5> 핵심 보조 민감도 분석의 요약

보조분석	주요 결과	허용되는 해석
국소잔차	최대 |잔차|=.091 (SA1−CF3)	국소적 내용중첩 가능성; 적합도 원인의 직접 분해 아님
성별 층화	SRMR 남 .031 / 여 .034	SRMR 유사성; 정식 측정동일성 아님
과다사용 인식	CF .601 / SA .549 / PR .460	동일 설문 내 수렴정보; 임상·객관 준거 아님

V. 논의

1. 기존 국가척도에 대한 재검토의 의미

본 연구는 2016년 개편 이후 지속적으로 사용되어 온 스마트폰 과의존 척도의 후속 검토에서 세 가지 점을 명확히 했다. 첫째, 원 개발연구가 확립한 세 내용영역을 전제로 단일요인·2요인·상관 3요인 구조를 동일한 WLSMV 틀에서 직접 비교하여 내용구조의 상대적 자료지지를 평가하였다. 둘째, 수정 문항−총점 상관과 신뢰도 같은 문항·점수 수준 정보와 최종 3요인 모형의 robust 절대적합, 요인상관·HTMT를 분리해 읽음으로써 '일관성'과 '구별 가능성'을 같은 개념으로 취급하지 않았다. 셋째, 이 결과를 실제 점수 사용과 연결하여 기술적 총점, 영역 프로파일, 독립적 하위점수의 해석 범위

를 구분하였다. 따라서 본 연구의 기여는 새로운 요인이나 진단도구를 제안하는 데 있지 않고, 제도화된 국가척도에서 각 분석결과가 어떤 점수 해석을 뒷받침하는지를 구체화한 데 있다. 이현엽 외(2023)의 병사 표본과 이문수 외(2024)의 청소년 척도-R이 세 내용영역의 재현 가능성을 보여주었다면, 본 연구의 robust RMSEA=.106, 높은 요인상관과 HTMT=.887은 세 영역을 독립 점수로 강하게 해석하기에는 추가 근거가 필요함을 간접적으로 시사한다. 이는 하위영역의 독립성을 직접 검증한 결과는 아니다.

이 구분은 기존 활용연구의 결과를 부정하기보다 그 해석 조건을 명료하게 한다. 박정혜(2019)처럼 공식 위험군을 이용하거나 유승엽(2026)처럼 하위영역을 실질변수로 사용하는 연구에서는 점수 사용이 분석의 출발점이 된다. 본 연구는 그 출발점의 측정학적 근거를 점검하여, 세 영역을 내용 프로파일로 제시하는 것과 서로 독립적인 개인차 점수로 사용하는 것은 동일한 추론이 아님을 보였다. 즉 본 연구의 추가적 가치는 통계기법의 수 자체가 아니라, 원 개발연구가 마련한 척도와 분류체계를 제도화 이후의 점수 사용 증거와 다시 연결했다는 데 있다.

2. 세 내용영역과 문항응답 합산값의 해석 수준

본 연구의 핵심 결과는 내용구조와 점수의 독립성이 같은 문제가 아니라는 점이다. M3의 상대적 우세는 10문항을 조절실패·현저성·문제적 결과의 세 내용영역으로 구분하는 설명이 검토한 비교모형보다 더 적절하다는 것을 보여준다. 그러나 robust RMSEA=.106과 높은 요인상관, HTMT=.887은 이 세 영역이 서로 선명하게 분리되는 독립 척도라고 보기 어렵다는 점을 동시에 보여준다. 한편 전체 10문항의 α=.910, ordinal α=.937은 문항응답 합산값의 높은 내적 일관성을 보여주지만, 그 자체로 엄격한 단일차원성이나 진단적 타당성을 입증하지 않는다. 따라서 구조적으로 세 영역이 상대적으로 유용하다는 결과와 10문항을 기술적으로 합산할 수 있다는 사실은 서로 양립할 수 있지만, 어느 쪽도 강한 진단·분류 해석으로 자동 확장되어서는 안 된다.

이 결과는 '총점과 하위척도 중 어느 하나만 선택해야 한다'는 이분법보다 해석의 수준을 구분할 필요가 있음을 시사한다. 본 분석자료와 현행 자기보고 10문항의 범위에서 합산점수는 응답 전반을 요약하는 기술적 종합지표로 제한하여 사용할 수 있고, 세 하위영역은 그 응답이 어떤 종류의 어려움으로 구성되는지를 설명하는 내용정보로 활용할 수 있다. 그러나 합산점수를 하나의 잠재특성에 대한 정밀 점수, 개인 진단점수, 정책적 위험분류의 재검증 결과로 해석할 근거는 현재 분석에 포함되어 있지 않다. 마찬가지로 높은 요인상관과 HTMT는 하위영역 간 판별가능성의 한계를 간접적으로 보여주므로, 각 하위영역을 총점과 독립된 개인별 위험등급이나 별개의 진단점수로 강하게 사용하는 것은 현재 자료만으로 충분히 정당화되지 않는다. 따라서 독립성 자체의 직접 검증으로 볼 수는 없다.

3. 보조분석이 보여주는 경계와 실천적 함의

보조분석은 주결론의 방향을 바꾸지 않았다. 일부 문항쌍의 국소잔차는 조절실패와 현저성의 내용이 실제 경험에서 중첩될 수 있음을 보여주었고, 성별 분리 적합에서 남성 SRMR=.031, 여성 SRMR=.034로 유사하여 기본 적합 양상이 특정 성별에 국한되지 않을 가능성을 시사했다. 주관적 과다사용 인식과 세 영역의 관련성 역시 동일 설문 안에서 수렴하는 방향을 보였다. 그러나 어느 결과도 측정동일성, 독립 외부준거 타당도나 인과관계를 입증하지 않는다. 따라서 보조분석은 '추가 결론'을 생산하는 층이 아니라 주 측정구조와 점수 해석의 경계를 확인하는 자료로 한정하였다.

실무적으로도 적용범위를 본 분석자료와 현행 자기보고 10문항에 한정할 필요가 있다. 첫째, 10문항 합산값은 개인의 응답 전반을 요약하는 기술적 지표로 제시할 수 있지만, 본 연구는 기존 공식 위험군 절단점의 정확성이나 임상적 진단성을 재검증한 연구가 아니다. 둘째, 조절실패·현저성·문제적 결과는 어떤 유형의 어려움이 상대적으로 두드러지는지 설명하는 영역 프로파일로 활용할 수 있다. 셋째, 각 하위점수를 서로 독립된 위험등급이나 임상적 절단점으로 사용하는 것은 현재 결과만으로 충분히 정당화되지 않는다. 또한 개인가중치가 없는 본 분석으로 전국 모집단의 위험분포를 재추정하거나, 연령정보 없이 생애주기 간 점수의 직접 비교 가능성을 가정해서는 안 된다. 따라서 상담·예방교육·자기점검에서 하위영역을 활용하더라도 '독립 진단점수'보다는 '내용적 보조정보'로 취급하는 것이 더 신중하다.

4. 한계와 후속연구

첫째, 본 연구는 2022년의 단일 횡단면 자기보고 자료를 사용하였으므로 인과방향, 시간적 순서, 매개효과뿐 아니라 이후 연도에서 동일한 측정구조가 재현되는지 까지 검증할 수 없다. 따라서 2022년 분석결과를 2026년 현재의 이용행태나 최근 추세를 직접 대표하는 것으로 해석해서는 안 된다. 둘째, 확보한 비식별 마이크로데이터에는 개인별 연령, 교육, 소득, 직업, 지역, 개인가중치가 포함되어 있지 않아 연령별 측정동일성, 인구사회학적 공변량 조정, 가중 모집단 추정을 수행하지 못했다. 따라서 결과는 현재 분석표본 내부의 측정구조와 점수 해석에 관한 것이며 전국 모집단의 평균이나 유병률 추정으로 확장해서는 안 된다.

셋째, 본 연구는 이미 개발된 국가척도의 사전 정의된 구조를 후속 검토하는 목적이므로 EFA를 핵심 분석으로 사용하지 않았다. 문항별 평균, 수정 문항–총점 상관, 표준화 부하량을 제시하고 사전 정의된 경쟁모형을 확인적 요인분석(confirmatory factor analysis, CFA) 방식으로 비교하였다. 이는 완전한 척도 재개발 절차와 구별되는 의도적인 설계 선택이다. 문항삭제 시 신뢰도 변화와 상·하위 집단 문항변별도는 현재 자료에서도 산출 가능하지만, 본 연구는 척도 재개발 목적이 아니어서 핵심 분석에 포함하지 않았다. 향후 독립 표본이나 다른 연도 자료에서는 EFA–CFA 교차검증을 추가할 필요가 있다. 넷째, 성별 층화 결과는 정식 ordinal measurement invariance를 대체하지 않으며, 향후 연령과 성별 정보를 포함한 다집단 WLSMV 측정동일성 검정이 필요하다.

다섯째, 하위영역의 독립성에 관한 판단은 잠재요인 상관과 HTMT에 기초한 간접적 판단이다. 이 지표들은 영역 간 판별가능성을 평가하는 데 유용하지만, 하위점수에 일반 수준과 구별되는 고유 정보가 얼마나 남는지를 직접 추정하지는 않는다. 따라서 본 연구는 하위점수의 독립성이 없다고 단정하지 않고, 현재 자료만으로 강한 독립적 해석을 지지하기 어렵다는 수준으로 결론을 제한하였다. 향후 연구에서는 동일한 서열형 추정틀에서 추가 점수분해 분석을 수행하여 이를 재검토할 필요가 있다. 여섯째, 자료에 임상진단, 객관적 사용로그, 독립적 기능손상 기준이 없어 임상적 준거타당도를 평가하거나 ROC/AUC를 통해 위험군 절단점을 재설정할 수 없었다. 주관적 과다사용 인식 역시 단일 자기보고 문항이므로 외부 준거로 보기 어렵다. 향후 연구에서는 임상 또는 기능손상 기준, 기기로그, 종단자료를 결합하여 총점과 하위영역의 예측타당도와 분류정확도를 검증할 필요가 있다.

Ⅵ. 결론

2022년 스마트폰 과의존 실태조사의 자기보고 10문항을 후속 검토한 결과, 문항별 수정 문항–총점

상관은 .635~.715, 최종 3요인 모형의 표준화 부하량은 .770~.887로 나타났고, 조절실패·현저성·문제적 결과의 상관 3요인 모형은 동일한 WLSMV 틀에서 사전에 정한 단일요인과 일부 결합 2요인보다 상대적으로 더 충분한 자료지지를 받았다. 그러나 최종 M3의 robust CFI=.952, TLI=.933, RMSEA=.106, 높은 요인상관(.742~.880)과 HTMT(최대 .887)는 절대적합과 영역 간 판별가능성에 한계가 있음을 보여주었다. 전체 10문항의 Cronbach α=.910과 ordinal α=.937은 문항응답 합산값의 높은 내적 일관성을 보여주지만, 이 결과만으로 엄격한 단일차원성, 진단적 타당성 또는 공식 위험분류 기준의 타당성을 확인할 수는 없다. 하위영역의 신뢰도도 양호했으나 높은 요인상관과 HTMT는 독립성의 한계를 간접적으로 시사하며 각 하위점수를 독립된 개인차 점수로 강하게 해석할 근거는 부족하다.

따라서 이 척도는 '엄격한 단일차원 척도' 또는 '서로 독립적인 세 하위척도' 중 하나로 단순화하기보다 점수별 해석범위를 구분하는 것이 타당하다. 본 분석 자료와 현행 자기보고 10문항의 범위에서 10문항 합산값은 응답 전반을 요약하는 기술적 종합지표로 제한하여 해석할 수 있고, 조절실패·현저성·문제적 결과는 응답을 구성하는 내용영역을 설명하는 프로파일로 보는 편이 더 신중하다. 반면 합산값을 단일 잠재특성 점수·개인 진단점수·공식 절단점의 재검증 결과·생애주기 간 직접 비교점수로 사용하는 것과, 각 하위점수를 서로 독립된 개인별 위험등급으로 사용하는 것에는 추가적인 외부준거, 판별타당성, 분류정확도와 측정동일성 검증이 필요하다. 특히 하위점수에 대한 이러한 제한은 높은 요인상관과 HTMT에 기초한 간접적 판단이며, 독립성 자체를 직접 검증한 결과로 해석해서는 안 된다. 연령정보가 없는 혼합연령 표본이므로 이 해석을 모든 생애주기 집단에 자동으로 일반화해서도 안 된다. 본 연구의 기여는 원 개발연구를 대체하는 새로운 타당화 체계를 제시한 것이 아니라, 제도화된 국가척도에서 구조적 상대비교, 문항·점수 일관성, robust 절대적합과 판별가능성을 구분하여 실제 점수 사용의 해석 범위 및 조건을 명확히 했다는 데 있다.

참고문헌

과학기술정보통신부, 한국지능정보사회진흥원 (2023). 2022년 스마트폰 과의존 실태조사 보고서. 한국지능정보사회진흥원.

권선중, 엄나래 (2016). 스마트폰 과의존 척도 개편 최종보고서. 한국정보화진흥원(편), 2016년 정보문화포럼 정책연구보고서 (pp. 253-281). 한국정보화진흥원.

박정혜 (2019). 스마트폰 과의존 위험군 청소년의 스마트폰 사용행태와 의존에 대한 자기인식. Journal of The Korean Data Analysis Society, 21(4), 2095-2105. https://doi.org/10.37727/jkdas.2019.21.4.2095

유승엽 (2026). 스마트폰 과의존이 대학생의 학업 지연행동에 미치는 영향: 자기효능감의 매개효과. 미래기술융합논문지, 5(1), 181-191. https://doi.org/10.23152/JATC.2026.05.01.181

이문수, 정여주, 김종민, 연규진, 김재훈, 김미림, 김승수, 류지원, 차정인 (2024). 청소년 스마트폰 과의존 척도-R 신규 개발 및 타당화. 상담학연구, 25(4), 117-138. https://doi.org/10.15703/kjc.25.4.202408.117

이현엽, 강성록, 현승주, 최윤영, 고재원 (2023). 병사 스마트폰 과의존 척도 요인구조 분석 및 개인적·조직적 보호 요인. 한국군사학논집, 79(3), 45-72. https://doi.org/10.31066/kjmas.2023.79.3.003

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares

robust estimation techniques for ordinal data. *Structural Equation Modeling, 21*(3), 425–438. https://doi.org/10.1080/10705511.2014.915373

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior, 97,* 60–66. https://doi.org/10.1016/j.chb.2019.03.006

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43,* 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kardefelt-Winther, D. (2014). A conceptual and methodological critique of Internet addiction research: Towards a model of compensatory Internet use. *Computers in Human Behavior, 31,* 351–354. https://doi.org/10.1016/j.chb.2013.10.059

Panova, T., & Carbonell, X. (2018). Is smartphone addiction really an addiction? *Journal of Behavioral Addictions, 7*(2), 252–259. https://doi.org/10.1556/2006.7.2018.49

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour, 5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in multivariate statistical analysis: A Festschrift for Heinz Neudecker* (pp. 233–247). Kluwer Academic Publishers.

투고일자: 2026. 8. 22.
심사일자: 2026. 10. 1.
게재확정일자: 2026. 10. 2.

미래사회
Journal of Future Society
2026. Vol. 17, No. 3, 23～38.
https://doi.org/10.22987/jifso.2026.17.3.23

Measurement Structure and Score Interpretation of the Smartphone Overdependence Scale

A Follow-up Psychometric Evidence Review of the 2022 National Survey Self-Report Data

Shuaipeng Chen

Kookmin University

This study conducted a follow-up psychometric review of the 10-item self-report Smartphone Overdependence Scale used in Korea's 2022 national survey, focusing on its measurement structure and the defensible interpretation of total and domain scores. The analytic sample included 23,097 participants with complete responses. Three prespecified structures—a one-factor model, a two-factor model, and a correlated three-factor model of self-control failure, salience, and problematic consequences—were compared using weighted least squares with mean and variance adjustment (WLSMV) for four-category ordinal items. Corrected item-total correlations ranged from .635 to .715 and standardized loadings from .770 to .887. While the correlated three-factor model had the strongest relative support, absolute fit remained limited (robust comparative fit index (CFI)=.952, Tucker-Lewis index (TLI)=.933, root mean square error of approximation (RMSEA)=.106), with high factor correlations (.742-.880) and heterotrait-monotrait ratio (HTMT)=.887 between salience and problematic consequences. The 10-item total showed high internal consistency (Cronbach α=.910; ordinal α=.937). However, these results do not establish strict unidimensionality, diagnostic validity, or the validity of official classification cutoffs. Thus, the total score is best treated as a descriptive composite. High factor correlations and HTMT indirectly suggest that the three domain scores should not be interpreted as strongly independent; they are better suited for use as content profiles. Finally, age information was unavailable, which prevented life-stage invariance testing.

Key words: Smartphone Overdependence, Measurement Structure, Score Interpretation, WLSMV, National Survey

부록

ULS bifactor 보조 민감도 진단

추정체계가 다른 분석이 본문의 핵심 결론 형성에 개입하지 않도록 bifactor 결과는 부록에만 제시하였다. 자기보고 10문항 완전응답자(N=23,097)의 4범주 응답에서 polychoric 상관행렬을 구성하고, 모든 문항이 일반요인에 적재되면서 각 문항이 해당 내용영역(CF, SA, PR)의 특정요인에 추가로 적재되는 직교 bifactor 구조를 비가중최소제곱법(unweighted least squares, ULS)으로 적합하였다. 일반요인과 특정요인, 그리고 특정요인 상호 간 상관은 0으로 고정하였고 별도의 회전은 적용하지 않았다. 보조 재검산은 Python 3.13과 SciPy 1.17.0의 비선형 최소제곱 최적화로 수행했으며, polychoric 상관의 비대각 원소와 모형암시 상관 간 잔차제곱합을 최소화하였다. 여러 초기값에서 동일한 해로 수렴했고 모든 잔차분산은 0 이상이었다. 아래 지표는 WLSMV M1~M3의 적합도 순위나 연구문제 1·2의 직접 근거로 사용하지 않고, 추정방법이 다른 부록의 보조적 참고 정보로만 별도로 제시하였다.

<표 A1> ULS bifactor 민감도 진단

지표	값	해석 범위
ECV	.775	공통분산 중 일반요인 비중
총점 ω	.958	총점의 모형기반 신뢰도
ωH	.887	총점 신뢰분산 중 일반요인 기여
ωHS−CF	.262	일반요인 통제 후 CF의 고유 신뢰정보
ωHS−SA	.025	일반요인 통제 후 SA의 고유 신뢰정보
ωHS−PR	.202	일반요인 통제 후 PR의 고유 신뢰정보

주: 이 표는 본문의 WLSMV 주분석과 추정방법이 다른 보조 민감도 결과이다. 공통분산 설명비율(explained common variance, ECV)=.775, ωH=.887, ωHS=.025~.262의 패턴은 공통요인의 기여가 크고 하위영역의 독립적 신뢰정보가 제한적이라는 방향을 보여주지만, 본문 결론은 이 값들에 의존하지 않는다.