Check for updates

Development and Application of User Experience Evaluation indicators for OTT Services

Seoha Park[1], Hyo-Jin Kang[2*]

[1]Department of Future Convergence Technology Engineering, Master Graduate, Graduate School of Sungshin Women's University, Seoul, Korea

[2]Department of Service Design Engineering, Associate Professor, Sungshin Women's University, Seoul, Korea

Abstract

Background As competition among over-the-top (OTT) services intensifies, user experience (UX) has become a decisive factor in user retention and the long-term sustainability of services. Prior OTT research, however, has largely centered on isolated constructs (technology acceptance, satisfaction, and continuance intention), which constrains its capacity to capture the contextual, multidimensional nature of actual OTT use. Because OTT services are subscription-based, content-centered platforms, an integrated UX evaluation framework that jointly addresses usability, experiential value, and long-term acceptance is required.

Methods This study develops and validates an integrated UX evaluation indicator system tailored to the OTT context. A conceptual framework, UVA (Usability–User Value–User Acceptance) was established by synthesizing the shared cognitive–affective–behavioral structure underlying the technology acceptance model (TAM), the ABC model, and the Stimulus–Organism–Response (SOR) framework. Indicators were derived through a four-stage process (literature integration, user observation and interviews, expert review, and survey-based validation) producing a system of three dimensions, ten main indicators, and twenty-five sub-indicators. Using survey responses from 347 OTT users, exploratory factor analysis (EFA) was conducted to examine the structural validity of the usability and user value dimensions. The user acceptance dimension (satisfaction, stickiness, word-of-mouth intention) was excluded from the EFA because each sub-indicator was measured with a single item. The validated indicators were then applied to four OTT services, and between-service UX differences were compared using nonparametric statistics (Friedman and Mann–Whitney U).

Results The EFA showed that the usability and user value indicators formed theoretically coherent factor structures, and reliability analyses indicated acceptable-to-good internal consistency. Ease of contents navigation and contents localization quality, both surfaced through user observation rather than the literature, were confirmed as critical UX elements in actual OTT use. The small-scale pilot application revealed statistically significant between-service differences in customization flexibility, contents diversity, contents recency, social relatedness, stickiness, and word-of-mouth intention, indicating the preliminary discriminative potential of the indicator system.

Conclusions This study proposes an OTT-specific, multidimensional UX evaluation indicator system that reflects the contextual nature of OTT experiences. Although EFA-based validation was limited to the usability and user value dimensions, the findings provide empirical support for the structural coherence of the proposed indicators. By extending the results into usability-centered design guidelines, the study positions UX evaluation not merely as a measurement instrument but as a practical framework for service diagnosis, planning, and design improvement.

Keywords OTT Service, UX Evaluation Indicators, UVA Framework, Exploratory Factor Analysis, User Observation, Design Guideline

http://dx.doi.org/10.15187/adr.2026.11.39.4.299

Received : Feb. 23. 2026 ; **Reviewed :** Aug. 20. 2026 ; **Accepted :** Aug. 21. 2026
pISSN 1226-8046
eISSN 2288-2987

Copyright : This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted educational and non-commercial use, provided the original work is properly cited.

This research was expanded from the author's master's thesis, "Development and Application of UX Evaluation Metrics for OTT Services" (Master's thesis), Sungshin Women's University, Seoul, Korea. This paper was partly supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea(NRF-2023S1A5A8076896) and Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government(MOTIE) (RS-2025-02263062, HRD Program for Industrial Innovation.)

*Corresponding author: Hyo-Jin Kang (hjkang@sungshin.ac.kr)

Citation: Park, S., & Kang, H.-J. (2026). Development and Application of User Experience Evaluation indicators for OTT Services. *Archives of Design Research, 39*(4), 299-326.

1. 연구의 배경 및 목적

OTT(Over-the-Top) 서비스는 사용자가 원하는 시간과 장소에서 다양한 디바이스를 통해 콘텐츠를 소비할 수 있도록 하며, 기존 방송 중심의 미디어 소비 구조를 근본적으로 변화시켜 왔다. 특히 글로벌 OTT 플랫폼의 확산은 콘텐츠 유통의 국경을 허물고, 개인화 추천 알고리즘과 구독 기반 서비스 모델을 통해 사용자의 일상 속 미디어 소비 방식을 재구성하고 있다. 이러한 환경 변화 속에서 OTT 서비스 간 경쟁은 콘텐츠 확보나 기능적 완성도를 넘어, 사용자 경험(User Experience, UX)을 중심으로 심화되고 있다.

그러나 기존 OTT 서비스 관련 연구는 주로 기술수용모델(Technology Acceptance Model, TAM)을 기반으로 사용 의도, 만족도, 지속사용의도와 같은 개별 변인을 중심으로 수행되어 왔다. 이러한 접근은 서비스에 대한 사용자의 수용 여부를 설명하는 데에는 유효하나, 실제 서비스 이용 과정에서 형성되는 복합적이고 맥락적인 사용자 경험을 구조적으로 평가하는 데에는 한계를 지닌다. 특히 OTT 서비스는 콘텐츠 탐색 과정, 추천 경험, 시청 환경, 자막 및 번역 품질과 같이 다층적인 경험 요소가 상호작용하며 사용자 경험을 구성함에도 불구하고, 이를 종합적으로 반영한 통합적 평가 지표 체계는 충분히 정립되지 않았다.

또한 OTT 서비스는 일회적 사용이 아닌 장기적·반복적 이용을 전제로 하는 구독 기반 서비스로서, 단기적 사용성 평가만으로는 사용자 경험의 전체 구조를 설명하기 어렵다. 이는 사용자 경험을 기능적 사용성 차원에 국한하지 않고, 콘텐츠 소비 과정에서 인식되는 가치와 서비스와의 장기적 관계 형성까지 포함하는 확장된 평가 틀이 필요함을 시사한다.

이에 본 연구는 OTT 서비스 이용 과정에서 발생하는 인지적·정서적·행동적 경험을 통합적으로 반영할 수 있는 사용자 경험 평가 지표 체계를 개발하는 것을 목적으로 한다. 이를 위해 기술수용모델(TAM), 정서–행동–인지(ABC) 모델, 자극–유기체–반응(SOR) 모형의 공통 구조를 통합한 UVA(Usability–User Value–User Acceptance) 프레임워크를 상위 분석 틀로 설정하고, 문헌 연구와 사용자 탐색 연구를 통해 OTT 서비스 맥락에 적합한 UX 평가 지표를 도출·검증·적용한다. 나아가 본 연구는 개발된 평가 지표를 기반으로 실무에서 활용 가능한 UX 디자인 가이드라인을 제시함으로써, OTT 서비스 UX 평가를 단순한 측정 도구를 넘어 설계 및 개선을 위한 실천적 도구로 확장하고자 한다.

이러한 연구 목적을 구체화하기 위해 본 연구는 지표 개발, 구조적 타당성 검증, 서비스 간 적용, 실무적 활용에 이르는 다음의 연구 질문(Research Questions)을 설정하였다.

RQ1. OTT 서비스 맥락에서 사용성, 사용자 가치, 사용자 수용도를 포괄하는 통합적 UX 평가 지표 체계는 무엇인가?

RQ2. 제안된 UX 평가 지표는 UVA 상위 차원에 부합하는 요인 구조를 형성하는가?

RQ3. 제안된 UX 평가 지표는 실제 OTT 서비스 간 사용자 경험 차이를 유의하게 구분해낼 수 있는가?

RQ4. 제안된 UX 평가 지표는 서비스 설계 및 개선에 활용 가능한 UX 디자인 가이드라인으로 확장될 수 있는가?

2. 이론적 배경

2. 1. 맥락적·관계적 품질로서의 사용성 평가

사용성(Usability)은 인간–컴퓨터 상호작용(HCI) 분야에서 핵심 개념으로, 전통적으로 사용자가 시스템을 통해 목표를 얼마나 효과적·효율적·만족스럽게 달성할 수 있는지를 나타내는 품질 속성으로 정의되어 왔다. 초기

연구에서는 사용성을 시스템 자체의 고유한 특성으로 이해하는 경향이 강했으나, 이후 연구의 진전에 따라 특정 사용자, 과업, 환경 간 상호작용 속에서 형성되는 맥락적 개념으로 확장되어 왔다.

Shackel(1981/2009)은 사용성을 특정 사용자 집단이 특정 과업을 수행하고 특정 환경에 놓이는 조건 속에서 발현되는 인간 기능적 능력으로 규정하며, 사용성을 시스템의 내재적 속성이 아닌 관계적 결과로 보았다. 이러한 관점은 ISO 9241-11(ISO, 2018)에서 '지정된 사용자가 지정된 사용 맥락에서 효과성, 효율성, 만족도를 가지고 목표를 달성할 수 있는 정도'로 표준화되었으며, Bevan과 Macleod(1994)는 이를 '맥락 속 사용 품질(quality of use)'로 개념화하여 사용성이 사용 경험 속에서 평가되어야 함을 강조하였다.

한편 Nielsen(1994)은 실무적 관점에서 사용성을 학습성, 효율성, 기억 용이성, 오류 관리, 만족도의 다차원적 구성 요소로 제시함으로써, 인터페이스 설계 및 사용성 평가에서의 적용 가능성을 확장하였다. 이후 연구들 또한 사용성을 사용자, 과업, 환경 간 상호작용 속에서 창발적으로 나타나는 경험적 품질로 규정하며, 사용성을 정태적 성능 지표가 아닌 경험적 결과로 이해해야 한다는 점을 공통적으로 제시하고 있다.

그러나 이러한 이론적 확장에도 불구하고, 전통적인 사용성 평가는 여전히 과업 수행 여부, 반응 속도, 오류 발생과 같은 기능적·행동적 지표에 초점을 두는 경향이 강하다. 이로 인해 즐거움, 몰입, 신뢰와 같은 정서적 반응이나 장기적 이용 맥락에서 형성되는 경험적 가치를 충분히 설명하는 데에는 한계가 존재한다. 특히 구독 기반으로 지속적 이용이 전제되는 OTT 서비스 환경에서는 사용성만으로 사용자 경험의 전체 구조를 포착하기 어렵다.

이러한 맥락에서 사용성은 사용자 경험을 구성하는 핵심적인 기초 요소이지만, 그 자체로는 충분하지 않은 부분적 차원으로 기능한다. 즉, 사용성은 주로 인지적·기능적 상호작용 경험을 설명하는 역할을 수행하며, 사용자에게 인식되는 가치나 장기적 수용 및 행동으로의 연결을 설명하기 위해서는 보다 확장된 평가 틀이 요구된다. 이는 사용성을 사용자 가치와 사용자 수용도의 차원과 함께 통합적으로 구조화할 필요성을 시사하며, 본 연구에서 제안하는 OTT 서비스 사용자 경험 평가를 위한 이론적 프레임워크의 이론적 출발점이 된다.

2. 2. UVA(Usability-Value-Acceptance) 프레임워크

사용자 경험을 체계적으로 평가하기 위한 이론적 프레임워크는 크게 두 흐름으로 전개되어 왔다. 하나는 시스템 수용 및 성과 중심으로 사용자 행동을 설명하는 모델(Davis, 1989; DeLone & McLean, 2003; Venkatesh, Thong, & Xu, 2012)이고, 다른 하나는 사용자 경험을 인지적·정서적·행동적 과정의 복합 구조로 이해하는 모델(Mehrabian & Russell, 1974; Rosenberg & Hovland, 1960)이다. 그러나 기존 대표적 프레임워크들은 콘텐츠 탐색, 개인화 추천, 장기 구독을 핵심으로 하는 OTT 서비스 맥락에서 일정한 한계를 지닌다.

DeLone과 McLean(2003)의 정보시스템 성공 모델(IS Success Model)은 정보품질, 시스템품질, 서비스품질이 사용자 만족과 시스템 사용을 거쳐 순효익(Net Benefit)으로 이어지는 구조를 제시한다. 그러나 이 모델은 콘텐츠 소비 과정에서 형성되는 즐거움, 몰입, 심미성, 사회적 연결감과 같은 경험적 가치를 독립적인 평가 차원으로 충분히 설명하지 못한다. Venkatesh 등(2012)의 UTAUT2(the Extended Unified Theory of Acceptance and Use of Technology) 역시 성과 기대, 노력 기대, 사회적 영향, 촉진 조건 등을 통해 사용 의도를 설명하는 데 강점이 있으나, 콘텐츠 품질에 대한 경험적 가치 인식이나 구독 기반 서비스에서 형성되는 장기적 관계 경험을 세밀하게 포착하는 데에는 한계가 있다. 특히 OTT 서비스는 초기 수용 이후 반복적 시청, 구독 유지, 추천 행동이 중요하므로, 수용 의도 중심 모델만으로는 실제 이용 과정에서 발생하는 다차원적 UX 구조를 충분히 설명하기 어렵다.

이에 본 연구는 기술수용모델(TAM), 정서–행동–인지(ABC) 모델, 자극–유기체–반응(SOR) 모형을 검토하였다. 이 세 모델은 각각 정보시스템, 사회심리학, 소비자 행동이라는 서로 다른 학문 분야에서 제안된 이론으로, 원래 OTT 서비스 UX 평가를 목적으로 고안된 것은 아니다. 그러나 이 모델들을 면밀히 살펴보면, 각각이 이용자의 경험을 인지적 평가→정서적 반응→행동적 결과의 유사한 단계로 포착하고 있음을 확인할 수 있다. 이들을 통합함으로써, 단일 분야 모델의 한계를 넘어 OTT 서비스 UX 평가에 적합한 다차원적 이론 틀로 발전시킬 수 있다.

기술수용모델(TAM)은 지각된 용이성(Perceived Ease of Use)과 지각된 유용성(Perceived Usefulness)이 사용자의 태도와 행동 의도에 영향을 미친다고 설명한다(Davis, 1989). OTT 서비스 맥락에서 이는 콘텐츠 탐색 구조의 직관성, 기능 조작의 용이성, 정보 인지의 명확성 등 사용자가 서비스를 이해하고 사용할 수 있는 정도로 구체화된다.

ABC 모델은 태도가 인지적(Cognitive), 정서적(Affective), 행동적(Behavioral) 요소의 상호작용을 통해 형성된다고 보며(Rosenberg & Hovland, 1960), 사용자 경험이 단일 차원이 아닌 다차원적 구조임을 강조한다. OTT 서비스에서 사용자는 콘텐츠 품질, 추천 적합성, 인터페이스 심미성, 몰입감 등을 종합적으로 경험하며, 이러한 경험은 서비스에 대한 가치 인식으로 이어진다.

SOR 모형은 서비스 환경이라는 외부 자극이 사용자의 인지·정서적 상태를 변화시키고, 그 결과로 행동적 반응이 나타나는 구조를 제시한다(Mehrabian & Russell, 1974). OTT 서비스 맥락에서 서비스 기능, 콘텐츠 구성, 추천 알고리즘, 사회적 기능은 자극(Stimulus)으로 작용하고, 사용자의 만족, 몰입, 신뢰와 같은 내적 반응(Organism)을 거쳐 지속 이용 의도와 추천 의도(Response)로 이어진다.

단일 모델이나 기존 IS Success Model·UTAUT2만으로는 OTT 서비스 이용자가 콘텐츠 탐색·개인화·반복 구독 과정에서 경험하는 다차원적 UX 구조를 충분히 포착할 수 없다. UVA 프레임워크는 이러한 한계를 보완하기 위해, [Figure 1]과 같이 세 모델이 공통적으로 전제하는 인지–정서–행동의 순차적 흐름을 OTT 서비스 맥락에 맞게 사용성–사용자 가치–사용자 수용도의 3차원 평가 구조로 재구성한 것이다.

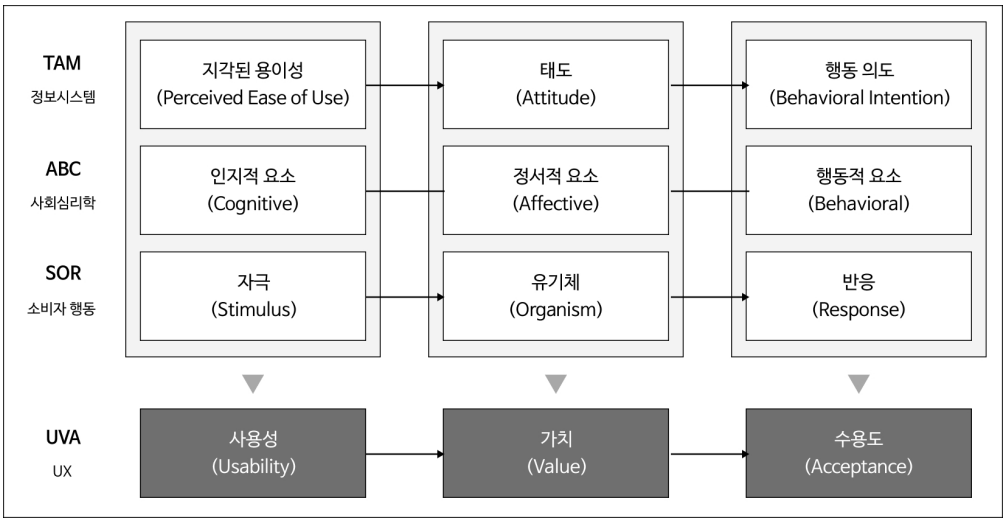

Figure 1 Integration of TAM, ABC, and SOR models into the UVA framework for OTT service UX evaluation

UVA 프레임워크에서 사용성은 인터페이스와 상호작용 과정에서 경험되는 인지적·기능적 측면을 의미하며, 사용자 가치는 콘텐츠 품질, 개인화 적합성, 정서적 만족과 같은 경험의 질적 평가를 포괄한다. 사용자 수용도는 만족도, 지속성, 추천 의도 등 서비스와의 장기적 관계 형성을 반영하는 행동적 결과 차원에 해당한다.

3. 연구 방법

본 연구는 OTT 서비스 사용자 경험을 체계적으로 평가하기 위한 UX 평가 지표 체계를 구축하기 위해, 개발–검증–적용–활용의 네 단계로 구성된 연구 절차를 설계하였다. 각 단계는 상호 독립적으로 수행되기보다, 이전 단계의 결과를 다음 단계에 반영하는 순환적 구조를 갖는다. 전체 연구 절차는 [Figure 2]에 제시하였다.



Figure 2 Research Flow Chart

3. 1. 사용자 경험 평가 지표 개발

3. 1. 1. 문헌 기반 UX 평가 지표 수집 및 통합

첫 번째 단계에서는 OTT 서비스 및 인접 디지털 서비스 영역에서 수행된 선행연구를 대상으로 사용자 경험 평가 지표를 수집·정리하였다. 문헌 조사는 2025년 2월 6일부터 5월 28일까지 총 112일간 Google Scholar, DBpia, RISS, arXiv 등 주요 학술 데이터베이스를 활용하여 수행하였다. 검색어는 'OTT', '사용성 평가', '사용자 경험', 'OTT UX', 'usability evaluation indicators', 'user experience evaluation', 'OTT platform', 'streaming service UX' 등을 조합하여 사용하였다. 키워드 검색을 통해 초기 문헌을 식별한 이후에는 선정된 문헌의 참고문헌 목록을 검토하여 관련 연구를 추가적으로 확보하는 연쇄 표집법(referential sampling)을 병행하였다. 이를 통해 데이터베이스 검색만으로는 포착되지 않는 이론적 원천 문헌 및 인접 영역의 핵심 연구를 보완적으로 수집하였다.

문헌 선정은 식별(identification), 선별(screening), 포함(inclusion)의 절차를 참고하여 진행하였다. 먼저 검색어 조합과 연쇄 표집을 통해 도출된 문헌의 제목, 초록, 키워드를 검토하여 사용자 경험 평가 지표 또는 사용성 구성 요소와 관련성이 있는 연구를 1차 후보로 선별하였다. 이후 전문 검토 단계에서 연구 대상, 평가 지표 제시 여부, 연구 방법의 명확성, OTT 또는 인접 디지털 서비스 맥락과의 관련성을 기준으로 최종 문헌 30편을 선정하였다. 선정 문헌은 연구 대상과 서비스 범위에 따라 세 가지 유형으로 분류하였다. 특정 도메인에 한정되지 않은 일반적인 인터랙티브 시스템의 사용자 경험·사용성을 다룬 연구는 GEN-UX 유형으로, OTT 또는 영상 스트리밍 서비스를 직접적으로 대상으로 한 연구는 OTT-UX 유형으로, OTT와 인접한 개발·서비스 환경의 UX·사용성을 다룬 연구는 DEV·SVC-UX 유형으로 구분하였다. 문헌 목록과 유형별 분류는 [Appendix 1]에 제시하였다.

지표 추출 과정에서는 각 문헌에서 제시된 평가 차원, 세부 지표, 조작적 정의, 설문 문항을 수집하였다. 이후 의미적으로 중복되거나 유사한 항목은 하나의 개념 단위로 통합하였다. 지표 통합 기준은 다음과 같다. 첫째, 동일한 사용자 경험 요소를 설명하는 지표는 더 포괄적인 상위 개념으로 통합하였다. 둘째, 명칭은 다르지만 조작적 정의가 유사한 지표는 동일 지표로 병합하였다. 셋째, OTT 서비스 이용 맥락과 직접적으로 연결되지 않거나 본 연구의 평가 목적과 관련성이 낮은 지표는 제외하였다. 넷째, 통합 또는 제외 여부가 모호한 지표는 FGI 및 전문가 검토 과정에서 논의하여 최종 반영 여부를 결정하였다.

3. 1. 2. 사용자 관찰 연구 및 전문가 검토

문헌 기반으로 도출된 초기 지표 체계가 실제 OTT 서비스 이용 맥락을 충분히 반영하는지를 검토하기 위해, OTT 서비스 이용 경험이 있는 사용자 8명을 대상으로 사용자 관찰 연구와 사후 인터뷰를 수행하였다. 참여자들은 1일간 일상적인 OTT 서비스 이용 과정을 기록하였으며, 콘텐츠 탐색 방식, 추천 경험, 시청 환경, 만족 및 불편 요소를 중심으로 경험을 서술하였다. 사후 인터뷰를 통해 기록된 경험의 맥락과 정서적 반응을 보완적으로 수집하였다.

수집된 질적 데이터는 발화 단위 분석을 통해 문헌 기반 UX 평가 지표와 대응시켜 분석되었으며, 반복적으로 언급되거나 기존 지표로 충분히 설명되지 않는 경험 요소를 중심으로 지표를 보완·확장하였다. 1차 정리된 지표를 기반으로 2025년 9월 19일, HCI 및 서비스디자인 전공 석·박사 연구원 4명이 참여한 FGI를 약 90분간 오프라인으로 실시하였다. 이후 2차 정리된 지표를 바탕으로 2025년 10월 4일부터 10월 10일까지 HCI/서비스디자인 석사 연구원 1명, 박사 연구원 1명, OTT 서비스 실무 경력자 1명, 서비스 기획자 1명 등 총 4명이 참여한 전문가 피드백을 진행하였다. 이를 통해 일부 지표의 명칭과 정의를 수정·보완하여 최종 3개 상위 차원, 10개 주지표, 25개 세부 지표의 평가 체계를 확정하였다.

3. 2. 사용자 경험 평가 지표 검증

두 번째 단계에서는 개발된 사용자 경험 평가 지표 체계의 구조적 타당성을 검증하기 위해 설문 조사를 실시하였다. 본 조사에 앞서 2025년 10월 8일부터 10월 10일까지 3일간 파일럿 테스트를 실시하였다. 파일럿 테스트에는 HCI/서비스디자인 연구원 2명, OTT 서비스 실무자 2명, 실제 OTT 사용자 3명 등 총 7명이 참여하였다. 설문은 Google Form을 활용하여 5점 리커트 척도로 구성되었으며, 문항의 이해도와 중복성을 점검한 후 일부 표현을 수정·보완하였다.

본 설문 조사는 2025년 10월 13일부터 10월 16일까지 총 4일간 진행되었으며, OTT 서비스 이용 경험이 있는 일반 사용자 347명을 대상으로 실시되었다. 설문지는 인구통계학적 문항 6개와 상세 지표 문항 25개로 구성되었다. 상세 지표 문항은 [Table 4]에 제시된 25개 세부 지표의 개념적 정의를 기반으로 구성하였다. 각 세부 지표당 1문항씩 작성하였으며, 'OTT 서비스가 [해당 지표가 측정하는 경험]을 하나요?'의 질문 형식으로 조작적으로 정의하였다. 모든 문항은 5점 리커트 척도(1 = 전혀 그렇지 않다, 5 = 매우 그렇다)로 응답하도록 구성하였으며, 전체 문항 목록은 [Appendix 2]에 제시하였다.

본조사 응답자는 OTT 서비스 이용 경험이 있는 일반 사용자로 구성되었다. [Table 1]에서 확인할 수 있듯이, 연령·성별·학력 측면에서 특정 집단에 편중되지 않은 비교적 균형 잡힌 인구통계학적 구성을 보였다.

Table 1 Survey Respondent Profiles

구분	항목				
연령	10대	20대	30대	40대	50대 이상
	8.1%(28명)	29.7%(103명)	23.3%(81명)	25.6%(89명)	13.3%(46명)
성별	여성	남성	기타		
	54.5%(189명)	45.2%(157명)	0.3%(1명)		
최종 학력	고등학교 졸업	전문학사 졸업	학사 졸업	대학원 이상	
	27.4%(95명)	12.1%(42명)	47.3%(164명)	7.2%(25명)	
디지털 친숙도	매우 친숙하지 않음	친숙하지 않음	보통	친숙함	매우 친숙함
	0.3%(1명)	2.6%(9명)	13.0%(45명)	44.7%(155명)	39.5%(137명)
OTT 사용 빈도	월 4회 미만	주 1-2회	주 3-4회	주 5-6회	매일 시청
	9.8%(34명)	18.7%(65명)	25.4%(88명)	16.1%(56명)	30.0%(104명)
사용해본 OTT (복수 응답 허용)	넷플릭스	쿠팡플레이	티빙	디즈니+	왓챠
	93.4%(324명)	71.2%(247명)	68.9%(239명)	49.3%(171명)	32.0%(111명)

응답자의 디지털 활용 능력은 전반적으로 높은 수준이었으며, OTT 사용 빈도는 주 3회 이상이 71.5%로 실제 이용 경험이 충분한 집단이었다. 사용 경험이 있는 OTT 서비스는 넷플릭스, 쿠팡플레이, 티빙이 중심을

이루었으나, 애플TV(8.6%), 아마존 프라임(2.9%), 웨이브(2.6%), HBO MAX(1.4%), 라프텔(1.4%) 등 다양한 플랫폼도 일부 포함되어, 특정 서비스에 편중되지 않은 이용 맥락을 반영하고 있었다.
수집된 응답 데이터를 기반으로 SPSS를 활용한 탐색적 요인분석(Exploratory Factor Analysis, EFA)을 수행하였다. 이를 통해 사용성 및 사용자 가치 차원의 세부 지표들이 이론적으로 설정한 상위 차원에 따라 의미적으로 일관된 요인 구조를 형성하는지를 검토하였다. 단, 사용자 수용도 차원의 만족도, 지속성, 추천 의도는 단일 문항으로 구성되어 EFA 대상에서 제외되었으므로, 본 검증은 UVA 프레임워크 전체가 아닌 사용성 및 사용자 가치 차원에 한정된다.

3. 3. 사용자 경험 평가 지표 예비 적용

세 번째 단계에서는 검증된 사용자 경험 평가 지표를 실제 OTT 서비스를 대상으로 예비 적용하였다. 평가 참여자인 HCI 및 서비스디자인 관련 석박사급 연구원 10명은 동일한 과업 조건하에서 실제 시중에서 사용되는 OTT 서비스 4종(아마존프라임, 넷플릭스, 티빙, 왓챠플레이)을 이용한 후 UX 평가 설문에 응답하였다.

3. 4. 사용자 경험 평가 지표 활용

마지막 단계에서는 UX 평가 결과를 바탕으로, 사용성 차원을 중심으로 한 UX 디자인 가이드라인을 도출하였다. 도출된 가이드라인은 UX 평가 지표를 단순한 측정 도구에 그치지 않고, 서비스 설계 및 개선 과정에서 직접 활용할 수 있도록 설계 기준의 형태로 재구성되었다.
이러한 활용 단계는 UX 평가 결과가 서비스 개선 전략으로 환류되는 구조를 갖도록 함으로써, UX 평가를 사후 검증 중심의 도구에서 기획–설계–운영 전반에 활용 가능한 실천적 도구로 확장하는 데 목적이 있다.

4. 연구 결과

본 장에서는 OTT 서비스 사용자 경험(UX) 평가 지표의 개발–검증–적용–활용과정에서 도출된 연구 결과를 단계별로 제시한다. 특히 문헌 기반 분석과 사용자 관찰 데이터를 통합하여 지표 체계를 정교화한 결과와, 검증된 지표를 실제 OTT 서비스에 적용하여 확인한 사용자 경험 특성을 중심으로 서술한다.

4. 1. UX 평가 지표 개발 결과

4. 1. 1. 문헌 기반 UX 평가 지표 통합 결과

먼저 OTT 서비스 및 인접 디지털 서비스 영역에서 수행된 선행연구를 대상으로 사용자 경험 평가 지표를 수집·정리하였다. 수집된 지표들은 의미적 유사성과 개념적 범위를 기준으로 통합되었으며, UVA 프레임워크를 상위 분류 체계로 적용하여 사용성, 사용자 가치, 사용자 수용도의 세 차원으로 재분류되었다.
이 과정에서 문헌마다 상이하게 사용되던 지표 명칭과 정의를 비교·조정함으로써, OTT 서비스 맥락에서 해석 가능한 개념 단위로 정제하였다. 의미적으로 중복되거나 동일한 경험을 설명하는 지표들은 통합되었으며, OTT 서비스 특성과 직접적으로 연결되지 않는 지표는 제외하였다. 문헌 기반 UX 평가 지표의 통합 결과와 분류 체계는 [Table 2]에 제시하였다.

Table 2 Literature-Based Integration of Initial UX Evaluation indicators

구분	주지표	상세지표	문헌 연구 도출 지표
사용성 (Usability)	사용편의성 (Ease of Use)	인지용이성 (Ease of Perception)	지각성, 인지적 용이성, 이해가능성, 명료성, 가독성, 지각 한계, 최소 기억 부하
		조작용이성 (Ease of Operation)	최소 조작, 사용의 용이성, 조작성, 통제와 자유, 매핑, 사용자 중심 설계
	인터랙션 적합성 (Interaction Suitability)	원활성 (Seamlessness)	일관성, 단순성 및 최소주의, 복구 가능성, 오류 수용성
		지원성 (Supportiveness)	사용자 안내, 도움성, 피드백
		정보전달성 (Informativeness)	정보전달성, 의사소통성, 주목성, 정확성
	상황 인지 특성 (Context-Awareness)	콘텐츠 개인화 (Contents Personalization)	개인화, 개인 맞춤성
		서비스 개인화 (Service Personalization)	사용자 태도 기반 적응, 서비스 맥락 적응
		사용자 맞춤형 유연성 (Customization Flexibility)	사용의 유연성, 맞춤화 유연성, 적응성
	콘텐츠 품질 (Contents Quality)	콘텐츠 다양성 (Contents Diversity)	다양성, 콘텐츠 다양성
		콘텐츠 최신성 (Contents Recency)	최신성
		콘텐츠 화질 (Contents Video Quality)	영상 품질, 화질, 3D 품질
		콘텐츠 음질 (Contents Audio Quality)	음질, 음성 품질
사용자 가치 (User Value)	기능적 가치 (Functional Value)	신뢰성 (Trustworthiness)	신뢰성
		안정성 (Stability)	안정성
		반응성 (Responsiveness)	반응성
		보안 (Safety)	안전성, 보안성, 프라이버시 보호
	감정적 가치 (Emotional Value)	즐거움 (Pleasure)	즐거움, 유희성, 오락성, 주관적 즐거움, 정서적 경험, 정서적 미디어 만족, 현실 도피, 시간 보내기
		몰입성 (Immersion)	몰입성, 몰입 성향, 감각적 경험
		심미성 (Aesthetics)	심미성, 지각된 심미성, 디자인, 조형성
	사회적 가치 (Social Value)	사회적 관계성 (Social Relatedness)	사회적 연결감, 사용자 간 상호작용 개념
사용자 수용도 (User Acceptance)	만족도 (Satisfaction)		만족도, 사용자 만족, 만족성
	지속성 (Stickiness)		지속 사용 의도, 이용 의도, 지속성
	추천 의도 (Word-of-Mouth Intention)		충성도, 구매 후 의도

4. 1. 2. 사용자 관찰 데이터 기반 초기 UX 지표 언급 분석 결과

문헌 기반으로 도출된 초기 UX 평가 지표 체계의 실제 사용 맥락 적합성을 검토하기 위해 사용자 관찰 연구를 수행하였다. 참여자들은 1일간 일상적인 OTT 서비스 이용 과정에서의 콘텐츠 탐색 행태, 추천 경험, 시청 환경, 만족 및 불편 요소를 기록하였으며, 사후 인터뷰를 통해 경험의 맥락과 정서적 반응을 보완적으로 수집하였다. 수집된 질적 데이터는 발화 단위 분석을 통해 초기 UX 평가 지표와 대응시켜 분석되었다. 사용자 관찰 데이터 기반 지표 언급 분석 결과는 [Table 3]에 제시하였으며, 각 지표가 실제 사용자 경험에서 어느 참여자에게서 언급되었는지와 어떤 맥락에서 등장하는지를 확인하였다.

Table 3 User Observation-Based Analysis of Initial UX indicator Mentions

카테고리	주지표	상세지표	User1	User2	User3	User4	User5	User6	User7	User8
사용성 (Usability)	사용편의성 (Ease of Use)	인지용이성 (Ease of perception)	–	O	O	O	O	–	O	O
		조작용이성 (Ease of operation)	O	O	O	–	O	O	O	O
		*탐색용이성 (Ease of navigation)	O	O	–	O	O	O	–	O
	인터랙션 적합성 (Interaction Suitability)	원활성 (Seamlessness)	O	O	O	O	O	O	O	O
		지원성 (Supportiveness)	O	O	O	O	O	O	O	O
		정보전달성 (Informativeness)	–	O	O	–	–	O	–	O
	상황 인지 특성 (Context-Awareness)	콘텐츠 개인화 (Contents Personalization)	O	O	O	O	O	O	O	O
		서비스 개인화 (Service Personalization)	O	O	O	O	O	O	O	O
		사용자 맞춤형 유연성 (Customization Flexibility)	O	O	O	–	O	–	O	O
	콘텐츠 품질 (Contents Quality)	콘텐츠 다양성 (Contents Diversity)	O	O	O	O	O	O	O	O
		콘텐츠 최신성 (Contents Recency)	O	O	–	–	–	–	O	O
		콘텐츠 화질 (Contents Video Quality)	O	O	O	O	O	O	O	O
		콘텐츠 음질 (Contents Audio Quality)	O	O	–	–	O	–	O	O
		*콘텐츠 번역 품질 (Contents Localization Quality)	–	O	O	O	O	–	O	O
사용자 가치 (User Value)	기능적 가치 (Functional Value)	신뢰성 (Trustworthiness)	O	O	–	–	O	O	O	O
		안정성 (Stability)	O	O	O	O	O	O	O	O
		반응성 (Responsiveness)	–	O	O	–	–	O	–	O
		보안 (Safety)	O	O	O	O	O	O	O	O
	감정적 가치 (Emotional Value)	즐거움 (Pleasure)	O	O	O	O	O	O	O	O
		몰입성 (Immersion)	O	O	O	O	O	O	O	O
		심미성 (Aesthetics)	O	O	O	O	–	O	–	–
	사회적 가치 (Social Value)	사회적 관계성 (Social relatedness)	O	O	O	O	O	O	O	O
사용자 수용도 (User Acceptance)	만족도 (Satisfaction)		O	O	O	O	O	O	O	O
	지속성 (Stickiness)		O	O	O	O	O	O	O	O
	추천 의도 (Word-of-Mouth Intention)		O	O	O	O	O	O	O	O

분석 결과, 일부 경험 요소는 문헌 기반 지표 체계에서 충분히 독립적인 지표로 다루어지지 않았음에도 불구하고, 실제 사용자 경험에서는 반복적으로 강조되는 것으로 나타났다.

특히 콘텐츠 탐색 용이성과 콘텐츠 번역 품질은 사용자 관찰 데이터에서 핵심적인 사용자 경험 요소로 빈번하게 언급되었으며, 이는 기존 문헌 기반 지표 체계의 보완 필요성을 시사하였다. 해당 결과는 문헌 중심 UX 평가 지표가 실제 OTT 서비스 이용 경험의 맥락을 충분히 반영하지 못할 수 있음을 보여준다.

4. 1. 3. 최종 OTT 서비스 UX 평가 지표 체계 도출

문헌 기반 지표 통합 결과와 사용자 관찰 데이터 분석 결과를 종합한 후, 전문가 검토를 통해 UX 평가 지표 체계를 최종적으로 정제하였다. 전문가 검토 과정에서는 지표 간 위계 구조, 주지표와 세부 지표 간 관계, 지표 명칭과 정의의 명확성, 그리고 OTT 서비스 맥락과의 적합성이 중점적으로 논의되었다.

Table 4 Derivation of an OTT Service User Experience Evaluation indicator System

사용성 Usability		
U.1. 사용편의성 Ease of Use		OTT 서비스가 사용자가 쉽게 이해하고 사용할 수 있게 설계된 정도
	U.1.1. 인지용이성 Ease of perception	OTT 서비스를 통해 사용자가 OTT 서비스의 사용 방식과 과정을 쉽게 인지하고 이해할 수 있는 정도
	U.1.2. 조작용이성 Ease of operation	OTT 서비스를 통해 사용자가 OTT 서비스의 사용 방식과 과정을 물리적으로 쉽게 조작할 수 있는 정도
	U.1.3. 탐색용이성 Ease of navigation	OTT 서비스를 통해 사용자가 원하는 콘텐츠를 얼마나 쉽고 빠르게 찾을 수 있는 정도
U.2. 인터랙션 적합성 Interaction Suitability		OTT 서비스가 사용자와의 상호작용 과정에서 원활한 커뮤니케이션을 제공하는 정도
	U.2.1. 원활성 Seamlessness	OTT 서비스가 과업 간, 기기 간, 다중 모드에서 끊김 없고 자연스러운 사용자 경험을 제공하는 정도
	U.2.2. 지원성 Supportiveness	OTT 서비스가 서비스 사용 중 발생한 문제를 돕고, 순조로운 사용을 위한 정보 및 기능을 지원하는 정도
	U.2.3. 정보전달성 Informativeness	OTT 서비스가 모든 관련 정보를 명확하고 효과적으로 전달하는 정도
U.3. 상황 인지 특성 Context-Awareness		OTT 서비스가 사용자의 상황과 선호를 인식하여 개인화된 콘텐츠와 경험을 제공하는 정도
	U.3.1. 콘텐츠 개인화 Contents Personalization	OTT 서비스가 사용자의 취향과 니즈에 따라 콘텐츠를 개인화하여 제공하는 정도
	U.3.2. 서비스 개인화 Service Personalization	OTT 서비스가 사용자에 맞춰 편리하게 개인화된 정도
	U.3.3. 사용자 맞춤형 유연성 Customization Flexibility	OTT 서비스가 사용자가 자신이 원하는 대로 설정할 수 있는 정도
U.4. 콘텐츠 품질 Contents Quality		OTT 서비스가 제공하는 콘텐츠의 시청 품질과 다양성을 유지하며 몰입감 있는 경험을 제공하는 정도
	U.4.1. 콘텐츠 다양성 Contents Diversity	OTT 서비스가 제공하는 콘텐츠가 다양한 정도
	U.4.2. 콘텐츠 최신성 Contents Recency	OTT 서비스가 제공하는 콘텐츠가 최신 콘텐츠를 제공하는 정도
	U.4.3. 콘텐츠 화질 Contents Video Quality	OTT 서비스가 제공하는 콘텐츠의 화질이 뛰어난 정도
	U.4.4. 콘텐츠 음질 Contents Audio Quality	OTT 서비스가 제공하는 콘텐츠의 음질이 뛰어난 정도
	U.4.5. 콘텐츠 번역 품질 Contents Localization Quality	OTT 서비스의 다국어 콘텐츠의 자막 번역이 자연스럽고 정확한 정도
사용자 가치 User Value		
V.1. 기능적 가치 Functional Value		OTT 서비스가 안정적이고 신뢰할 수 있는 기술적 환경을 통해 사용자의 서비스 이용을 효과적으로 지원하는 정도
	V.1.1. 신뢰성 Trustworthiness	OTT 서비스의 전반적인 기능이 신뢰가 가는 정도

	V.1.2. 안정성 Stability	OTT 서비스가 안정적으로 작동하는 정도
	V.1.3. 반응성 Responsiveness	OTT 서비스의 전반적인 반응 속도가 빠른 정도
	V.1.4. 보안 Safety	OTT 서비스가 사용자의 개인 정보를 안전하게 관리하고 프라이버시를 보장하는 정도
V.2. 감정적 가치 Emotional Value		OTT 서비스 이용 과정에서 사용자가 느끼는 유희성, 몰입감, 심미적 만족 등 정서적 경험의 가치
	V.2.1. 즐거움 Pleasure	OTT 서비스의 전반적인 사용경험이 사용자에게 유희나 즐거움을 제공하는 정도
	V.2.2. 몰입성 Immersion	OTT 서비스를 몰입하여 사용하면서 시간을 보내고 지루함을 해소하는 정도
	V.2.3. 심미성 Aesthetics	OTT 서비스의 전반적인 인터페이스가 사용자에게 심미적인 만족감을 주는 정도
V.3. 사회적 가치 Social Value		OTT 서비스 사용을 통해 타인과 사회적 연결감을 형성하는 정도
	V.3.1. 사회적 관계성 Social Relatedness	OTT 서비스 사용을 통해 타인과의 상호작용에 대한 만족감을 주는 정도
사용자 수용도 User Acceptance		
A.1. 만족도 Satisfaction		OTT 서비스의 전반적인 사용 경험에 대해 만족하는 정도
A.2. 지속성 Stickiness		OTT 서비스를 장기적으로 이용·유지하고 시청할 의향
A.3. 추천 의도 Word-of-Mouth Intention		OTT 서비스를 다른 사람에게 추천할 의향

그 결과 일부 지표는 통합·수정되었으며, 최종적으로 UVA 프레임워크를 상위 구조로 하는 3개 차원, 10개 주지표, 25개 세부 지표로 구성된 OTT 서비스 사용자 경험 평가 지표 체계가 도출되었다. 최종 UX 평가 지표 체계는 [Table 4]에 제시하였다.

4. 2. UX 평가 지표 검증 결과

최종 확정된 25개 세부 지표를 기반으로 구성된 설문지의 응답 데이터를 활용하여, 개발된 UX 평가 지표 체계의 구조적 타당성을 검증하기 위해 탐색적 요인분석(EFA)을 수행하였다. 분석은 SPSS를 활용하였으며, 추출 방법은 주성분 분석(Principal Component Analysis), 회전 방법은 카이저 정규화가 적용된 베리맥스(Varimax) 직교 회전 방식을 사용하였다. Varimax 직교 회전을 선택한 이유는 본 연구의 UVA 프레임워크에서 각 주지표가 개념적으로 구별되는 경험 요소를 측정하도록 설계되었기 때문이다. 요인 간 독립성이 이론적으로 전제될 경우 직교 회전은 각 요인의 해석 가능성을 높이는 장점이 있다.

본 연구의 UVA 이론 구조는 사용성(Usability), 사용자 가치(User Value), 사용자 수용도(User Acceptance)의 단계적 관계를 전제로 한다. 각 단계의 세부 지표는 서로 다른 경험 차원을 설명하도록 설계되었기 때문에, 모든 지표를 통합하여 분석할 경우 요인 간 교차적재나 구조적 혼동이 발생할 가능성이 존재하였다. 이러한 이론적·구조적 특성을 반영하여 사용성 지표군과 사용자 가치 지표군을 분리하여 독립적으로 요인분석을 실시하였다.

한편 사용자 수용도 영역의 만족도, 지속성, 추천 의도는 각각 단일 문항으로 구성되어 요인분석을 적용하기 위한 최소 항목 수 요건을 충족하지 못하였다. 따라서 사용자 수용도 영역은 EFA 대상에서 제외하였다. 이에 따라 본 연구의 EFA 결과는 UVA 프레임워크 전체가 아니라 사용성 및 사용자 가치 차원의 구조적 타당성을 검토한 결과로 해석되어야 한다.

4. 2. 1. 사용성 카테고리 검증 결과

사용성 관련 14개 세부 지표에 대한 분석 결과, 표본 적합도(KMO)는 .927로 나타났으며, Bartlett의 구형성 검정 결과는 χ^2 = 1984.197(df = 91, p 〈 .001)로 확인되어 요인분석에 적합한 자료임이 검증되었다.

요인 수 결정은 고윳값 기준, 누적 분산 설명력, 이론적 구조와의 정합성, 요인 해석 가능성을 종합하여 이루어졌다. 초기 고윳값 1 이상 기준에서는 2개 요인이 도출되었으나, 이 기준만으로는 사용편의성, 인터랙션 적합성, 상황인지 특성, 콘텐츠 품질로 구분되는 이론적 주지표들이 과도하게 통합될 가능성이 있었다. 이에 본 연구는 스크리 플롯의 변화 양상과 이론적 구조, 요인별 적재 문항의 해석 가능성을 함께 고려하여 4요인 모형을 최종 모형으로 채택하였다. 4요인 모형의 누적 분산 설명력은 64.825%로 나타났다. 각 요인의 구성 및 적잿값은 [Table 5]에 제시하였다.

Table 5 Final EFA Results for the Usability Category

KMO의 표본 적합도(MSA)						0.927
Bartlett의 구형성 검정		근사 카이제곱				1984.197
		자유도				91
		유의확률				〈0.001
요인	요소	성분				공통성
		1	2	3	4	
요인1	콘텐츠 다양성	0.727	0.161	0.228	0.171	0.636
	콘텐츠 화질	0.678	0.315	0.237	0.147	0.637
	콘텐츠 최신성	0.669	0.459	−0.081	0.093	0.673
	콘텐츠 음질	0.636	0.087	0.277	0.311	0.585
	콘텐츠 번역 품질	0.567	0.102	0.438	0.218	0.572
요인2	인지용이성	0.200	0.766	0.202	0.074	0.672
	조작용이성	0.238	0.722	0.181	0.043	0.612
	탐색용이성	0.203	0.692	0.157	0.321	0.648
요인3	서비스 개인화	0.300	0.261	0.743	0.159	0.735
	콘텐츠 개인화	0.440	0.180	0.634	0.149	0.651
	사용자 맞춤형 유연성	−0.039	0.507	0.596	0.167	0.642
요인4	원활성	0.177	0.176	0.091	0.852	0.796
	정보전달성	0.401	0.183	0.374	0.534	0.620
	지원성	0.287	0.090	0.477	0.528	0.597

사용성 카테고리의 각 요인에 대한 신뢰도 분석 결과, 콘텐츠 품질 요인은 α = .818, 사용편의성 요인은 α = .743, 상황인지 특성 요인은 α = .744, 인터랙션 적합성 요인은 α = .721로 나타났다. 일반적으로 Cronbach's α .70 이상은 수용 가능한 수준의 내적 일관성을 의미하므로, 사용성 카테고리의 모든 요인은 양호한 신뢰도를 확보한 것으로 판단된다. 특히 콘텐츠 품질 요인은 가장 높은 신뢰도를 보였으며, 이는 콘텐츠 다양성, 최신성, 화질, 음질, 번역 품질이 OTT 서비스 이용 과정에서 비교적 일관된 콘텐츠 경험 차원으로 인식되고 있음을 보여준다.

4. 2. 2. 사용자 가치 카테고리 검증 결과

사용자 가치 관련 8개 세부 지표에 대한 분석 결과, KMO 값은 .850, Bartlett의 구형성 검정 결과는 χ^2 = 679.900(df = 28, p 〈 .001)로 나타나 요인분석의 적합성이 확보되었다.

요인 수 결정은 사용성 카테고리와 동일하게 고윳값 기준, 스크리 플롯 변화 양상, 이론적 구조와의 정합성, 요인 해석 가능성을 종합하여 이루어졌다. 초기 고윳값 1 이상 기준에서는 2개 요인이 도출되었으나, 사용자 가치 차원은 이론적으로 기능적 가치, 감정적 가치, 사회적 가치의 세 주지표로 구성되어 있다. 따라서 본 연구는 사용자 가치의 의미 구조를 반영하고 각 요인의 해석 가능성을 확보하기 위해 요인 수를 3개로 고정하여 추가 분석을 실시하였다. 최종적으로 3요인 모형이 채택되었으며, 누적 분산 설명력은 65.780%로 나타났다. 각 요인의 구성 및 적잿값은 [Table 6]에 제시하였다.

Table 6 Final EFA Results for the User Value Category

KMO의 표본 적합도(MSA)				0.850	
Bartlett의 구형성 검정		근사 카이제곱		679.900	
		자유도		28	
		유의확률		〈0.001	
요인	요소	성분			공통성
		1	2	3	
요인1	몰입성	0.830	0.047	0.104	0.702
	유희성	0.809	0.223	0.001	0.704
	반응성	0.568	0.500	−0.002	0.572
	심미성	0.516	0.463	0.237	0.537
요인2	보안	−0.018	0.728	0.320	0.633
	안정성	0.197	0.722	−0.003	0.560
	신뢰성	0.347	0.701	0.045	0.613
요인3	사회적 관계성	0.101	0.136	0.956	0.942

사용자 가치 카테고리의 신뢰도 분석 결과, 감정적 가치 요인은 α = .755로 나타나 양호한 내적 일관성을 보였다. 기능적 가치 요인은 α = .652로 나타나 일반적인 기준인 .70에는 다소 미치지 못하였으나, 탐색적 연구 단계에서 허용 가능한 수준으로 판단하였다. 이는 기능적 가치를 구성하는 보안, 안정성, 신뢰성이 모두 서비스의 기능적 품질과 관련되지만 사용자가 인지하는 경험 양상이 완전히 동일하지 않을 수 있음을 시사한다. 사회적 가치 요인은 사회적 관계성 단일 문항으로 측정되었기 때문에 Cronbach's α를 산출하지 않았다. 전체 25개 지표의 Cronbach's α는 .937로 나타나 전체 척도 수준에서는 매우 높은 내적 일관성을 보였다.

종합적으로 사용성 및 사용자 가치 차원의 세부 지표들은 이론적으로 설정한 상위 차원과 대체로 일관된 요인 구조를 형성하였다. 다만 기능적 가치 요인의 신뢰도(α = .652)는 일반적 기준(.70)에 다소 못 미쳐, 해당 요인이 이론적 구조에 완전히 부합한다고 단정하기보다는 잠정적으로 부합하는 경향을 보인다고 해석하는 것이 타당하다. 또한 사용자 수용도 차원의 세부 지표는 단일 문항으로 구성되어 EFA에서 제외되었으므로, 본 연구의 구조적 타당성 검증은 UVA 프레임워크의 사용성 및 사용자 가치 차원에 한정된다. 향후 연구에서는 사용자 수용도 차원의 다문항 척도를 개발하고 확인적 요인분석(CFA)을 통해 전체 UVA 프레임워크의 구조적 타당성을 추가 검증할 필요가 있다.

4. 3. UX 평가 지표 적용 결과

검증된 UX 평가 지표를 실제 OTT 서비스 4개를 대상으로 소규모 예비 적용한 결과, 서비스별로 상이한 사용자 경험 특성이 나타나는 경향이 확인되었다. 프리드먼 검정을 통해 지표별 서비스 간 차이를 분석한 결과, 사용자 맞춤형 유연성, 콘텐츠 다양성, 콘텐츠 최신성, 사회적 관계성, 지속성, 추천 의도에서 통계적으로 유의한 차이가 나타났다($p < .05$). 유의한 차이를 보인 6개 지표에 대한 서비스 간 순위는 [Table 7]에 제시하였으며, 세부 기술통계 및 검정 결과는 Appendix 3와 Appendix 4에 제시하였다.

Table 7 Service Rankings by App Based on Significant UX indicators from the Friedman Test

지표	순위							
사용자 맞춤형 유연성 (Customization Flexibility)	넷플릭스(4.4)	=	티빙(4.4)	〉	왓챠플레이(3.5)	=	아마존프라임(3.5)	
콘텐츠 다양성 (Contents Diversity)	넷플릭스(4.5)	〉	왓챠플레이(4.2)	〉	아마존프라임(3.4)	〉	티빙(3.3)	
콘텐츠 최신성 (Contents Recency)	넷플릭스(4.7)	〉	티빙(4.3)	〉	왓챠플레이(3.9)	〉	아마존프라임(3.7)	
사회적 관계성 (Social Relatedness)	왓챠플레이(3.5)	〉	티빙(2.9)	〉	넷플릭스(2.1)	〉	아마존프라임(1.8)	
지속성 (Stickiness)	티빙(3.9)	〉	넷플릭스(3.6)	〉	왓챠플레이(3.3)	〉	아마존프라임(2.5)	
추천 의도 (Word-of-Mouth Intention)	티빙(3.8)	=	넷플릭스(3.8)	〉	왓챠플레이(3.4)	〉	아마존프라임(2.6)	

사용성 차원에서는 전문 연구원 집단의 평가에서 넷플릭스가 인지용이성, 조작용이성, 탐색용이성, 원활성 등 전반적인 인터랙션 품질 측면에서 상대적으로 높은 점수를 보이는 경향이 나타났다.

콘텐츠 관련 지표에서는 넷플릭스가 콘텐츠 다양성과 최신성에서 상대적으로 높은 순위를 보였다. 이는 글로벌 콘텐츠 수급 구조와 신작 업데이트 전략이 사용자 인식에 반영되었을 가능성을 시사한다. 아마존프라임은 정보전달성에서 상대적으로 높은 평가를 받았는데, 시청 중 배우, 등장인물, OST 정보를 확인할 수 있는 X-Ray 기능이 정보 이해를 보조하는 역할을 한 것으로 해석된다.

사회적 관계성에서는 왓챠플레이가 상대적으로 높은 점수를 보였다. 이는 왓챠파티 및 리뷰 중심 커뮤니티 기능이 함께 시청하거나 의견을 공유하는 경험을 지원하는 서비스 특성과 관련된 것으로 추정된다.

지속성과 추천 의도에서는 티빙과 넷플릭스가 상대적으로 높은 순위를 기록하였다. 이는 콘텐츠 만족과 사용 편의성이 장기 이용 의향 및 추천 행동으로 연결될 수 있음을 시사한다.

4. 4. UX 평가 지표 활용 결과

본 연구는 UX 평가 지표를 실무에서 활용 가능한 설계 기준으로 확장하기 위해, 사용성 차원을 중심으로 UX 디자인 가이드라인을 도출하였다. 도출된 가이드라인은 UX 평가 지표를 서비스 설계 및 개선 과정에서 직접 적용할 수 있도록 디자인 방향성과 구체적 솔루션 형태로 재구성되었다.

특히 가이드라인 도출 과정에서는 실제 OTT 서비스에서 제공되는 기능 사례를 참고하여 지표와 설계 요소 간의 연결 가능성을 명확히 하였다. 예를 들어, 콘텐츠 음질 지표와 관련하여 아마존 프라임의 'Dialogue Boost' 기능을 사례로 제시함으로써, 음성 명료도 향상을 위한 기능 설계 방향을 구체화하였다. 이와 같이 실제 서비스 기능을 근거로 제시함으로써, 평가 지표가 추상적 기준에 머무르지 않고 구현 가능한 설계 전략으로 전환될 수 있음을 보여주고자 하였다.

해당 가이드라인은 화면 구성, 인터랙션 설계, 기능 개선 과정에서 즉시 활용 가능한 기준을 제시하며, UX 평가 결과를 서비스 개선 전략으로 연결하는 순환 구조를 갖는다. 사용성 중심 가이드라인은 [Table 8]에 제시하였다.

Table 8 UX Design Guidelines Based on User Experience Evaluation indicators

주지표	상세지표	디자인 방향성 및 솔루션
사용편의성 (Ease of Use)	인지용이성 (Ease of perception)	사용자가 화면에 제시된 정보의 의미와 기능을 즉시 해석할 수 있도록 시각적 단서와 구조를 명확히 제공해야 함
		• 메뉴 아이콘 + 텍스트 라벨 병기 • 정보 우선순위의 시각적 위계 확립 • 신규 사용자 온보딩 제공 • 일관된 색상·아이콘 체계 유지
	조작용이성 (Ease of operation)	사용자가 최소한의 학습으로도 원하는 동작을 수행할 수 있도록 조작 체계를 일관되고 직관적으로 설계해야 함
		• 재생·일시정지·구간 이동 등 핵심 기능 단순화 • 더블탭 이동 등 직관적 조작 도입 • 제스처 표준화(밝기·음량·타임라인 드래그) • 오조작 방지 위한 터치 영역 확보
	탐색용이성 (Ease of navigation)	사용자가 목표 콘텐츠에 도달하기까지의 경로를 단순화하고 탐색 중 인지적 부담을 최소화해야 함
		• 카테고리 구조 최소 단계화 • 추천·검색·탐색 경로 분리 • 자동완성·인기/연관 검색 제공 • 필터·정렬 기능 직관적 배치
인터랙션 적합성 (Interaction Suitability)	원활성 (Seamlessness)	기기·모드·과업 전환 과정에서 끊김 없는 경험을 보장하여 사용 흐름의 연속성과 몰입감을 유지해야 함
		• 기기 간 재생 기록 실시간 동기화 • 모드 전환 시 화면비 유지 • 프리로딩 기반 끊김 최소화 • 이어보기·백그라운드 재생 강화
	지원성 (Supportiveness)	사용자가 문제 상황에서 즉시 해결책에 접근할 수 있도록 예측 가능한 도움 구조와 명확한 지원 체계를 제공해야 함
		• 오류 유형별 해결 가이드 제공 • 고객센터 접근 경로 단축 • 설정 화면 내 도움말 배치 • 채팅형 실시간 지원 제공
	정보전달성 (Informativeness)	사용자가 의사결정에 필요한 정보를 적절한 시점과 맥락에 맞게 제공하여 탐색·선택 과정의 효율성을 높여야 함
		• 장르·상영시간·출연진 등 핵심 정보 상단 배치 • 단계적 정보 노출 설계 • 장면별 정보(OST·캐릭터 등) 제공
상황 인지 특성 (Context-Awareness)	콘텐츠 개인화 (Contents Personalization)	사용자의 시청 패턴과 선호를 기반으로 개인별 맥락에 맞춘 적합한 콘텐츠를 지속적으로 제안해야 함
		• 시청 기록·선호 장르 기반 추천 정밀화 • 추천 이유 제공 • 사용자 평가 기반 알고리즘 학습 반영 • 상황별 맞춤 추천 제공
	서비스 개인화 (Service Personalization)	사용자의 설정·환경·기기 특성을 반영하여 서비스 전반을 개인의 사용 맥락에 맞게 조정할 수 있도록 해야 함
		• 설정(자막·속도 등) 프로필별 저장 • 홈 화면 구성 사용자 편집 허용 • 테마·알림·품질 설정 개인화 • 기기 간 동기화 유지
	사용자 맞춤형 유연성 (Customization Flexibility)	사용자가 원하는 방식으로 시청 환경을 세밀하게 제어할 수 있도록 다양한 설정을 자유롭게 조합할 수 있게 해야 함
		• 자막·오디오 스타일 세부 조정 • 화면 비율·배속·화질 옵션 제공 • 실시간 즉시 적용 UI 제공 • 접근성 옵션 강화

콘텐츠 품질 (Contents Quality)	콘텐츠 다양성 (Contents Diversity)	다양한 취향과 문화적 배경을 가진 이용자가 만족할 수 있도록 폭넓은 스펙트럼의 콘텐츠 선택권을 보장해야 함
		• 국가·언어·장르 균형 확보 • 부족 장르 보완 큐레이션 • 세분화된 필터 제공 • 소수 취향 콘텐츠 추천 강화
	콘텐츠 최신성 (Contents Recency)	최신 콘텐츠를 빠르게 발견하고 접근할 수 있도록 신작 중심의 시각적 강조와 갱신 체계를 유지해야 함
		• 최신 콘텐츠 전용 영역 고정 배치 • 출시일 기준 정렬 기능 제공 • 신작 알림·사전 공개 기능 강화 • 신작 카테고리 시각적 강조
	콘텐츠 화질 (Contents Video Quality)	시청 환경에 맞춰 최적의 화질을 제공하고 품질 선택권을 명확히 제시하여 시청 품질 경험을 안정적으로 관리해야 함
		• 자동 화질 조정 고도화 • 화질 등급 명확 표시 • 화질 고정 옵션 제공 • HDR·Dolby Vision 등 프리미엄 화질 정보 제공
	콘텐츠 음질 (Contents Audio Quality)	다양한 시청 환경에서도 청취 품질이 유지되도록 음향 모드를 최적화하고 사용자 제어권을 확보해야 함
		• 대사 명료도 강화 기능(예: Dialogue Boost) • 다중 음향 모드 제공 • Dolby Atmos 등 고품질 포맷 지원 • 시청 중 즉시 조정 UI 제공
	콘텐츠 번역 품질 (Contents Localization Quality)	문화적 맥락과 언어적 정확성을 확보한 번역을 제공하여 사용자 이해도와 몰입도를 저해하지 않도록 해야 함
		• 전문 번역·검수 체계 강화 • 장르별 용어 가이드 구축 • 오역 신고 기능 제공 • 자막·더빙 번역의 일관성 유지 • 문화 맥락 반영

5. 논의 및 시사점

본 연구는 OTT 서비스 사용자 경험을 통합적으로 평가하기 위한 UX 평가 지표 체계를 개발하고, 이를 구조적·실증적으로 검토하였다. 특히 OTT 서비스 UX를 사용성, 사용자 가치, 사용자 수용도의 다차원적 구조로 재정의하고, 문헌 연구, 사용자 관찰 연구, 전문가 검토, 설문 기반 검증을 통해 지표 체계를 단계적으로 정교화했다는 점에서 의의를 갖는다.

5. 1. 통합적 UX 평가 지표 체계의 도출

본 연구는 OTT 맥락에 적합한 UX 평가 지표 체계를 도출하고자 하였다(RQ1). 문헌 통합과 사용자 관찰 연구를 통해 사용성, 사용자 가치, 사용자 수용도로 구성된 UVA 기반 지표 체계를 확정하였다. 특히 콘텐츠 탐색용이성과 콘텐츠 번역 품질은 실제 이용 맥락에서 핵심 요소로 반복적으로 도출되었으며, 이는 OTT 서비스 UX가 인터페이스 사용성을 넘어 콘텐츠 탐색 및 이해 경험을 포함하는 복합적 구조임을 시사한다.

이러한 결과는 기존 이론적 논의와의 연계 속에서 의미를 갖는다. Morville(2005)이 정보 환경 전반에서 논의한 발견가능성(findability) 개념은 사용자가 원하는 정보를 얼마나 쉽게 발견할 수 있는가를 핵심으로 한다. 본 연구에서 탐색용이성이 사용자 관찰 데이터에서 반복적으로 언급된 경험 요소로 도출된 것은, findability 개념이 OTT 콘텐츠 소비 맥락에서도 사용자 경험의 핵심 요소로 작동함을 보여준다. 특히 OTT 서비스는 방대한 콘텐츠를 제공하기 때문에 사용자가 원하는 콘텐츠를 발견하는 과정에서 인지적 부담이 커질 수 있다. 따라서 탐색용이성은 OTT UX의 차별적 평가 요소로서 독립적인 지표로 체계화될 필요가 있다.

또한 콘텐츠 번역 품질이 사용성 차원의 독립적 세부 지표로 도출된 것은 이론적으로 주목할 만한 결과이다.

기존 사용성 평가 이론과 디지털 서비스 UX 연구에서는 언어적·문화적 지역화 품질이 독립적인 평가 차원으로 충분히 체계화되지 않은 경우가 많았다. 그러나 글로벌 OTT 콘텐츠 소비가 일상화된 환경에서 사용자는 자막 번역의 부정확성, 문화적 맥락의 손실, 장르별 표현의 어색함을 실제 이용 과정의 핵심 불편 요소로 인식한다. 본 연구는 이러한 지역화 품질이 단순한 콘텐츠 보조 요소가 아니라 OTT UX 평가 체계에서 독립적으로 고려되어야 할 사용성 지표임을 제안한다.

5. 2. 구조적 타당성 검증의 의미

탐색적 요인분석 결과, 사용성 및 사용자 가치 차원의 세부 지표들은 이론적으로 설정한 상위 차원과 대체로 일관된 요인 구조를 형성하였다(RQ2). 이는 UVA 프레임워크의 사용성 및 사용자 가치 차원에 해당하는 지표들이 개념적 모델을 넘어 실제 응답 데이터에서도 대체로 구조적 정합성을 나타냄을 시사한다. 다만 기능적 가치 요인과 같이 신뢰도가 상대적으로 낮은 요인에 대해서는 구조적 부합성을 단정적으로 해석하기보다 잠정적 증거로 이해할 필요가 있다.

신뢰도 분석 결과 역시 도출된 요인 구조의 안정성을 뒷받침한다. 사용성 카테고리의 콘텐츠 품질(α = .818), 사용편의성(α = .743), 상황인지 특성(α = .744), 인터랙션 적합성(α = .721)은 모두 수용 가능한 수준 이상의 내적 일관성을 보였다. 사용자 가치 카테고리의 감정적 가치 또한 α = .755로 양호한 신뢰도를 나타냈다. 기능적 가치는 α = .652로 일반적인 기준인 .70에는 미치지 못했으나, 탐색적 연구 단계에서는 수용 가능한 수준으로 해석할 수 있다. 이는 보안, 안정성, 신뢰성이 모두 기능적 가치와 관련되어 있으나 사용자가 체감하는 경험 차원에서는 다소 구분될 수 있음을 시사한다.

5. 3. 서비스 간 UX 차이 분석의 함의

소규모 예비 적용 결과, 제안된 UX 평가 지표 체계는 실제 OTT 서비스 간 UX 특성의 차이를 일정 수준에서 포착하는 것으로 나타났다(RQ3). 사용자 맞춤형 유연성, 콘텐츠 다양성, 콘텐츠 최신성, 사회적 관계성, 지속성, 추천 의도에서 서비스 간 차이가 확인되었으며, 이는 OTT 서비스 간 경쟁이 단순히 콘텐츠 보유량에 의해서만 결정되는 것이 아니라 탐색 구조, 개인화 전략, 업데이트 방식, 사회적 기능 등 경험 설계의 차이에 의해 형성될 수 있음을 시사한다.

다만 본 적용 단계는 HCI 및 서비스디자인 연구원 10명을 대상으로 한 소규모 사례 적용이므로, 서비스 간 UX 우열을 일반화하는 결과로 해석해서는 안 된다. 본 결과는 제안된 지표 체계가 서비스별 UX 특성의 차이를 탐색적으로 드러낼 수 있는 가능성을 보여주는 예비적 근거로 이해하는 것이 타당하다.

5. 4. 실무적 시사점

본 연구는 UX 평가 결과를 사용성 중심 디자인 가이드라인으로 확장함으로써, 평가–분석–개선으로 이어지는 실천적 활용 구조를 제안하였다(RQ4). 특히 실제 서비스 기능 사례를 기반으로 가이드라인을 제시함으로써, 지표와 설계 전략 간의 연결 가능성을 구체화하였다.

실무적으로 본 지표 체계는 OTT 서비스의 UX 진단 도구로 활용될 수 있다. 서비스 기획자와 디자이너는 각 지표별 평가 결과를 통해 사용자가 어느 지점에서 어려움이나 만족을 경험하는지 파악할 수 있으며, 이를 기반으로 개선 우선순위를 설정할 수 있다. 예를 들어 탐색용이성 점수가 낮은 서비스는 검색, 추천, 카테고리 구조, 필터 기능을 우선적으로 점검할 수 있으며, 콘텐츠 번역 품질 점수가 낮은 서비스는 자막 검수 체계, 장르별 용어 가이드 등을 개선 전략으로 설정할 수 있다.

5. 5. 연구의 한계 및 향후 연구 방향

본 연구는 다음과 같은 한계를 지닌다.

첫째, UVA 프레임워크의 구조적 검증 범위가 제한적이다. 본 연구의 EFA는 사용성 및 사용자 가치 차원에 한정하여 구조적 타당성을 검증하였다. 사용자 수용도 차원의 세부 지표인 만족도, 지속성, 추천 의도는 각각 단일 문항으로 구성되어 요인분석의 최소 항목 수 요건을 충족하지 못하였으므로 EFA에서 제외되었다. 따라서 본 연구 결과만으로 UVA 프레임워크 전체의 구조적 타당성이 완전히 검증되었다고

보기는 어렵다. 향후 연구에서는 사용자 수용도 차원의 다문항 척도를 개발하고, 확인적 요인분석(CFA) 또는 구조방정식모형(SEM)을 통해 사용성, 사용자 가치, 사용자 수용도 간의 관계를 검증할 필요가 있다.

둘째, 서비스 적용 단계의 표본과 해석 범위에 한계가 있다. 본 연구의 실제 OTT 서비스 4종에 대한 UX 평가 지표 적용은 HCI 및 서비스디자인 전공 연구원 10명을 대상으로 이루어졌다. 이들은 인터페이스 사용성에 대한 전문적 감수성을 보유하고 있어 평가의 일관성과 분석적 관찰에는 장점이 있으나, 일반 OTT 이용자와 체계적으로 다른 평가 경향을 보일 수 있다. 따라서 본 적용 결과는 일반 사용자 경험의 차이를 확정적으로 보여주는 결과가 아니라, 지표 체계의 변별력 가능성을 확인하기 위한 소규모 예비 적용으로 해석되어야 한다. 향후 연구에서는 일반 OTT 이용자를 대상으로 동일한 지표 체계를 적용한 반복 검증을 실시함으로써 본 예비 적용 결과의 일반화 가능성을 확인할 필요가 있다.

셋째, 사회적 가치 차원의 측정 한계가 존재한다. 사회적 가치 차원의 유일한 세부 지표인 사회적 관계성은 단일 문항으로 측정되어 내적 일관성 신뢰도 분석이 불가능하였다. 또한 요인분석에서 단독 요인을 형성하였기 때문에, 향후 연구에서는 사회적 상호작용, 공동 시청, 리뷰 및 평가 공유, 커뮤니티 참여, 지인 기반 추천 등 사회적 시청 경험을 세분화한 다문항 척도를 개발할 필요가 있다.

넷째, 본 연구는 SVOD(Subscription Video On Demand, 구독형 주문형 비디오) 중심의 영상 스트리밍 OTT 서비스를 주요 대상으로 하였다. 따라서 광고 기반 AVOD, FAST, 스포츠 중계형 OTT, 교육형 OTT, 음악 스트리밍 서비스 등 다른 형태의 콘텐츠 소비 환경까지 본 지표 체계가 동일하게 적용될 수 있는지는 추가 검증이 필요하다. 향후 연구에서는 서비스 유형과 문화권을 확장하여 본 지표 체계의 일반성과 맥락 적합성을 검토할 필요가 있다.

다섯째, 본 연구에서 제안한 UX 디자인 가이드라인은 지표 정의 및 실제 서비스 사례에 기반하여 도출되었을 뿐, 가이드라인의 실제 적용 효과나 타당성에 대한 별도의 검증 절차를 거치지 않았다. 따라서 이는 검증된 설계 규범이라기보다 지표 체계의 실무적 활용 가능성을 보여주는 예시적 제안으로 해석되어야 하며, 향후 연구에서는 가이드라인 적용에 따른 실제 사용성 개선 효과를 실험적으로 검증할 필요가 있다.

6. 결론

본 연구는 OTT 서비스 사용자 경험을 통합적으로 평가하기 위한 UX 평가 지표 체계를 개발하고, 이를 실증적으로 검토하였다. 연구 결과, OTT 맥락에 적합한 다차원적 지표 체계를 도출하였으며, 사용성 및 사용자 가치 차원의 지표들이 이론적 구조에 부합하는 요인 구조를 형성함을 확인하였다. 다만 사용자 수용도 차원은 단일 문항 구성의 한계로 이번 EFA 검증 범위에서 제외되었으며, 향후 연구를 통해 보완될 필요가 있다.

본 연구의 주요 기여는 다음과 같다. 첫째, OTT UX를 사용성 중심의 단일 평가에서 벗어나 사용자 가치와 사용자 수용도까지 포괄하는 다차원적 구조로 제안하였다. 둘째, 콘텐츠 탐색용이성과 콘텐츠 번역 품질이 실제 OTT 이용 맥락에서 핵심 경험 요소로 작동함을 확인하였다. 이는 findability 개념이 OTT 콘텐츠 소비 환경에서도 중요한 UX 평가 요소로 작동함을 보여주며, 글로벌 콘텐츠 소비가 확대되는 환경에서 localization quality를 독립적인 UX 지표로 고려할 필요성을 제기한다. 셋째, 개발된 지표를 실제 OTT 서비스에 예비 적용함으로써, 서비스별 UX 특성을 구조적으로 진단할 수 있는 가능성을 확인하였다.

실무적으로 본 연구는 UX 평가 결과를 디자인 가이드라인으로 확장함으로써, 평가 결과가 서비스 개선 전략으로 환류될 수 있는 구조를 제안하였다. 이는 OTT 서비스 기획자와 디자이너가 특정 서비스의 문제 영역을 지표 단위로 진단하고, 개선 우선순위를 설정하는 데 활용될 수 있다. 결론적으로 본 연구는 OTT 서비스 사용자 경험을 맥락적이고 다차원적인 구조로 평가하기 위한 이론적 틀과 실무적 도구를 함께 제안했다는 점에서 의의를 갖는다.

References

1. Ahn, J. H., & Kim, S. I. (2016). 비모수적 통계방법을 이용한 모바일 페이스북과 인스타그램의 사용자 경험 차이 비교 – 감성인터페이스 모형을 중심으로 – [A comparison of user experience differences between mobile Facebook and Instagram using nonparametric statistical methods: Focusing on the emotional interface model]. *디지털융복합연구[Journal of Digital Convergence], 14*(11), 481–488. https://doi.org/10.14400/JDC.2016.14.11.481
2. T. Ajith, N., P., S., Periaiya, S., & Mathew, L. S. (2025). Experience matters: Exploring the impact of user experience on stickiness and loyalty in OTT platforms. *International Journal of Human–Computer Interaction, 41*(18), 11363–11377. https://doi.org/10.1080/10447318.2024.2443245
3. Baharuddin, R., Singh, D., & Razali, R. (2013). Usability dimensions for mobile applications–A review. *Research Journal of Applied Sciences, Engineering and Technology, 5*(6), 2225–2231. https://doi.org/10.19026/rjaset.5.4776
4. Batavia, A. I., & Hammer, G. S. (1990). Toward the development of consumer-based criteria for the evaluation of assistive devices. *Journal of Rehabilitation Research and Development, 27*(4), 425–436. https://doi.org/10.1682/jrrd.1990.10.0425
5. Bevan, N. (1995). Measuring usability as quality of use. *Software Quality Journal, 4*(2), 115–130. https://doi.org/10.1007/BF00402715
6. Bevan, N., & Macleod, M. (1994). Usability measurement in context. *Behaviour & Information Technology, 13*(1–2), 132–145. https://doi.org/10.1080/01449299408914592
7. Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly, 25*(3), 351–370. https://doi.org/10.2307/3250921
8. Choi, J. W., & Kim, J. B. (2017). 전자상거래 애플리케이션의 사용성 평가 지표 [Usability evaluation indicators for e-commerce applications]. *예술인문사회 융합 멀티미디어 논문지[Asia-Pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology], 7*(11), 117–126. https://doi.org/10.14257/ajmahs.2017.11.81
9. Choo, Y. J. (2007). *사용성 향상을 위한 그래픽 유저 인터페이스(GUI) 디자인에 관한 연구: 종합병원 키오스크의 시각적 구성요소를 중심으로 [A study on graphic user interface (GUI) design for usability improvement: Focusing on the visual components of general hospital kiosks]* [Master's thesis, Hongik University].
10. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008
11. DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems, 19*(4), 9–30. https://doi.org/10.1080/07421222.2003.11045748
12. Desmet, P. (2018). Measuring emotion: Development and application of an instrument to measure emotional responses to products. In M. Blythe & A. Monk (Eds.), *Funology 2: From usability to enjoyment* (2nd ed., pp. 391–404). Springer. https://doi.org/10.1007/978-3-319-68213-6_25
13. Desmet, P., & Hekkert, P. (2007). Framework of product experience. *International Journal of Design, 1*(1), 57–66.
14. DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Sage.
15. 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. https://doi.org/10.1037/1082-989X.4.3.272
16. Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing Interactive Systems* (pp. 261–268). ACM. https://doi.org/10.1145/1013115.1013152
17. Grant, J. S., & Davis, L. L. (1997). Selection and use of content experts for instrument development. *Research in Nursing & Health, 20*(3), 269–274. https://doi.org/10.1002/(SICI)1098-240X(199706)20:3〈269::AID-NUR9〉3.0.CO;2-G
18. Hassenzahl, M. (2010). *Experience design: Technology for all the right reasons*. Morgan & Claypool. https://doi.org/10.2200/S00261ED1V01Y201003HCI008

19. Hassenzahl, M. (2018). The thing and I: Understanding the relationship between user and product. In M. Blythe & A. Monk (Eds.), *Funology 2: From usability to enjoyment* (2nd ed., pp. 301–313). Springer. https://doi.org/10.1007/978-3-319-68213-6_19
20. Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J. Ziegler & G. Szwillus (Eds.), *Mensch & Computer 2003: Interaktion in Bewegung* (pp. 187–196). B.G. Teubner. https://doi.org/10.1007/978-3-322-80058-9_19
21. Hassenzahl, M., & Tractinsky, N. (2006). User experience-A research agenda. *Behaviour & Information Technology, 25*(2), 91–97. https://doi.org/10.1080/01449290500330331
22. Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967–988. https://doi.org/10.1177/014920639502100509
23. Holbrook, M. B. (1994). The nature of customer value: An axiology of services in the consumption experience. In R. T. Rust & R. L. Oliver (Eds.), *Service quality: New directions in theory and practice* (pp. 21–71). Sage. https://doi.org/10.4135/9781452229102.n2
24. International Organization for Standardization. (2018). *ISO 9241-11: Ergonomics of human-system interaction-Usability definitions and concepts*.
25. ISO/IEC. (2001). *ISO/IEC 9126: Software engineering-Product quality*.
26. Jang, J., & Yi, M. Y. (2019). Determining and validating smart TV UX factors. *International Journal of Human-Computer Studies, 130*, 58–72. https://doi.org/10.1016/j.ijhcs.2019.05.001
27. Jung, H. J., Kim, K. M., Cho, S., & Ko, Y. J. (2015). 서비스디자인 플랫폼을 위한 사용성 평가지표 연구 [A study on usability evaluation indicators for a service design platform]. *한국과학예술포럼[Korea Science & Art Forum], 20,* 409–419. https://doi.org/10.17548/ksaf.2015.06.20.409
28. Keiningham, T. L., Cooil, B., Aksoy, L., Andreassen, T. W., & Weiner, J. (2007). The value of different customer satisfaction and loyalty metrics in predicting customer retention, recommendation, and share-of-wallet. *Managing Service Quality, 17*(4), 361–384. https://doi.org/10.1108/09604520710760526
29. Kim, B. H. (2015). *중등학교 STEAM 교육의 교수역량 평가지표 개발 및 타당화[Development and validation of teaching competency evaluation indicators for secondary STEAM education]* [Doctoral dissertation, Korea National University of Education].
30. Kim, J. H. (2025). *LLM 기반 교육용 챗봇의 사용성 평가 지표 개발 및 적용[Development and application of usability evaluation indicators for LLM-based educational chatbots]* [Master's thesis, Sungshin Women's University].
31. Kim, S. (2020). *OTT 플랫폼 UI 사용성이 사용자 만족도 및 충성도에 미치는 영향 [Effects of OTT platform UI usability on user satisfaction and loyalty]* [Master's thesis, Chung-Ang University].
32. Kim, Y. J., & Kim, T. W. (2025). 국내・외 OTT 플랫폼 콘텐츠 썸네일 프리뷰 재생방식의 사용성 및 선호도 연구[Usability and preference of thumbnail preview playback in OTT platforms]. *커뮤니케이션디자인학연구[Journal of Communication Design], 90*, 81–94.
33. Kim, Y. K., & Park, S. H. (2023). 자율주행 셔틀 인포테인먼트 서비스 사용성 평가 지표 개발 [Development of self-driving shuttle infotainment service usability assessment index]. *Journal of Korea Multimedia Society, 26*(9), 1182–1196. https://doi.org/10.9717/kmms.2023.26.9.1182
34. Lee, J. W., Choi, Y. E., & Lee, S. I. (2012). 고령자를 위한 병원 안내시스템 설계 및 평가 [Design and evaluation of a hospital guidance system for older adults]. *대한산업공학회 춘계공동학술대회 논문집[Proceedings of the Korean Institute of Industrial Engineers Spring Joint Conference]*.
35. Lee, K. (2007). *모바일 컨텐츠의 사용자 만족 및 사용의도에 영향을 미치는 요인 [Factors influencing user satisfaction and usage intention of mobile content]* [Master's thesis, Chungnam National University].
36. Mehrabian, A., & Russell, J. A. (1974). *An approach to environmental psychology*. MIT Press.
37. Morville, P. (2005). *Ambient findability*. O'Reilly Media.
38. Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.
39. Ofcom. (2024). *Media Nations 2024*. https://www.ofcom.org.uk/media-use-and-attitudes/media-habits-adults/media-nations-2024

40. Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing, 64*(1), 12-40.
41. Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg & C. I. Hovland (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1-14). Yale University Press.
42. Seomun, H. (2020). *시청자 경험을 고려한 OTT 서비스 인터페이스 디자인 연구[A study on OTT service interface design considering viewer experience]* [Master's thesis, Yonsei University].
43. Shackel, B. (2009). Usability-Context, framework, definition, design and evaluation. *Interacting with Computers, 21*(5-6), 339-346. https://doi.org/10.1016/j.intcom.2009.04.007 (Original work published 1981)
44. Sheth, J. N., Newman, B. I., & Gross, B. L. (1991). Why we buy what we buy: A theory of consumption values. *Journal of Business Research, 22*(2), 159-170. https://doi.org/10.1016/0148-2963(91)90050-8
45. Shin, S. M., & Kim, T. G. (2022). OTT 서비스의 모바일 앱 UI 디자인 사용성 연구 - 20대 사용자를 중심으로- [A study on the usability of mobile app UI design for OTT services: Focusing on users in their 20s]. *한국콘텐츠학회논문지[The Journal of the Korea Contents Association], 22*(5), 125-135. https://doi.org/10.5392/JKCA.2022.22.05.125
46. Sweeney, J. C., & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing, 77*(2), 203-220. https://doi.org/10.1016/S0022-4359(01)00041-0
47. UNCTAD. (2021). *Digital Economy Report 2021: Cross-border data flows and development-For whom the data flow*. United Nations. https://unctad.org/publication/digital-economy-report-2021
48. Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157-178. https://doi.org/10.2307/41410412
49. Vermeeren, A. P. O. S., Law, E. L.-C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010). User experience evaluation methods: Current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 521-530). ACM. https://doi.org/10.1145/1868914.1868973
50. Won, J. W. (2021). OTT 서비스의 콘텐츠 유형별 정보인지 디자인 연구 [A study on information cognition design by content type of OTT services]. *커뮤니케이션디자인학연구[Journal of Communication Design], 75*, 143-152.
51. Woodruff, R. B. (1997). Customer value: The next source for competitive advantage. Journal of the Academy of Marketing Science, 25(2), 139-153. https://doi.org/10.1007/BF02894350
52. Yeom, D. S. (2022). OTT 서비스에 대한 사용자 경험이 지속적 이용 의도에 미치는 영향: 이용 만족의 매개효과를 중심으로 [The effect of user experience on continuance intention of OTT services: Focusing on the mediating effect of user satisfaction]. *디지털융복합연구[Journal of Digital Convergence], 20*(5), 513-523. https://doi.org/10.14400/JDC.2022.20.5.513

Appendix

[Appendix 1] User Experience Evaluation indicators Literature Collection

구분	제목	문헌유형[1)]	저자 (발행연도)	지표	개수
1	Toward the Development of Consumer-Based Criteria for the Evaluation of Assistive Devices	GEN-UX	Batavia & Hammer (1990)	Affordability, Compatibility, Repairability, Dependability, Durability, Ease of Assembly, Maintenance, Effectiveness, Flexibility, Learnability, Operability, Comfort, Security, Portability	14
2	A Guide to Usability: Human Factors in Computing	GEN-UX	Preece, J (1993)	Effectiveness, Efficiency, Safety, Functionality, Learnability, Memorability	6
3	Usability Engineering	GEN-UX	Nielsen (1994)	Efficiency, Memorability, Learnability, Error Tolerance, Satisfaction	5
4	A Proposed Index of Usability	GEN-UX	Lin, Choong & Salvendy (1997)	Compatibility, Consistency, Flexibility, Learnability, Minimal Action, Minimal Memory Load, Perceptual Limitation, User Guidance	8
5	Ergonomic Requirements for Office Work with Visual Display Terminals	GEN-UX	ISO (1998)	Effectiveness, Efficiency, Satisfaction	3
6	Affordance, Conventions, and Design	GEN-UX	Norman (1999)	Visibility, Feedback, Mapping, Consistency, Control and Freedom, User-Centered Design, Error Prevention, Simplicity and Minimalism, Recoverability, Utility	10
7	Measuring Usability with the USE Questionnaire	GEN-UX	Lund (2001)	Usefulness, Ease of Use, Ease of Learning, Satisfaction	4
8	Software Engineering, Product Quality	GEN-UX	ISO/IEC (2001)	Understandability, Learnability, Operability, Attractiveness	4
9	Ambient Findability	GEN-UX	Morville (2005)	Usefulness, Usability, Satisfaction, Reliability, Findability, Accessibility, Value	7
10	Service Quality, Perceived Value, Customer Satisfaction, and Post-Purchase Intention	DEV·SVC-UX	Kuo et al. (2009)	Service Quality (contents, Navigation, Design, Reliability), Perceived Value, Satisfaction, Post-Purchase Intention	8
11	Usability Dimensions for Mobile Applications	GEN-UX	Baharuddin et al. (2013)	Efficiency, Effectiveness, Satisfaction, Usefulness, Pleasure, Learning, Simplicity, Intuition, Understandable, Attractiveness	10
12	OTT 플랫폼 UI 사용성이 사용자 만족도 및 충성도에 미치는 영향	OTT-UX	김수 (2020)	학습 용이성, 사용 효율성, 기억 용이성, 최소한의 에러, 주관적 즐거움, 만족도, 충성도	7
13	시청자 경험을 고려한 OTT 서비스 인터페이스 디자인 연구	OTT-UX	서문홍익 (2020)	일관성, 간결성, 접근성, 효율성, 심미성	5
14	OTT 서비스의 콘텐츠 유형별 정보인지 디자인 연구	OTT-UX	원종욱 (2021)	가시성, 주목성, 정확성, 조형성	4
15	OTT 서비스의 모바일 앱 UI 디자인 사용성 연구	OTT-UX	신새미, 김태균 (2022)	유용성, 사용성, 검색성, 매력성, 신뢰성, 가치성	6
16	OTT 서비스에 대한 사용자 경험이 지속적 이용 의도에 미치는 영향	OTT-UX	염동섭 (2022)	기능성, 신뢰성, 사용성, 편리성, 유희성, 이용 만족, 지속적 이용 의도	7
17	What Drives User Stickiness and Satisfaction in OTT Video Streaming Platforms	OTT-UX	Periaiya & Nandukrishna (2024)	Entertainment, Escape, Modality, Navigability, Passing Time, Personalization, Affective Media Satisfaction, Stickiness	8
18	국내·외 OTT 플랫폼 콘텐츠 썸네일 프리뷰 재생방식의 사용성 및 선호도 연구	OTT-UX	김예지, 김태완 (2025)	탐색용이성, 정보전달성, 시각적 만족도, 사용편의성, 흥미유발성, 선호도	6

1) 문헌 유형은 GEN-UX(도메인 비한정 일반 인터랙티브 시스템), OTT-UX(OTT·영상 스트리밍 직접 대상), DEV·SVC-UX(OTT 인접 개발·서비스 환경)의 세 가지로, 연구 대상과 서비스 범위를 기준으로 구분한 것이다.

19	Experience Matters: Exploring the Impact of User Experience on Stickiness and Loyalty in OTT Platforms	OTT-UX	Ajith et al. (2025)	Cognitive Experience, Affective Experience, Sensorial Experience, Behavioral Experience, Stickiness, Loyalty	6
20	Measuring Perceived Quality of Speech and Video	DEV·SVC-UX	Watson & Sasse (1998)	Speech Quality, Video Quality	2
21	사용성 향상을 위한 그래픽 유저 인터페이스(GUI) 디자인에 관한 연구	GEN-UX	추영지 (2007)	일관성, 경제성, 안정성, 심미성, 간결성, 의사소통성	6
22	모바일 콘텐츠의 사용자 만족 및 사용의도에 영향을 미치는 요인	DEV·SVC-UX	리강 (2007)	정확성, 다양성, 오락성, 유용성, 자기효능감, 혁신성향, 몰입성향, 프라이버시위험, 경제적위험, 시간적위험, 사용자만족, 지속사용의도	12
23	고령자를 위한 병원 안내 시스템 설계 및 사용성 평가	DEV·SVC-UX	이정원 외 (2012)	Learnability, Efficiency, Memorability, Errors, Satisfaction	5
24	서비스디자인 플랫폼을 위한 사용성 평가지표 연구	DEV·SVC-UX	정회준 외 (2015)	유효성, 효율성, 학습의 용이성, 기억의 용이성, 사용의 용이성, 사용의 유연성, 에러, 조작성, 매력성, 만족도, 사용자의 태도	11
25	비모수적 통계방법을 이용한 모바일 페이스북과 인스타그램의 사용자 경험 차이 비교	DEV·SVC-UX	안지현, 김승인 (2016)	Meaningful, Pleasurable, Convenient, Usable, Reliable, Functional	6
26	전자상거래 애플리케이션의 사용성 평가 지표	DEV·SVC-UX	최지원, 김종배 (2017)	Efficiency, Effectiveness, Flexibility, Accuracy	4
27	Determining and Validating Smart TV UX Factors	DEV·SVC-UX	Jang & Mun (2019)	Ease of Adaptation, Perceived Aesthetics, Cognitive Easiness, Contents Diversity, Customization Flexibility, Helpfulness, Applicability, Picture Quality, Playfulness, Responsiveness, Security, Social Relatedness, Sound Quality, Stability, 3D Quality, Voice Command Usability, Satisfaction, Usage Intention	18
28	디지털 교통안전표지판의 사용성평가 지표 개발 연구	DEV·SVC-UX	배예나 (2023)	[유용성(Usefulness)] (1) 지각성(Perceptibility) – 가시성(Visibility), 가독성(Legibility), 직관성(Intuitiveness) (2) 기능성(Functionality) 효용성(Utility), 일치성(Conformity), 보편성(Universality), 안전유발성(Safety Induction), 행동유도성(Affordance) [만족성(Satisfaction)] (1) 감성(Sensibility) 조화성(Reconciliability), 심미성(Aesthetic Quality) (2) 신뢰성(Trust) 수용성(Acceptability), 가치성(Value), 확장성(Extendability)	19
29	자율주행 셔틀 인포테인먼트 서비스 사용성 평가 지표 개발	DEV·SVC-UX	김윤경, 박선홍 (2023)	Effectiveness, Intuitiveness, Efficiency, Convenience, Learnability, Operability, Readability, Aesthetics, Satisfaction, Comfort	10
30	LLM 기반 교육용 챗봇의 사용성 평가 지표 개발 및 적용	DEV·SVC-UX	김지효 (2025)	명료성(Clarity), 투명성(Transparency), 맥락 적합성(Contextual Conformity), 최신성(Up to Dateness), 가시성(Visibility), 직관성(Intuitiveness), 환경적 접근성(Environmental Accessibility), 인지적 접근성(Cognitive Accessibility), 프라이버시 보호(Privacy Protection), 오류 관리(Error Management), 윤리성(Ethicality), 개인 맞춤성(Personalization), 과정 제시성(Presentation of Progress), 피드백(Feedback), 적응성(Adaptiveness), 공감(Empathy), 인간다운 자연스러움(Human Naturalness), 몰입성(Immersion), 학습 동기부여(Learning Motivation), 자기 주도성(Self Directedness), 이해 가능성(Understandability), 학습 효과성(Learning Effectiveness), 학습 효율성(Learning Efficiency), 학습 신뢰성(Learning Credibility), 학습 흥미성(Learning Interest), 심미성(Aesthetics), 친밀감(Intimacy), 자기효능감(Self Efficacy), 만족도(Satisfaction), 태도(Attitude), 지속 사용 의도(Continuous Intention)	31

[Appendix 2] Survey Questionnaire Items for the 25 UX Evaluation Sub-indicators

지표	설문
인지용이성 Ease of perception	OTT 서비스의 사용 방식과 과정을 쉽게 인지하고 이해할 수 있나요?
조작용이성 Ease of operation	OTT 서비스의 사용 방식과 과정을 물리적으로 쉽게 조작할 수 있나요? (서비스 내 메뉴, 버튼을 통한 조작과 리모컨, 휴대기기 등 외부 기기 조작을 포함)
탐색용이성 Ease of navigation	OTT 서비스에서 원하는 콘텐츠를 쉽고 빠르게 찾을 수 있나요?
원활성 Seamlessness	OTT 서비스가 과업 간, 기기 간, 다중 모드에서 끊김 없고 자연스러운 사용자 경험을 제공하나요? (서비스를 사용할 때, 스마트폰/TV 등 여러 기기를 오갈 때, PIP 모드를 사용할 때 등을 포함)
지원성 Supportiveness	OTT 서비스가 서비스 사용 중 발생한 문제를 돕고, 순조로운 사용을 위한 정보 및 기능을 지원하고 있나요?
정보전달성 Informativeness	OTT 서비스가 모든 관련 정보를 명확하고 효과적으로 전달하고 있나요? (서비스, 콘텐츠 관련 모든 정보)
콘텐츠 개인화 Contents Personalization	OTT 서비스가 나의 취향과 니즈에 따라 콘텐츠를 개인화하여 제공하고 있나요?
서비스 개인화 Service Personalization	OTT 서비스가 나에게 맞춰 편리하게 개인화되어 있나요?
사용자 맞춤형 유연성 Customization Flexibility	OTT 서비스를 내가 원하는 대로 설정할 수 있나요? (자막, 재생 속도 등 시청 환경과 알림, 프로필 관리 등 서비스 자체 설정 포함)
콘텐츠 다양성 Contents Diversity	OTT 서비스가 다양한 콘텐츠를 제공하고 있나요? (국가, 장르 등 여러 종류의 콘텐츠를 다양하게 제공하고 있는지)
콘텐츠 최신성 Contents Recency	OTT 서비스가 최신 콘텐츠를 제공하고 있나요?
콘텐츠 화질 Contents Video Quality	OTT 서비스가 제공하는 콘텐츠의 화질은 뛰어난가요?
콘텐츠 음질 Contents Audio Quality	OTT 서비스가 제공하는 콘텐츠의 음질은 뛰어난가요?
콘텐츠 번역 품질 Contents Localization Quality	OTT 서비스가 제공하는 다국어 콘텐츠의 자막 번역이 자연스럽고 정확한가요?
신뢰성 Trustworthiness	OTT 서비스의 전반적인 기능을 신뢰할 수 있다고 느끼나요? (재생, 검색, 추천, 연결, 연령별 접근 제한 등의 기능이 안정적으로 작동하는지)
안정성 Stability	OTT 서비스가 안정적으로 작동하나요? (멈춤, 끊김, 접속 오류, 버퍼링 등이 없이 잘 작동하는지)
반응성 Responsiveness	OTT 서비스의 전반적인 반응 속도가 빠른가요?
보안 Safety	OTT 서비스가 나의 개인 정보를 안전하게 관리하고 프라이버시를 보장한다고 느끼나요?
유희성 Pleasure	OTT 서비스의 전반적인 사용경험이 나에게 유희나 즐거움을 제공하나요?
몰입성 Immersion	OTT 서비스를 몰입하여 사용하면서 시간을 보내고 지루함을 해소하나요?
심미성 Aesthetics	OTT 서비스의 전반적인 인터페이스는 시각적으로 만족스럽나요? (화면 구성, 시각적 일관성, 브랜드 이미지 등)
사회적 관계성 Social Relatedness	OTT 서비스 사용을 통해 타인과의 상호작용에 대한 만족감을 느끼나요? (실시간 채팅, 왓챠파티 등 플랫폼 내 기능을 사용한 소통)
만족도 Satisfaction	OTT 서비스의 전반적인 사용 경험에 만족하나요?
지속성 Stickiness	OTT 서비스를 장기적으로 이용·유지하고 시청할 의향이 있나요?
추천 의도 Word-of-Mouth Intention	OTT 서비스를 다른 사람에게 추천할 의향이 있나요?

[Appendix 3] Descriptive Statistics and Friedman Test Results for OTT App Services

평가 지표		대상 서비스			
		아마존프라임	넷플릭스	티빙	왓챠플레이
인지용이성 (Ease of perception)	평균 (표준편차)	3.7 (1.06)	4.3 (0.68)	4.2 (0.63)	3.5 (1.08)
	평균 순위	2.25	3	2.9	1.85
	유의확률		0.056		
조작용이성 (Ease of operation)	평균 (표준편차)	3.7 (0.95)	4.4 (0.52)	4.4 (0.70)	4 (0.67)
	평균 순위	1.95	2.85	2.8	2.4
	유의확률		0.1088		
탐색용이성 (Ease of navigation)	평균 (표준편차)	4.1 (0.74)	4.5 (0.71)	4 (0.82)	4.1 (0.74)
	평균 순위	2.35	3	2.3	2.35
	유의확률		0.3916		
평균		3.83	4.4	4.2	3.87
원활성 (Seamlessness)	평균 (표준편차)	3.5 (1.35)	4.2 (1.03)	3.8 (1.03)	3.7 (1.25)
	평균 순위	2.15	2.9	2.35	2.6
	유의확률		0.4201		
지원성 (Supportiveness)	평균 (표준편차)	3.2 (0.79)	3.7 (0.82)	3.1 (0.74)	3.1 (0.88)
	평균 순위	2.5	2.95	2.25	2.3
	유의확률		0.3521		
정보전달성 (Informativeness)	평균 (표준편차)	4.1 (1.20)	4 (0.94)	3.9 (0.74)	3.2 (1.48)
	평균 순위	2.95	2.5	2.6	1.95
	유의확률		0.1848		
평균		3.6	3.97	3.6	3.33
콘텐츠 개인화 (Contents Personalization)	평균 (표준편차)	3.2 (0.79)	4.1 (0.74)	3.8 (0.92)	4 (0.82)
	평균 순위	1.75	2.9	2.55	2.8
	유의확률		0.066		
서비스 개인화 (Service Personalization)	평균 (표준편차)	3.1 (0.88)	3.8 (0.92)	3.6 (0.52)	3.4 (0.70)
	평균 순위	1.85	2.9	2.75	2.5
	유의확률		0.096		
사용자 맞춤형 유연성 (Customization Flexibility)	평균 (표준편차)	3.5 (1.08)	4.4 (0.70)	4.4 (0.84)	3.5 (0.85)
	평균 순위	1.9	3.05	3	2.05
	유의확률		*0.0373		
평균		3.27	4.1	3.93	3.63
콘텐츠 다양성 (Contents Diversity)	평균 (표준편차)	3.4 (0.97)	4.5 (0.53)	3.3 (0.82)	4.2 (0.42)
	평균 순위	2	3.3	1.8	2.9
	유의확률		*0.0021		
콘텐츠 최신성 (Contents Recency)	평균 (표준편차)	3.7 (0.68)	4.7 (0.48)	4.3 (0.68)	3.9 (0.57)
	평균 순위	1.85	3.4	2.7	2.05
	유의확률		*0.0031		
콘텐츠 화질 (Contents Video Quality)	평균 (표준편차)	4.5 (0.71)	4.3 (0.68)	4.2 (0.79)	4.1 (0.88)
	평균 순위	2.85	2.55	2.4	2.2
	유의확률		0.1989		
콘텐츠 음질 (Contents Audio Quality)	평균 (표준편차)	4.6 (0.84)	4.1 (0.74)	4.2 (0.63)	4.1 (0.57)
	평균 순위	3.15	2.25	2.4	2.2
	유의확률		0.0669		

콘텐츠 번역 품질 (Contents Localization Quality)	평균 (표준편차)	4.1 (1.10)	4.1 (0.74)	3.5 (0.53)	3.6 (0.84)
	평균 순위	2.95	2.95	1.95	2.15
	유의확률	0.0508			
평균		4.06	4.34	3.9	3.98
신뢰성 (Trustworthiness)	평균 (표준편차)	3.6 (0.97)	4.1 (0.74)	4 (0.67)	3.9 (0.88)
	평균 순위	2.1	2.7	2.65	2.55
	유의확률	0.635			
안정성 (Stability)	평균 (표준편차)	4.3 (0.95)	4.2 (1.03)	4.1 (0.99)	4.1 (0.74)
	평균 순위	2.7	2.65	2.35	2.3
	유의확률	0.6444			
반응성 (Responsiveness)	평균 (표준편차)	4.3 (0.68)	4.5 (0.71)	4 (1.05)	4 (0.82)
	평균 순위	2.6	2.95	2.3	2.15
	유의확률	0.1962			
보안 (Safety)	평균 (표준편차)	3.3 (0.68)	3.3 (0.68)	3.3 (0.68)	3.3 (0.82)
	평균 순위	2.5	2.5	2.5	2.5
	유의확률	1			
평균		3.88	4.03	3.85	3.83
유희성 (Pleasure)	평균 (표준편차)	3.6 (0.97)	4.1 (0.99)	3.9 (0.99)	3.9 (0.88)
	평균 순위	2	2.9	2.55	2.55
	유의확률	0.1806			
몰입성 (Immersion)	평균 (표준편차)	3.4 (1.51)	3.8 (1.40)	3.7 (1.49)	3.5 (1.18)
	평균 순위	2.2	2.7	2.6	2.5
	유의확률	0.7262			
심미성 (Aesthetics)	평균 (표준편차)	2.7 (0.95)	3.8 (1.14)	3.4 (1.17)	3.2 (1.14)
	평균 순위	2	2.95	2.65	2.4
	유의확률	0.309			
평균		3.23	3.9	3.67	3.53
사회적 관계성 (Social Relatedness)	평균 (표준편차)	1.8 (1.14)	2.1 (1.20)	2.9 (1.10)	3.5 (1.35)
	평균 순위	1.65	2.15	2.95	3.25
	유의확률	*0.0093			
평균		1.65	2.15	2.95	3.25
만족도 (Satisfaction)	평균 (표준편차)	3.2 (0.63)	3.9 (0.74)	4 (0.67)	3.5 (0.85)
	평균 순위	2	2.75	2.95	2.3
	유의확률	0.2069			
지속성 (Stickiness)	평균 (표준편차)	2.5 (1.18)	3.6 (1.51)	3.9 (1.20)	3.3 (0.95)
	평균 순위	1.6	2.95	3.15	2.3
	유의확률	*0.0087			
추천 의도 (Word-of-Mouth Intention)	평균 (표준편차)	2.6 (0.84)	3.8 (1.14)	3.8 (1.14)	3.4 (1.26)
	평균 순위	1.6	3	2.95	2.45
	유의확률	*0.0135			
평균		2.77	3.77	3.9	3.4

[Appendix 4] Mann–Whitney U Test Results by OTT App Service

사용성 (Usability)													
주지표	상세지표	아마존프라임–넷플릭스		아마존프라임–티빙		아마존프라임–왓챠		넷플릭스–티빙		넷플릭스–왓챠		티빙–왓챠	
		M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률
사용편의성 (Ease of Use)	인지용이성 (Ease of perception)	34	0.203	37.5	0.316	56	0.664	54.5	0.734	72	0.087	69.5	0.127
	조작용이성 (Ease of operation)	28	0.077	28.5	0.092	41	0.487	48	0.898	66	0.178	66	0.199
	탐색용이성 (Ease of navigation)	34.5	0.217	53.5	0.808	50	1.000	67.5	0.166	65.5	0.217	46.5	0.808
인터랙션 적합성 (Interaction Suitability)	원활성 (Seamlessness)	34.5	0.235	44.5	0.695	46.5	0.814	62	0.360	61	0.401	52	0.906
	지원성 (Supportiveness)	33	0.175	54	0.775	52	0.901	71	0.096	70	0.109	47.5	0.866
	정보전달성 (Informativeness)	57	0.598	61.5	0.378	69.5	0.136	56	0.653	66	0.221	62.5	0.343
상황 인지 특성 (Context–Awareness)	콘텐츠 개인화 (Contents Personalization)	22	***0.027**	34	0.212	26	0.060	60.5	0.423	53.5	0.808	43	0.601
	서비스 개인화 (Service Personalization)	29.5	0.096	29	0.090	36.5	0.286	54	0.773	60	0.440	57	0.576
	사용자 맞춤형 유연성 (Flexibility)	25.5	0.057	26	0.062	50	1.000	48	0.899	78.5	***0.025**	77	***0.035**
콘텐츠 품질 (Contents Quality)	콘텐츠 다양성 (Contents Diversity)	17.5	***0.010**	52.5	0.872	25	***0.036**	87.5	***0.003**	65	0.185	20	***0.009**
	콘텐츠 최신성 (Contents Recency)	14	***0.004**	27.5	0.071	41	0.460	66.5	0.168	83	***0.006**	66.5	0.167
	콘텐츠 화질 (Contents Video Quality)	59	0.476	61	0.384	63	0.303	53	0.837	56	0.655	53	0.840
	콘텐츠 음질 (Contents Audio Quality)	70	0.102	69	0.119	73	0.062	46.5	0.800	50.5	1.000	54.5	0.720
	콘텐츠 번역 품질 (Contents Localization Quality)	53.5	0.810	70	0.120	66	0.221	72.5	0.068	66	0.205	45	0.706

사용자 가치 (User Value)													
주지표	상세지표	아마존프라임–넷플릭스		아마존프라임–티빙		아마존프라임–왓챠		넷플릭스–티빙		넷플릭스–왓챠		티빙–왓챠	
		M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률
기능적 가치 (Functional Value)	신뢰성 (Trustworthiness)	34.5	0.231	37	0.313	39.5	0.422	54	0.769	55	0.707	51	0.966
	안정성 (Stability)	52	0.901	56.5	0.622	60.5	0.411	54	0.776	57	0.598	53	0.839
	반응성 (Responsiveness)	41	0.476	56.5	0.627	60.5	0.415	63.5	0.286	67.5	0.166	52	0.905
	보안 (Safety)	46	0.754	46	0.754	45.5	0.737	50	1.000	48.5	0.934	48.5	0.934
감정적 가치 (Emotional Value)	유희성 (Pleasure)	34.5	0.221	41.5	0.515	41.5	0.493	56.5	0.631	58	0.538	50.5	1.000
	몰입성 (Immersion)	41	0.500	43.5	0.623	49.5	1.000	52.5	0.873	59	0.504	57	0.603
	심미성 (Aesthetics)	23	***0.039**	33.5	0.210	36	0.286	60	0.455	65.5	0.240	54	0.785
사회적 가치 (Social Value)	사회적 관계성 (Social Relatedness)	41.5	0.515	22.5	***0.035**	17.5	***0.012**	29.5	0.115	22	***0.034**	35	0.259

사용자 수용도 (User Acceptance)												
주지표	아마존프라임–넷플릭스		아마존프라임–티빙		아마존프라임–왓챠		넷플릭스–티빙		넷플릭스–왓챠		티빙–왓챠	
	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률	M/W U	유의확률
만족도 (Satisfaction)	25.5	***0.049**	21	***0.019**	39.5	0.408	46	0.769	63	0.310	67	0.175
지속성 (Stickiness)	27.5	0.086	20	***0.020**	30.5	0.132	46	0.778	64	0.279	71	0.094
추천 의도 (Word–of–Mouth Intention)	16	***0.009**	16	***0.009**	29	0.106	50	1.000	61	0.401	61	0.401

OTT 서비스의 사용자 경험 평가 지표 개발 및 적용

박서하[1], 강효진[2*]

[1]성신여자대학교 대학원 미래융합기술공학과, 석사졸업생, 서울, 대한민국

[2]성신여자대학교 서비스디자인공학과, 부교수, 서울, 대한민국

초록

연구배경 OTT(Over-the-Top) 서비스가 다양한 디바이스와 사용 맥락 전반에서 미디어 소비의 핵심 플랫폼으로 자리 잡으면서, 사용자 경험(UX)이 서비스 유지와 장기적 경쟁력을 좌우하는 핵심 요인으로 부상하였다. 그러나 기존 연구는 기술 수용, 만족도, 지속 이용 의도 등 개별 변인 중심 접근에 머물러, 콘텐츠 탐색·개인화·경험적 가치가 결합된 OTT 서비스의 복합적·맥락적 사용자 경험을 통합적으로 평가하는 데 한계가 있었다. OTT 서비스가 구독 기반의 콘텐츠 중심 플랫폼이라는 점에서, 사용성·경험적 가치·장기적 수용을 함께 포괄하는 통합적 UX 평가 틀이 요구된다.

연구방법 본 연구는 OTT 서비스 맥락에 적합한 사용자 경험 평가 지표 체계를 개발하고 실증적으로 검증하였다. 기술수용모델(TAM), 정서–행동–인지(ABC) 모델, 자극–유기체–반응(SOR) 모형이 공통적으로 전제하는 인지–정서–행동 구조를 통합하여, 사용성–사용자 가치–사용자 수용도로 구성된 UVA 프레임워크를 상위 분석 틀로 설정하였다. 문헌 통합, 사용자 관찰 및 인터뷰, 전문가 검토, 설문 기반 검증의 네 단계를 거쳐 3개 차원, 10개 주지표, 25개 세부 지표로 구성된 지표 체계를 도출하였다. OTT 이용 경험이 있는 사용자 347명의 설문 응답을 바탕으로 탐색적 요인분석(EFA)을 수행하여 사용성 및 사용자 가치 차원의 구조적 타당성을 검토하였으며, 만족도·지속성·추천 의도로 구성된 사용자 수용도 차원은 각 지표가 단일 문항으로 측정되어 EFA 대상에서 제외하였다. 이후 네 개 OTT 서비스를 대상으로 지표를 소규모 예비 적용하고, 프리드먼 검정과 만–휘트니 U 검정 등 비모수 통계 분석을 통해 서비스 간 UX 차이를 탐색적으로 비교하였다.

연구결과 탐색적 요인분석 결과, 사용성 및 사용자 가치 차원의 지표들은 UVA 이론 구조에 부합하는 요인 구조를 형성하였으며, 신뢰도 분석에서도 대체로 수용 가능한 수준 이상의 내적 일관성이 확인되었다. 특히 문헌이 아닌 사용자 관찰 과정에서 도출된 콘텐츠 탐색용이성과 콘텐츠 번역 품질이 실제 이용 맥락의 핵심 UX 요소로 확인되었다. 소규모 예비 적용 결과에서는 사용자 맞춤형 유연성, 콘텐츠 다양성, 콘텐츠 최신성, 사회적 관계성, 지속성, 추천 의도에서 서비스 간 통계적으로 유의한 차이가 나타나, 제안된 지표 체계의 예비적 변별 가능성을 시사하였다.

결론 본 연구는 OTT 서비스의 맥락적·다차원적 경험을 반영한 OTT 특화 UX 평가 지표 체계를 제안하고, 사용성 및 사용자 가치 차원을 중심으로 구조적 정합성을 실증적으로 검토하였다. EFA 기반 검증이 두 차원에 한정되었으나, 제안된 지표 체계의 구조적 일관성에 대한 실증적 근거를 확보하였다. 나아가 평가 결과를 사용성 중심의 디자인 가이드라인으로 확장함으로써, UX 평가를 단순한 측정 도구를 넘어 서비스 진단·기획·설계 개선을 위한 실천적 틀로 자리매김한다는 점에서 학술적·실무적 의의를 지닌다.

주제어 OTT 서비스, 사용자 경험 평가 지표, UVA 프레임워크, 탐색적 요인 분석, 서비스 비교, 사용자 관찰 연구, 디자인 가이드라인

본 연구는 저자의 석사학위논문 「OTT 서비스 사용자 경험 평가 지표 개발 및 적용」(성신여자대학교, 서울, 대한민국)을 확장한 것이다.

또한 이 논문은 교육부와 한국연구재단의 인문사회분야 신진연구자지원사업(NRF-2023S1A5A8076896)과 2026년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구이다.

(RS-2025-02263062, 2026년 디지털융합 디자인 전문인력양성사업)

*교신저자 : 강효진 (hjkang@sungshin.ac.kr)