

## 유튜브 스포츠 영상 콘텐츠의 소비 결정 요인 : 다층회귀모형을 통한 빅데이터 분석

김종호(서울대학교, 강사) · 김기한\*(서울대학교 체육교육과/스포츠과학연구소, 교수)

본 연구는 유튜브 스포츠 영상 콘텐츠의 소비 결정 요인을 분석하였다. 스포츠를 주요 소재로 영상을 제작하고 게시하는 124개의 유튜브 채널과 여기에 게시된 89,834개의 영상 콘텐츠를 빅데이터로 구성하여 분석하였다. 스포츠 영상 소비 행동을 유도할 수 있는 독립변수를 영상 콘텐츠 수준과 유튜브 채널 수준에서 구성하여 조회수에 미치는 영향을 다층모형 회귀분석(Multi-Level Regression)을 통하여 분석하였다. 구체적으로 영상 콘텐츠 수준의 변수로는 스포츠 요인(종목, 콘텐츠 유형), 콘텐츠 크리에이티브 요인(영상, 썸네일), 접근성 요인(검색어 수, 타 SNS 링크 수)이 있었으며, 채널 수준의 변수에는 제작자 요인(유형, 신뢰성, 전문성, 성별)과 운영 요인(보유 콘텐츠 수, 업로드 주기, 소재의 다양성)을 설정하였다. 주요 분석 결과로 콘텐츠 수준의 변수 중 스포츠 요인이 조회수에 미치는 영향이 유의미하였으며 동시에 영상 정보의 양과 썸네일의 형태, 콘텐츠 접근성이 조회수에 영향을 미치는 것으로 나타났다. 그리고 채널 수준의 변수의 영향력은 제작자의 성별과 유형, 인지도 신뢰성, 채널 운영 형태와 기간이 조회수에 유의한 영향을 미치는 것으로 나타났다. 이와 같은 결과는 유튜브 스포츠 채널 운영자 입장에서 콘텐츠를 제작하고 채널을 운영할 때 높은 조회수를 획득하기 위한 효과적인 커뮤니케이션 전략 수립에 필요한 정보를 제공한다는 점에서 의미가 있다. 또한 기존의 스포츠 미디어 연구를 유튜브 플랫폼에 확장하고 빅데이터 분석을 진행하였다는 점에서 학술적인 의미가 있다.

※ 주요어 : 유튜브, 스포츠 콘텐츠, 스포츠 미디어, 빅데이터, 다층모형 회귀분석, 웹크롤링

---

\* 교신저자 : kihan@snu.ac.kr

## 서론

### 1. 연구 배경 및 필요성

닐슨 코리아의 보고서 '유튜브 저널리즘'에 의하면 2019년 유튜브 플랫폼의 월평균 이용자 수는 약 19억 명에 이르며, 이들은 매일 10억 시간 분량의 동영상을 시청한다(닐슨, 2020). 개인 미디어와 모바일 스트리밍 기술이 보편화되면서 유튜브는 다양한 콘텐츠 생산과 소비문화를 창출하며 온라인 영상 소비 시장을 견인하고 있다(Khan, 2017).

'스포츠'는 유튜브의 핵심 콘텐츠 장르 중 하나이다. 2021년 4월 기준 세계 구독자 보유수 상위 10개의 유튜브 채널 중 2개(WWE: 7,620만, Dude Perfect: 5,590만)가 스포츠를 주요 소재로 한 채널이라는 점은 이를 뒷받침한다(뉴스 인플루언서, 2021. 3). 이와 같은 경향은 국내 유튜브도 유사한데, 2021년 4월 기준 피트니스 채널인 피지컬 갤러리는 306만 구독자, AllblancTV는 165만 구독자 수를 보유하고 있다. 해외에 대비하여 국내 스포츠 채널의 구독자 수는 상대적으로 작은 수치로 보이지만, 한국어 언어권의 시청자 수와 영문권 시청자 수의 절대적인 차이를 고려해 보았을 때, 이와 같은 수치는 결코 적은 수가 아니다.

이처럼 유튜브의 핵심 장르로 성장한 유튜브 스포츠 콘텐츠는 기성 스포츠 채널, 스포츠 구단 및 협회뿐만 아니라 유튜브를 전문으로 하는 MCN(Multi Channel Network) 사업자와 일반 개인에 이르기까지 다양한 참여자들이 스포츠 영상 콘텐츠를 제작하여 유튜브에 유통하고 있다. 자연히 유튜브 스포츠 영상 제작자들 사이에 시청자들의 시간과 관심을 확보하기 위한 경쟁이 치열하다.

유튜브 초창기 시절의 영상이 일반인이 제작한 비전문적 콘텐츠가 주를 이루었다면, 최근 유튜브 영상은 보다 전문적이고 전략적으로 기획 및 제작된다는 특징이 있다. 많은 유튜브 영상 제작자들은 시청자들의 관심과 소비를 이끌어내기 위해 가장 효과적인 영상 제작 요인을 사전에 분석하여 기획 단계부터 적용함으로써 조회수 극대화를 위한 노력을 한다. 예컨대 콘텐츠 장르와 내용에 따라 가장 적절한 키워드, 썸네일의 이미지와 글자 수,

최적의 영상 길이 등에 대한 노하우를 축적하여 영상 제작에 적용하고 있다.

스포츠 영상 제작도 이와 같은 전략적인 접근이 필요하다. 하지만 유튜브 스포츠 영상 콘텐츠 소비를 결정하는 주요 요인에 대한 연구는 아직까지 학계에 보고된 바 없다. 관련된 선행 연구는 대부분 TV를 중심으로 한 전통 미디어 소비에 대한 연구로 미디어스포츠 소비 동기에 관한 연구(김원제, 2016; 정영남, 2008; Raney, 2003, 2006; Gantz & Wenner, 1995; Shank & Lyberger, 2014)와 스포츠 TV 시청률 결정 요인에 관한 연구(Alavy, Gaskell, Leach & Szymanski, 2010; Paul & Weinbach, 2007; Raney, 2013)가 주를 이룬다. 이러한 선행 연구는 큰 틀에서는 스포츠미디어 소비의 영향요인을 이해하는 기초 정보를 제공한다는 점에서 유용하지만, 다른 한편으로는 유튜브 플랫폼 특유의 소비 환경에 대한 고려가 없다는 점에서 유튜브 스포츠 영상 콘텐츠의 소비 영향요인을 설명하기에는 충분하지 않다는 한계가 있다.

본 연구는 이와 같은 선행 연구의 한계를 보완하기 위하여 유튜브에 게시된 스포츠 영상 콘텐츠의 소비 결정 요인을 분석하고자 한다. 이를 위해 전통적 커뮤니케이션 이론을 바탕으로 유튜브 미디어 환경에서 스포츠 영상 소비 행동에 영향을 미칠 수 있는 변수를 영상 콘텐츠 수준과 유튜브 채널 수준에서 구성하고, 124개 유튜브 스포츠 채널의 89,834개의 영상 자료를 빅데이터로 구성하여 다층회귀모형을 분석하였다.

기존 체육학 연구에서 스포츠미디어에 대한 연구는 활발히 이루어져 왔다. 하지만 기성 스포츠 미디어 연구가 TV를 중심으로 한 전통 미디어 중심의 연구에 한정되었다는 한계가 있는 반면, 본 연구는 이와 같은 전통적 스포츠미디어 연구를 새롭게 각광받는 유튜브 스포츠 콘텐츠의 영역으로 확장함으로써 기성 체육학 연구를 통하여 다루어지지 않았던 신매체 기반의 스포츠 미디어 소비 행동에 대한 이해도를 넓힐 수 있다는 점에서 의미가 크다. 유튜브로 대변되는 디지털 미디어 소비 환경에서의 스포츠미디어 소비에 관한 연구는 전통 미디어를 기반으로 수행된 선행연구와 차별화되는 연구의 학술적·실무적 의미가 있다.

## 2. 이론적 배경

### 1) 유튜브 스포츠 콘텐츠와 소비

유튜브 스포츠 콘텐츠는 유튜브 플랫폼 상에서 스포츠를 주제 혹은 중심 소재로 제작하여 소비자에게 제공하는 영상 콘텐츠를 의미한다. 유튜브에서 분류한 주요 15개 콘텐츠 장르 중 하나인 유튜브 스포츠 콘텐츠는 경쟁 스포츠뿐만 아니라 라이프스타일 스포츠를 포괄하는 넓은 영역의 스포츠 주제를 다루고 있다(Stauff, 2009). 유튜브에서 스포츠는 복미에서 스포츠를 정의하기 위해 사용한 개념인 경쟁, 신체적 움직임, 규칙, 리그의 4대 요소를 포함해야만 하는 협소한 개념이 아닌 건강과 즐거움을 위한 신체적 움직임으로서 스포츠의 광의적인 개념으로 영역이 형성되어 있다(Stauff, 2009).

미디어 환경에 따라 변화한 오늘날의 유튜브 스포츠 콘텐츠는 그 유형과 형태가 매우 다양하게 분화되어 활발히 소비되고 있다. 유튜브 스포츠 콘텐츠는 신체활동, 스포츠 경기, 스포츠 스타, 시설, 용품 등과 같은 다양한 소재를 이용하고 있으며 목적에 따라 정보 제공, 오락 및 예능, 교육 등과 같은 다양한 형태로 콘텐츠가 제작되고 있다. 나이가 종목에 따라 소재와 유형은 더욱 세부적으로 분화된다(Burgess & Green, 2018; Stauff, 2009).

다양한 콘텐츠들이 게시된 유튜브 플랫폼의 웹 인터페이스(Web interface)에서 소비자들은 이용 동기와 취향에 맞는 콘텐츠를 선택하여 시청한다. 그리고 이러한 시청 행동은 유튜브의 추천시스템이나 구독시스템에 의해서 강화되는 경향을 보이기도 한다. 특히 취향 콘텐츠인 스포츠 콘텐츠는 이와 같은 현상이 더욱 강하게 발현되는 특징을 보이기도 한다. 이와 관련하여 Duffett, Petroșanu, Negricea 그리고 Edu(2019)는 유튜브 플랫폼에서의 미디어 소비 행동은 플랫폼을 이용하는 과정 전반에서 다양한 영향요인에 의해 변화할 수 있다고 주장하였다. 구체적으로 웹 인터페이스(Web interface)에서 콘텐츠를 선택할 때와 플레이어 인터페이스(Player interface)에서 시청을 지속할 때 인지적(cognitive), 감정적(affective), 행동적(behavior) 영향요인이 있음을 보고 하였다.

위계적 커뮤니케이션 이론(Hierarchy of effect

communication model)를 응용한 선행연구(Duffett et al., 2019; Gupta, Singh, & Sinha, 2017)에 의하면 첫 번째 인지영역(cognitive)에서 일어나는 관심(attention)과 흥미(interest)를 유도하는 데는 정보를 제공하는 정보원(source)과 메시지의 내용이 중요한 영향요인으로 작용한다. 그리고 두 번째 확인과 공감 단계에서 소비자들은 메시지(message)의 내용과 구성을 통해 자신의 기호와 동기에 적합한 콘텐츠인지를 확인하는 행동을 한다. 이러한 두 가지 조건이 충족된 이후 세 번째 욕구(desire)의 단계에서 시청만족이라는 감정이 생성된다. 그리고 마지막 행동(action) 단계에서 시청을 지속하거나, 연관 콘텐츠의 검색(search), 타인에게 추천(share)과 같은 시청 후 행동이 이어지게 된다.

이러한 배경을 근거로 본 연구에서는 이러한 커뮤니케이션 단계에 따른 소비자의 태도와 행동에 영향을 미치는 소비 영향요인을 도출하기 위해, 능동적인 소비자들의 정보처리 과정에서 근거로 활용되는 단서들을 파악하고자 시도하였다. 구체적으로 이를 위해 전통 커뮤니케이션 설득 모형인 Berlo(1960)의 SMCR(Sender, Message, Channel, Receiver) 모델을 기준으로 요인들을 도출하였다. <그림 1>과 같이 SMCR 모델은 콘텐츠의 제작자와 콘텐츠의 내용, 콘텐츠를 재생하는 채널의 특수성 그리고 소비자의 특성에 따라 미디어 효과가 상이하다는 것을 체계적으로 설명하는 전통 광고 커뮤니케이션 모델이다. 특히 SMCR 모델은 미디어 소비 행동에 미치는 다양한 영향요인을 도출할 근거를 제공해 준다는 점에서 의미가 있다.

우선 콘텐츠를 선택하는 환경인 웹 인터페이스에서는 소비자들의 콘텐츠 선택이 이뤄진다. 이때 콘텐츠 송신자(Sender)들은 소비자의 선택을 받기 위해 소비자의 눈

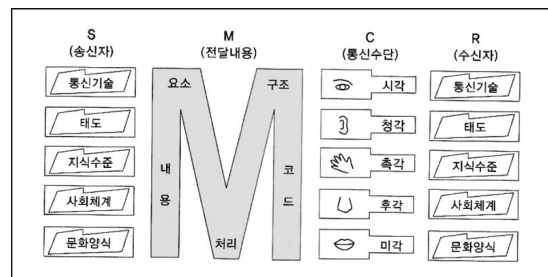


그림 1. SMCR 모델(Berlo, 1960)

길을 끌고 만족시킬 수 있는 콘텐츠를 제작하기 위한 다양한 크리에이티브(Creative) 전략을 구사한다. 한편 수신자(Receiver)들은 공신력 있는 송신자(Sender)가 게시한 콘텐츠 내용(Message)이 자신의 취향에 적합한지 그리고 이용 목적에 적합한 형태를 하고 있는지 등을 근거로 수많은 콘텐츠들 중 하나의 콘텐츠를 선택하게 된다. 선택된 콘텐츠가 재생되는 플레이어 인터페이스(Player interface)에서는 제공되는 영상의 내용(Message)이 기대했던 것과 일치하는지, 메시지를 전달하는 전달자(Sender)는 공신력 있고 매력이 있는지 등을 판단하며 지속적인 시청을 이어나간다. 메시지가 충분히 설득력 있어 소비자를 감화시키는 데 성공했다면 연관 동영상 시청하거나 댓글을 남기는 것과 같은 강한 시청 행동을 보이기도 한다. 이와 같은 현상은 유튜브 채널의 특수성(Channel)에 의해서 더욱 강화되는데, 시청자가 만족할 만한 콘텐츠를 추천해주거나, 타인에게 쉽게 공유할 수 있도록 하는 접근성 최적화 서비스를 대표적인 예시로 들 수 있다.

이처럼 유튜브 플랫폼에서 영상 콘텐츠 소비 행동은 커뮤니케이션 과정을 수행하고 있다고 볼 수 있다. 커뮤니케이션이란 메시지 송신자가 특정 목적을 달성하기 위해 정보를 메시지화 하여 수신자에게 전달하고 피드백(feedback) 받는 행위를 의미한다. 유튜브 플랫폼에서 콘텐츠 제작자들은 높은 성과지표를 달성하기 위해 콘텐츠를 제작하여 소비자에게 전달하고 있으며, 소비자들은 자신의 기호와 사용 동기에 적합한 콘텐츠를 시청하길 원한다. 특히 스포츠 콘텐츠 제작자들은 스포츠 장르의 내용적 특성과 스포츠 소비자의 특수성을 고려하여 소비자의 선택과 시청 만족을 유도해 오고 있다. 유튜브 플랫폼에서 스포츠 미디어 커뮤니케이션은 1) 콘텐츠 제작자라는 주체가 2) '소비자(Target Audience)'에게 스포츠와 관련된 3) 정보를 제공하고 지속적이고 강한 시청 행동을 설득하기 위해(To inform and persuade) 4) 다양한 제작 전략을 이용하여 시행하는 5) 비대면(Non-personal)적 커뮤니케이션 특성을 가지고 있다. 이때 정보의 원천(Sender/Source)과 내용(Message/Contents)은 커뮤니케이션 과정에서 매우 중요한 요소임을 알 수 있다.

이어지는 장에서는 유튜브 스포츠 콘텐츠 소비에 영향

을 미치는 영향요인을 체계적으로 분류하고, 각각의 영향요인이 소비에 미치는 영향의 기전에 대하여 설명하였다.

## 2) 유튜브 스포츠 소비 영향요인

앞서 언급된 바와 같이 유튜브 상에서 소비자들은 콘텐츠를 소비하는 과정에서 다양한 요인에 의해서 영향을 받게 된다. 본 장에서는 이와 같은 소비 행동에 영향을 미치는 영향요인을 구분하고, 각각의 영향요인이 특별히 스포츠 콘텐츠의 소비에 영향을 미치는 과정에 대해 기술하였다. 유튜브 스포츠 콘텐츠 소비 행동에 영향을 미치는 요인은 앞서 언급된 'SMCR 모델'을 근거로 콘텐츠 수준에서 형성되는 요인과 채널 수준에서 형성되는 요인을 체계적으로 구분하여 설정하였다. <그림 2>는 본 연구에서 설정된 유튜브 스포츠 콘텐츠 소비 요인의 체계를 요약하여 보여준다. 구체적으로 콘텐츠 수준의 소비 영향요인은 스포츠 요인, 콘텐츠 크리에이티브(Creative) 요인, 접근성 요인이 있으며, 채널 수준의 소비 영향요인은 제작자 요인과 운영 요인이 있다.

콘텐츠 수준의 소비 영향요인의 세부 내용은 다음과 같다. 첫 번째 콘텐츠 수준의 소비 영향요인인 스포츠 요인은 스포츠 콘텐츠의 메시지의 내용적 특성에서 발생하는 소비 영향요인을 의미한다. 특히 본 연구에서는 '스포츠 종목'과 '스포츠 콘텐츠 유형'을 콘텐츠 수준의 소비 영향요인(스포츠 요인)으로 설정하였다. 스포츠 콘텐츠 영역에서 종목은 일종의 주제로써 콘텐츠의 특성을 대표하는 속성이 될 수 있기 때문이다. 스포츠 콘텐츠의 중심 주제를 의미하는 종목은 서로 상이한 시청 지형을 형성하는 기초가 된다는 점에서, 유튜브 스포츠 콘텐츠 시청 결정 요인으로 고려될 수 있다. 우선 스포츠 종목이 서로 상이한 시청 지형을 형성하는 이유는 종목마다 서로 다른 네러티브 구조를 가지기 때문이다. McLean, Hurd 그리고 Rogers(2005)는 따라서 서로 상이한 내용을 가지고 있으며, 이에 따라 서로 다른 형태의 즐거움이 발생한다고 하였다. 이러한 현상은 미디어스포츠에서는 더욱 특별히 상이하게 발현될 수 있음을 선행 학자들은 주장하였다(박보현, 2008; 김기한, 윤리라, 방신용, 2012; Birrell & Loy, 1979). 따라서 본 연구에서는 스포츠 종목을 스포츠 요인의 하위 요인으로서 설정하게 되었다.

또한 본 연구에서는 종목뿐만 아니라 ‘스포츠 콘텐츠의 유형’, 즉, 정보가 전달되는 방식이 유튜브 콘텐츠 소비 행동에 영향을 미칠 것으로 보았다. 이는 Ramaprasad와 Hasegawa(1992)가 주장한 설득 커뮤니케이션에서의 메시지 실행 요인(execution)이 청자의 태도와 행동을 변화시킨다는 논리에 근거한 것이다. 메시지의 실행요인은 특정 소재를 어떤 방식으로 전달하는가와 관련되는데, 본 연구에서는 스포츠라는 메시지를 전달하는 방식에 초점을 두었다. 즉 스포츠라는 주제영역에서 특정 종목이라는 소재를 선택한 후 특정 목적을 달성하기 위해 메시지를 구성하는 것이다. 콘텐츠 유형은 일종의 소비를 유도하는 설득전략이라 할 수 있는데, 설득전략은 소비자의 태도를 변화시키기 위한 방향으로 설정되는 특징을 가진다. 즉 동일한 종목이라 하더라도 제작자의 목적에 따라 정보성을 가진 콘텐츠로 제작될 수 있는 반면 감정적인 즐거움을 주기 위한 콘텐츠로 제작될 수도 있다. 본 연구에서는 Raney(2006)가 제시한 스포츠 미디어 시청 동기를 기초로 스포츠 콘텐츠 유형을 1)레슨 및 정보형 영상, 2) 오락 및 엔터테인먼트 영상, 3) 경기 및 하이라이트 영상, 4) 프로스포츠 정보 및 분석 영상, 5) 커뮤니티형 영상, 6) 기타 유형 등으로 구분하였다.

두 번째 스포츠 콘텐츠 소비 영향요인으로는 콘텐츠 크리에이티브 요인이 있다. 콘텐츠 크리에이티브 요인은 영상과 썸네일로 구성된 영상 콘텐츠가 어떠한 내용과 형식으로 제작되었는가와 관련된 변수이다. 크리에이티브

요인은 광고 연구에서 사용하는 용어로 특정 광고의 내용과 소통 형식에 관한 제작 방향과 전략을 의미한다(이주호, 김기한, 2019; Stewart & Furse, 1986). 예컨대 광고에 유명인이 등장하는지 여부, 광고 카피의 내용 등은 광고의 주요 크리에이티브 요인에 해당하며, 이는 광고 소비와 효과에 결정적인 영향을 미친다(Chaffey & Smith, 2017; Erdogan, 1999). 이와 같은 크리에이티브 요인을 동영상 소비 환경에 적용할 경우 콘텐츠 크리에이티브 요인은 <그림 2>에서 제시하는 바와 같이 ‘영상 크리에이티브 요인’과 ‘썸네일 크리에이티브 요인’으로 구분할 수 있다. 영상 크리에이티브 요인은 동영상의 내용과 형식에 관련된 변수로 구체적으로 영상이 어떠한 내용을 얼마나 많이 담고 있는지에 관련된 내용적인 요소와 해당 영상의 길이와 같은 형식적인 요인을 모두 포괄한다. 한편 썸네일 크리에이티브 요인은 동영상의 표지 역할을 하게 되는 이미지의 내용 구성과 얼마나 자극적인 이미지를 사용했느냐와 같은 표현의 형식에 관한 변수가 있다. 이와 같은 영상 크리에이티브와 썸네일 크리에이티브와 관련된 다양한 요인은 영상 소비를 결정하는 주요한 요인이 된다. 소비자들은 썸네일과 제목을 통해 시청하고자 하는 콘텐츠가 어떤 내용을 담고 있는지 예상할 수 있으며, 여러 콘텐츠에 대한 외형정보가 한 페이지에 제시되는 인터페이스를 통해 적극적으로 탐색하고 비교해 가며 참여할 콘텐츠를 선택하기 때문이다(Hoiles, Krishnamurthy, & Aprem, 2016).

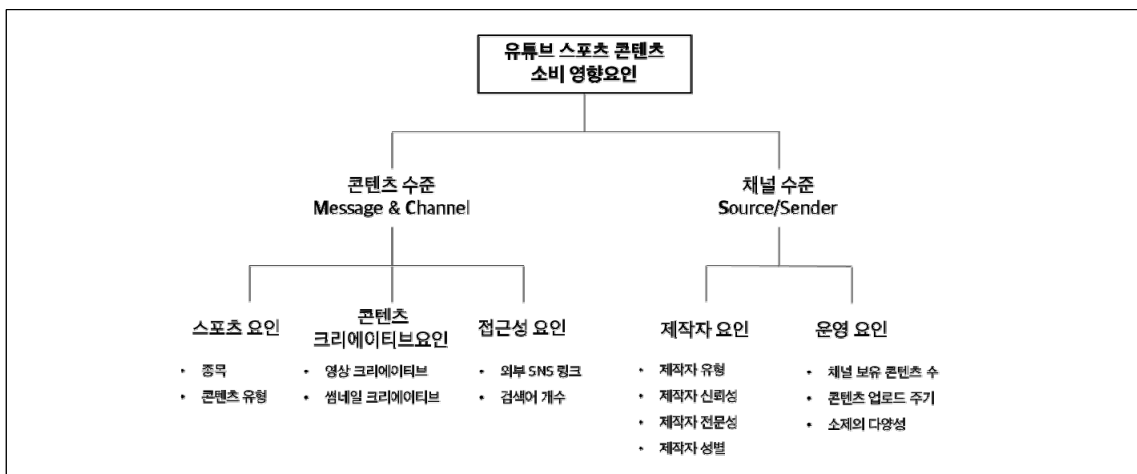


그림 2. 유튜브 스포츠 콘텐츠 소비 영향요인

세 번째로 콘텐츠에 대한 접근성을 소비 영향요인으로 설정하였다. 본 연구에서 유튜브 접근성 요인은 영상 콘텐츠의 조회수를 증가시키는 미디어 환경적 요인으로 정의된다. 환경적 요인은 콘텐츠 내용, 주제와는 관계없는 콘텐츠 외적인 요소이다. 유튜브는 개방형 콘텐츠 서비스로서 이용자들의 참여와 공유를 통해서 인기콘텐츠가 양산되는 환경적 특성을 가지고 있다. 이는 콘텐츠 간에 긴밀한 네트워크 구조가 구축되어 있으며, 네트워크는 이용자들의 사용패턴에 의해 구조화되는 시스템을 사용하고 있기 때문이다(Akrouf, Meriem, Yahia, & Eddine, 2013; Yoganarasimhan, 2012). 개방형 콘텐츠 소비 환경에서 인기 있는 콘텐츠가 되기 위해서는 많은 소비자에게 노출되어야 하며, 다양한 사람들에게 노출되어 전파 가능성이 높아져야 한다. 따라서 본 연구에서는 다양한 사람들에게 노출 가능성을 높이는 요인인 접근성 요인을 ‘콘텐츠에 설정된 검색어의 개수’, 외부 SNS 플랫폼에 링크된 횟수’를 하위 요인으로 설정하였다. 이는 선행연구에서 유튜브 접근성 요인을 콘텐츠의 노출성과 전파가능성으로 분류하여 정의하고 변수화 한 것에 기초한 결과이다(박병언, 임규건, 2015; Akrouf et al., 2013; Susarla, Oh, & Tan, 2012; Yoganarasimhan, 2012).

한편 채널 수준의 소비 영향요인은 제작자 요인과 운영 요인으로 구성되며 세부 내용은 다음과 같다. 우선 채널 제작자 요인은 앞선 콘텐츠 수준의 세 요인과는 달리 채널 수준에서 발생하는 소비 영향요인이다. 구체적으로 제작자 요인은 정보를 제작하고 전달하는 송신자의 정체성에 의해서 발생하는 소비 영향요인을 의미한다. 유튜브에서 정보 제작자이자 송신자는 크리에이터 혹은 채널 그 자체로 인식된다(이용관, 2019; Welbourne & Grant, 2016). 유튜브의 채널과 크리에이터는 콘텐츠 정보를 전달하는 일종의 정보 원천이며, 정보 원천에 대한 소비자의 인식과 태도는 시청 행동에 영향을 미친다. 정보원 공신력 모델(source credibility model)에 의하면 청자는 화자가 제공하는 메시지를 받아들일 때 화자에 대해 인지하는 공신력과 매력에 의해 정보 습득 태도와 정보에 대한 태도가 변화한다(Hovland, Janis, & Kelly, 1953; McGuire, 1985; Ohanian, 1990). 또한 콘텐츠를 제작하는 제작 회사에 대한 평판은 콘텐츠에 대한 태도를 변화시키기도 한다. 따라서 유튜브에서 화

자이자 제작자인 크리에이터에 의해서 형성되는 ‘제작자 요인’은 크리에이터의 성별, 크리에이터의 유형 그리고 크리에이터의 공신력(신뢰성/전문성)들을 통해서 측정될 수 있다. 다수의 선행연구들은 콘텐츠 제작자에 대한 소비자들의 태도는 주로 플레이어 인터페이스 상에서 등장하는 화자의 배경, 외모, 언어 그리고 행동 등에 의해서 형성된다고 밝히고 있다(Cunningham & Bright, 2012; Paek, Hove & Jeong, 2011; Welbourne & Grant, 2016).

마지막으로 채널 수준의 두 번째 소비 영향요인은 유튜브 제작자의 채널 운영 방식과 실적이다. 유튜브 제작자의 채널 운영 방식과 실적은 유튜브 소비자들의 태도에 영향을 미쳐 해당 채널에 게시된 영상 시청 행동에 영향을 미치게 된다는 점에서 유튜브 영상 콘텐츠의 소비 영향요인으로 설정되었다. 관련 선행 연구 중 Welbourne과 Grant(2016)가 수행한 연구에 따르면 채널의 운영 실적 및 성과는 채널 브랜드의 평판을 형성하고 긍정적인 평판은 콘텐츠의 설득력과 소비에 영향을 미친다고 한다. 또한 Coursaris와 Osch(2016)는 일종의 운영 활동이라 할 수 있는 콘텐츠 업로드량, 업로드 주기와 같은 지표들은 소비자가 채널을 신뢰하고 호의적인 태도를 형성하는데 중요한 단서라 주장하였다. 이를 근거로 본 연구에서는 크리에이터 혹은 채널의 유형과 운영 행동에 의해 축적된 채널의 실적을 ‘운영 요인’으로 조작적 정의하였다. 이러한 채널 운영 요인은 제작자가 콘텐츠를 업로드하고 소비자와 소통하며, 오랫동안 성실히 채널을 운영해 온 운영 관리 행동을 통해서 측정될 수 있다. 예를 들면 해당 채널에 얼마나 많은 콘텐츠가 업로드되었는지, 콘텐츠는 어떤 주기로 업로드되는지, 채널에 게시된 콘텐츠들의 소재가 얼마나 다양한지 등이 있다.

### 3) 유튜브 스포츠 시청 소비 성과지표: 조회수

유튜브 성과지표는 각각의 콘텐츠가 웹상에 게시되고 소비자들의 관심을 얻어 창출된 경제적, 사회적 가치를 의미한다. 유튜브 콘텐츠의 경제적 가치는 소비자들의 콘텐츠 열람 빈도와 시청 시간에 의해서 결정된다.

단순히 시청만 하던 과거 전통 미디어 시대에는 시청률을 대표적인 성과지표로 경제적 가치와, 사회적 파급

력을 평가하는 데 활용해 왔다. 그러나 시청률은 시청자들의 다양한 시청반응에 대한 정보를 포함한 포괄적 형태의 프로그램 가치를 평가할 수가 없다는 한계를 가지고 있다. 특히나 오늘날 플랫폼이 다양화되고, 시청 매체가 다각화된 디지털미디어 시대에는 사용기기가 다양하며, 시청 형태도 다양하기 때문에 시청률을 프로그램의 가치로 평가하기는 무리가 있다. 비교정형 PC나 스마트 기기를 통해 이뤄지는 시청 형태는 시청률에 포함되지 않으며, 앞서 언급된 다양한 시청 형태를 반영하지도 못한다(강명현, 2018).

한편 유튜브 영상과 같은 웹을 기반으로 한 콘텐츠의 성과 지표는 일반적으로 조회수와 같은 트래픽 데이터로 구성된다. 따라서 유튜브 성과는 오늘날 웹2.0에서의 미디어환경과 유튜브 플랫폼 고유의 특성이 반영된 형태의 성과지표로 측정되어야 한다. 이와 같은 배경에서 본 연구의 주요 종속 변인인 유튜브의 성과지표는 경제적 성과지표로서 조회수를 측정하였다.

구체적으로 경제적 성과지표는 콘텐츠나 채널이 유튜브 파트너십 프로그램을 통해서 얼마나 수익을 얻었는지를 나타내는 지표이다. 유튜브 파트너십 프로그램은 현재 콘텐츠의 조회수와 시청시간 그리고 광고 시청 빈도에 따라 크리에이터들에게 수익을 분배하고 있다. 즉 유튜브 크리에이터가 콘텐츠 게시를 통해 수익을 창출하기 위해서는 소비자의 선택, 지속적인 시청, 특정 광고의 시청 등과 같은 관심과 행동을 얻어야 한다. 따라서 경제적 성과지표는 조회수를 통해서 측정될 수 있다.

특히 관심의 경제이론(Attention economy theory)은 웹에서의 소비자의 관심이 금전적 수익으로 환원되는 현상을 설명하였는데, 유튜브에서는 조회수, 지속적인 시청과 같은 소비자들의 관심과 행동이 금전적 수익으로 환원되어 진다는 점에서 조회수를 기반으로 경제적 성과지표를 측정하는 것은 타당하다(Brynjolfsson & Oh, 2012; Manzerolle & Wisman, 2016).

### 3. 연구문제

본 연구는 유튜브 스포츠 영상 콘텐츠의 소비 결정요인을 영상 콘텐츠 수준과 유튜브 채널 수준의 독립변수를 중심으로 분석하기 위하여 아래와 같은 연구문제를

설정하였다. 본 연구에서 제시된 연구문제는 선행 연구에서 좀처럼 다루어지지 않았던 유튜브 플랫폼에서 유통되는 스포츠 영상 콘텐츠에 대한 시청 행동을 연구 대상으로 삼은 것으로 기성 체육학 연구에 보고된 스포츠 미디어 연구를 새로운 미디어 소비 환경에 확장하여 적용하였다는 데 의미가 크다.

RQ1: 유튜브 스포츠 콘텐츠의 조회수에 영향을 미치는 소비 영향요인은 콘텐츠 수준과 채널 수준에 모두 존재하는가?

RQ2: 유튜브 스포츠 콘텐츠의 조회수에 유의미한 영향을 미치는 콘텐츠 수준의 소비 영향요인은 무엇인가?

RQ3: 유튜브 콘텐츠 수준의 소비 영향요인을 통제한 상태에서 스포츠 콘텐츠의 조회수에 유의미한 영향을 미치는 채널 수준의 소비 영향요인은 무엇인가?

## 연구방법

### 1. 자료

#### 1) 분석 대상

본 연구의 최소 분석 단위(unit of analysis)는 유튜브 플랫폼 상에 게시된 개별 동영상 콘텐츠와 채널로 구분하여 2개의 수준으로 설정하였다. 콘텐츠 수준의 소비 영향요인과 채널 수준의 소비 영향요인을 모두 분석 대상으로 하기 때문이다. 구체적으로 스포츠와 관련된 콘텐츠와 채널에서 다양한 정보를 수집하기 위해 구독자 수, 콘텐츠 평균 조회수, 채널에 포함된 콘텐츠의 개수 등을 기준으로 상위 150개의 스포츠 관련 유튜브 채널과 각각의 채널에 배속된 영상 105,894개를 최초 분석 대상으로 선정하였다(2020년 8월 20일 검색 기준). 총 150개의 채널을 우선적으로 분석 대상으로 선정한 이유는 스포츠 콘텐츠의 높은 조회수 획득에 영향을 미치는 요인을 파악하기 위한 유의미한 자료를 확보하기 위해서이다. 구체적으로 이를 위해 수많은 채널들 중 유튜브

플랫폼에서 파트너 채널로서 운영되고 있으며, 이 중 스포츠에 집중된 콘텐츠를 업로드하는 적절한 채널을 분석 대상으로 선정하였다.

대상이 되는 채널과 콘텐츠는 대한민국을 주요 시장으로 유튜브 플랫폼에 등록된 채널로 한정하였다. 그리고 150개의 채널 중 스포츠로 등록되었지만 스포츠와 관련이 부족한 콘텐츠들을 주로 업로드한 26개의 채널들은 연구대상에서 제외되었다. 26개의 채널이 대상에서 제외된 이유는 스포츠와 관련 없는 콘텐츠가 다수 포함되어 있거나(12개 채널), 법적인 문제로 인해 채널 운영이 중단된 경우(6개), 콘텐츠 업로드가 중단된 지 5개월이 넘는 채널(8개)이 제외되었기 때문이다.

개별 콘텐츠의 경우 스포츠와 관련이 없는 경우이거나 게시 경과시간이 24주가 되지 않은 3,421개의 콘텐츠는 대상에서 제외하였다. 콘텐츠 게시 경과시간이 6달이 되지 않은 경우는 웹상에서 활발히 소비되고 있는 과정에 있기 때문이다. 이는 Cheng, Dale, Liu(2007, 2008)의 연구에서 제시한 콘텐츠의 'Active life span'에 근거한 결과이다. 해당 연구에서는 일반적으로 유튜브 콘텐츠의 조회수 성장은 대부분 150~170일 사이에 끝나는 것으로 보고하고 있다. 이후 수집된 102,473개 중 썸네일을 비공개로 해두거나(2,422개) 댓글을 비공개로 해 두어(1,154개) 주요변수의 수집이 불가능한 케이스 3,380개, 그리고 영상의 주요 소재가 스포츠가 아니어서 내용 분석을 통한 변수 작성이 불가능한 케이스 1,587개는 결측치를 보이는 자료로서 연구 대상에서 제외되었다. 즉, 본 연구에서 결측치로서 제외된 영상의 개수는 8,543개이다. 그리고 상·하위 2.5%의 조회수 이상값(outlier)을 보이는 자료 4,096개 또한 참고범위 산정법을 기준으로 분석 대상에서 제외하였다. 결과적으로 결측치 8,543개와 이상값 4,096개의 자료가 제외되어 최종적으로 89,834개의 자료가 분석 대상으로 결정되었다.

## 2) 자료의 구성

본 연구의 주요 변수와 측정에 관한 정보는 <표 1>과 같다. 주요 종속 변수인 성과지표는 조회수와 댓글수로 구성되어 있으며, 주요 독립변수는 4가지 형태의 소비 결정 요인으로 구성되어 있다. 구체적으로 4가지 형태의 소비 결정 요인은 앞서 제시된 바와 같이 스포츠 요인, 콘텐츠

요인, 채널 요인 그리고 접근성 요인으로 구성되어 있다.

독립변수 중 스포츠 요인은 콘텐츠 소비에 영향을 미치는 스포츠와 관련된 특화 요인으로서 1) 종목 유형, 2) 스포츠 콘텐츠 유형으로 구성되었다. 수집된 수많은 콘텐츠들의 종목은 총 35개(축구, 야구, 피트니스, 골프, 농구, 당구, 격투기, 배구, 태권도, 수영, 테니스, 빙상, 핸드볼, 씨름, 육상, 체조, 사이클, 유도, 펜싱, 탁구, 레슬링, 양궁, 유도, 역도, 복싱, 배드민턴, 근대 5종, 볼링, 레이싱, 하키, 치어리딩, 검도, 피겨스케이팅, 빙상, 철인 3종)이었다. 종목 자체를 분석 유목으로 설정하기엔 종목별 개체 수의 불균형과 결과해석의 문제가 있기 때문에 McLean, Hurd 그리고 Rogers(2005)가 제시한 스포츠 종목 유형 분류 기준에 기초하여 종목을 5가지 종목 유형으로 분류하였다. 분류된 종목 유형은 기록형 경쟁 스포츠(골프, 당구, 수영, 육상, 근대 5종, 볼링, 레이싱, 사이클, 빙상, 양궁, 역도, 철인 3종), 영역형 경쟁 스포츠(축구, 농구, 핸드볼, 하키), 네트형 경쟁 스포츠(배구, 테니스, 탁구, 배드민턴), 필드형 경쟁 스포츠(야구), 투쟁형 경쟁 스포츠(격투기, 태권도, 씨름, 유도, 펜싱, 레슬링, 복싱, 검도), 단련형 스포츠(피트니스, 요가, 필라테스), 표현 및 율동 스포츠(체조, 치어리딩, 댄스스포츠, 피겨스케이팅), 아웃도어 스포츠(클라이밍, 요트, 서핑)으로 구분된다. 한편 콘텐츠 유형의 경우 앞서 이론적 배경에서 언급된 바와 같이 Raney(2006)가 제시한 스포츠 미디어 시청 동기를 기준으로 1) 레슨 및 정보형 영상, 2) 오락 및 엔터테인먼트 영상, 3) 경기 및 하이라이트 영상, 4) 프로스포츠 정보 및 분석 영상, 5) 커뮤니티형 영상, 6) 기타 유형의 5가지로 구분하여 구성하였다.

주요 독립변수 중 두 번째인 콘텐츠 크리에이티브 요인은 영상 단위에서 측정되는 콘텐츠 크리에이티브 요인들이다. 앞서 제시된 이론적 배경에 근거하여 영상 크리에이티브 요인은 영상의 길이를 초로 환산하여 측정하였고, 영상에서 제공되는 음성 정보의 양은 분당 재생되는 단어의 수로 측정하였다. 썸네일 크리에이티브 요인은 제목의 길이, 썸네일에 포함된 텍스트의 수, 그리고 썸네일의 선정성과 폭력성으로 측정하였다. 이는 제작자가 콘텐츠를 제작할 때 고려되는 크리에이티브 요소 중 정보의 양과 형태를 측정하기 위해서이다.

주요 독립변수 중 세 번째는 접근성 요인이다. 콘텐츠



가 많은 선택을 받기 위해서는 보다 많은 소비자에게 노출되어야 한다. 접근성 요인은 콘텐츠가 많은 소비자에게 노출되도록 해주는 요인들을 의미하며, 1) 검색어를 다양하게 설정하거나 2) 타 SNS에 인용(link)되는 횟수를 측정하여 구성하였다. 이는 검색어를 다양하게 설정하여 웹 인터페이스에서 높은 빈도로 소비자에게 노출되는 경우와 외부 SNS에 링크되어 다른 채널에서의 유입 효과가 있음을 측정하기 위함에 목적이 있다.

네 번째는 제작자 요인이다. 제작자 요인은 콘텐츠 수준이 아니라, 콘텐츠를 제작하고 게시하는 송신자의 정

체성을 측정하는 변수들로 구성되었다. 구체적으로 제작자의 전문성, 신뢰성 그리고 성별을 측정하였다. 특히 전문성과 신뢰성의 경우 Ohanian(1990)이 제안한 정보원 공신력(Source Credibility), 소비자에게 형성된 정보 제공자의 신뢰성과 전문성을 측정하는 문항을 활용하였다.

마지막 다섯 번째는 운영 요인이다. 운영 요인은 채널의 운영 기간, 콘텐츠 평균 업로드 시간으로 측정되어졌다. 각각의 채널에 얼마나 많은 콘텐츠가 업로드되었는지는 채널의 규모를 측정하는 것이며, 콘텐츠 평균 업로드 주기는 게시자가 얼마나 성실히 콘텐츠를 업로드하

표 1. 연구의 변수와 측정

유형	요인	변수	정의 및 측정	수집 방식	
종속 변수	경제적 성과지표	조회수	콘텐츠 활성 수명기간 동안 누적된 조회수	웹 크롤링	
콘텐츠 수준 독립 변수	스포츠 요인	종목 유형	기록 경쟁=1, 영역형 경쟁=2, 네트형 경쟁=3, 필드형 경쟁=4, 투쟁형 경쟁=5, 단련형 경쟁=6, 표현 및 율동=7, 아웃도어=8, 기타=9	내용 분석	
		스포츠 콘텐츠 유형	레슨 및 정보형 영상=1, 오락 및 엔터테인먼트형 영상=2, 하이라이트 및 경기 영상=3, 프로스포츠 관련 정보 및 분석 영상=4, 선수, 팀과 관련된 팬 커뮤니티형 영상=5, 기타 영상=6		
	콘텐츠 요인	영상 크리에이터	영상 길이	영상의 길이를 초 단위로 환산한 값	웹 크롤링
		썸네일 크리에이터	음성 정보량	분당 단어 재생 수(word for minute_VidiQ제공)	2차 데이터
			성적 선정성	Google 마이닝 프로그램 이용 5점 척도	API 이미지 분석
			폭력성	Google 마이닝 프로그램 이용 5점 척도	
			썸네일 글자수	Google 마이닝 프로그램 이용음절 단위로 측정	웹 크롤링
		제목 글자수	크롤링 데이터를 음절 단위로 측정		
	접근성 요인	외부 SNS 링크 수	페이스북, 트위터에 링크된 횟수	2차 데이터	
		검색 개수	채널 게시자가 설정한 검색 색인어의 개수	웹 크롤링	
채널 수준 독립 변수	크리에이터 요인	전문성	크리에이터 공신력 중 전문성 관련 6문항의 평균 값(7점 척도) 'Authoritativeness', 'Competence', 'Qualification', 'Trained', 'Informed', 'Educated'	내용 분석 Ohanian (1990)	
		신뢰성	크리에이터 공신력 중 신뢰성 관련 6문항의 평균 값(7점 척도) 'Favorable disposition', 'Acceptance', 'Psychological safety', 'Perceived supportive climate'		
		성별	성별 비공개 및 알 수 없음=0, 남성=1, 여성=2, 혼성=3		-
	운영 요인	채널 개설 후 경과시간	최초 게시일로부터 경과시간(일 단위)	웹 크롤링	
		평균 콘텐츠 업로드 주기	일단위 콘텐츠 업로드 주기의 평균		
		제작자 유형	개인 제작자=1, 프로 구단=2, 방송사=3, 유튜브 전문 제작사=4	내용 분석	
통제 변수	콘텐츠 게시 경과시간	콘텐츠 게시 후 경과시간을 일 단위로 측정	웹 크롤링		
	채널 구독자수	채널 수준의 통제변수로서 채널의 구독자수 평균			
	업로드 영상 개수	해당 채널에 게시된 콘텐츠 개수			

능가를 측정한 것이다.

한편, 본 연구에서는 시간에 따른 성과지표의 변동과 채널이 보유한 구독자 수와 영상 개수와 같은 채널 공신력에 의한 성과지표 획득의 차이를 통제하기 위해서 콘텐츠 게시 경과시간, 채널 구독자 수, 채널에 포함된 영상의 개수를 통제변수로 설정하였다.

### 3) 자료수집 절차

자료수집을 위해 <표 2>와 같이 총 4단계의 절차를 걸쳤다. 첫 번째 단계에서는 최소 분석 단위가 되는 영상과 이와 관련된 모든 정보를 수집하기 위해 프로그래밍 언어인 파이썬(Python)을 이용하여 웹 크롤러(web crawler)를 개발하여 데이터베이스화하는 데에 사용하였다. 구체적으로 기초 데이터 선정과 수집을 위해 채널을 선정하고, 각각의 채널에 배속된 영상 URL을 웹 크롤러를 이용하여 데이터베이스에 저장하였다. 저장된 URL을 이용하여 썸네일 이미지 파일, 조회수, 댓글수, 영상 제목, 영상 길이, 음성 정보량, 검색어, 콘텐츠 게시일 등을 추가로 데이터베이스에 저장하였다. 이러한 기초 데이터 수집은 2020년 8월 20일에서 8월 27일까지 1주일 동안 이뤄졌다.

두 번째 단계에서는 데이터베이스에 저장된 데이터를 마이닝하여 비정형화된 형태의 데이터들을 정형화하였다. 구체적으로 첫 번째 단계에서 수집된 썸네일 이미지 파일을 마이닝 프로그램에 투입하여 이미지의 성적 선정성, 폭력성, 이미지에 포함된 글자 수와 같은 정보를 분석하였다. 구체적으로 구글 개발자 플랫폼에서 제공하는 Google Vision AI를 이용하여 이미지 분석을 실행하였

다. Google Vision API는 수억 개의 사전 이미지를 활용하여 이미지가 가진 유해성을 판별하고, 이미지에 포함된 객체를 인지하는 기능이 있다. 머신러닝 알고리즘에 의해 사용자 데이터를 학습데이터로 인식하여 정확도가 높아지는 시스템을 차용하고 있는데, 2020년 8월 사용 당시 베타서비스를 종료하고 정식서비스로서 사용자에게 제공되고 있다.

세 번째 단계에서는 앞선 두 단계에서 수집되기 어려운 자료들을 유료 분석 사이트를 통해 획득한 이후 데이터베이스에 저장하였다. 구체적으로 실시간으로 콘텐츠에 대한 통계치를 작성하고, 일 단위로 채널에 대한 다양한 통계 정보를 제공하는 유튜브 분석 전문 업체인 VidiQ.com과 Noxinfluence.com을 통해 웹 크롤링과 마이닝으로 수집하지 못한 자료들을 추가로 수집하였다. 결과적으로 세 번째 단계에서는 콘텐츠의 추정된 수익, 그리고 각각의 콘텐츠가 제공하는 음성량, 외부 SNS 링크 개수, 크리에이터가 설정한(사용자에게 노출되지 않는) 검색어 정보를 수집하여 데이터베이스에 추가하였다.

마지막 네 번째 단계에서는 기계적인 분석을 통해 측정되기 어려운 변수들을 내용 분석을 통해 데이터베이스에 저장하였다. 스포츠 종목, 종목의 유형, 그리고 콘텐츠 유형과 같은 정보는 콘텐츠의 내용을 확인하지 않고는 변수화하여 저장하기 어렵기 때문에 스포츠 마케팅 및 미디어를 주전공으로 하는 석사 및 박사 과정생 5명을 코더로 선정하여 내용 분석을 실시하여 각각의 변수를 측정하였다. 구체적으로 종목의 경우 크리에이터가 설정한 콘텐츠 검색어를 활용하여 1차적으로 종목을 코딩하였고, 이를 통해 코딩되지 않은 콘텐츠들의 경우 제

표 2. 자료수집 방법 및 절차

	방법	내용	사용 프로그램
1	데이터 크롤링 (Data Crawling)	파이썬(Python) 언어로 작성된 크롤링 프로그램을 이용한 유튜브 동영상 관련 비정형 데이터 수집	Python3.8.2, Bs4, Selenium, Pandas
2	데이터 마이닝 (Data Mining)	Google Cloud API에서 제공되는 데이터 마이닝 패키지 및 오픈 소스 API를 활용한 비정형 데이터의 정형화	Vision AI API, Excel Macro
3	유료 2차 데이터 수집	유튜브 데이터 전문 분석 업체의 자체 알고리즘을 활용한 관련 데이터 수집	Nox Influencer, VidiQ
4	내용 분석 (Contents Analysis)	코더 3인, 내용 분석을 통해 채널 관련 비정형 데이터를 변수로 분류	-

목을 활용하여 2차적으로 코딩하였다. 이 두 가지 경우를 통해서도 코딩이 되지 않는 경우에는 영상을 재생하여 종목을 확인하였다. 그리고 콘텐츠 유형의 경우 또한 위의 과정과 마찬가지로 검색어와 제목을 통해 코딩하였으며, 이 두 과정을 통해 코딩되지 않은 경우 영상을 재생하여 확인하는 과정을 거쳤다. 한편 크리에이터 요인 중 크리에이터의 전문성과 신뢰도의 경우 5명의 코더가 각각의 채널의 프로필과 대표콘텐츠를 시청한 후 Ohanian(1990)이 제시한 전문성과 신뢰성 측정 문항에 응답하는 형식으로 측정되었다.

## 2. 자료의 분석

본 연구에서는 유튜브 스포츠 콘텐츠 소비 영향요인들(콘텐츠 수준/채널 수준)이 조회수에 미치는 영향력을 검증하고자 다층모형 회귀 분석(Multi-level Modeling regression)을 실시하였다. 다층모형 회귀분석은 서로 상이한 분석수준의 변수들이 종속변수에 미치는 영향력을 통합적으로 분석하는 통계기법이다. 다층모형 회귀분석에서 분석수준은 개별 수준에서 그룹 수준으로 위계가 나뉘어져 있다. 이때 개별수준의 독립변수들이 종속변수에 미치는 영향력이 개별 수준을 넘어 그룹 수준의 맥락적 요인에 의해 증폭되거나 완화되는 정도를 추정하기 위해 개발된 분석기법이 바로 다층모형 회귀분석이다(Raudenbush & Bryk, 2002).

특히 본 연구에서는 콘텐츠 수준의 변수들이 채널수준의 변수들에 포함된 자료 형태를 띄고 있어, 대표적인 내포 모형(Embedded Model)인 확률 절편 모형(Random intercept model)을 사용하였다. 내포 모형은 1수준의 개별 자료가 2수준에 구분되어 내포될 때 사용할 수 있는 모델링 기법인데, 본 연구의 대상인 콘텐츠들은 각각의 채널에 구분되어 명확히 포함되기 때문이다. 이러한 분석 방법은 2수준의 회귀계수를 효율적으로 추정할 수 있게 해주며 단순준 회귀분석에서 빈번히 발생하는 집단 편이의 문제(aggregation bias)와 표준오차의 추정오류(Misestimated standard errors)를 효과적으로 통제할 수 있다는 장점이 있다(Raudenbush & Bryk, 2002).

구체적으로 스포츠 요인(종목, 콘텐츠 유형), 콘텐츠

크리에이티브 요인(영상/썸네일), 접근성 요인(검색어 수, 타 SNS 링크 수)은 콘텐츠 수준에서 영향을 미치는 소비 영향요인으로 분석에 투입되었고, 제작자 요인과 운영 요인은 채널 수준에서 영향을 미치는 소비 영향요인으로 분석에 투입되었다. 그리고 확률절편 모형 분석법에 따라 총 3가지 모형을 통해 분석을 진행하였다. 첫 번째는 <수식 1>과 같은 무조건 모형(Unconditional null model), 두 번째는 <수식 2>와 같은 개별 콘텐츠 수준 모형(Individual Contents level model), 세 번째는 <수식 3>과 같은 1수준 변수와 2수준 변수의 결합모형(Combined model)이다.

### 수식 1. 무조건 모형

$$Y_{ij} = \beta_{0j} + \epsilon_{ij}, \epsilon_i \sim N(0, \delta^2)$$

$$\beta_{0j} = \gamma_{00} + u_{0j}, u_{0j} \sim N(0, \tau_{00})$$

$Y_{ij}$  : j번째 채널의 i번째 영상의 종속변수 Y값  
 $\beta_{0j}$  : j번째 채널에 속한 콘텐츠들의 종속변수 Y 평균값  
 $\gamma_{00}$  : 영상들의 종속변수 Y에 대한 전체 평균  
 $\epsilon_{ij}$  : j번째 채널의 i번째 영상의 영상수준(1수준) 고유효과  
 $\epsilon_{ij}$  : j번째 채널의 채널 수준(2수준) 고유효과  
 $u_{0j}$  : 채널 내의 영상들의 개별적인 차에 의해 발생하는 종속변수 Y의 분산  
 $\delta^2$  : 채널별 차이에 의해 발생하는 종속변수 Y의 분산  
 $\tau_{00}$  : 채널별 차이에 의해 발생하는 종속변수 Y의 분산

### 수식 2. 1수준 모형

$$Y_{ij} = \beta_{0j} + \sum \beta_p X_{\pi j} + \epsilon_{ij}, \epsilon_i \sim N(0, \delta^2)$$

$$\beta_{0j} = \gamma_{00} + u_{0j}, u_{0j} \sim N(0, \tau_{00})$$

$$\beta_p = \gamma_{p0}$$

$\beta_p, \gamma_{p0}$  : p번째 영상단위(Level 1) 독립변수  $X_p$ 의 회귀계수  
 $X_{\pi j}$  : j번째 채널의 i번째 영상의 p번째 영상수준(Level 1) 독립변인  $X_p$ 값

### 수식 3. 2수준 모형

$$Y_{ij} = \beta_{0j} + \sum \beta_{pj} X_{\pi j} + \epsilon_{ij}, \epsilon_i \sim N(0, \delta^2)$$

$$\beta_{0j} = \gamma_{00} + \sum \gamma_{0q} Z_{qj} + u_{0j}, u_{0j} \sim N(0, \tau_{00})$$

$$\beta_p = \gamma_{p0}$$

$\gamma_{0q}$  : q번째 채널수준(Level 2) 독립변수  $Z_q$ 의 회귀계수

$Z_{qj}$  : j번째 채널의 q번째 채널수준(Level 2) 독립변인  $Z_q$  값

## 연구결과

### 1. 자료처리 및 주요변수의 기술통계

수집된 자료의 기술통계는 <표 3>과 같다. 명목형 변수로서 수집된 자료인 스포츠 종목, 콘텐츠 유형과 채널 제작자 성별, 유형, 채널 운영 종목 수는 빈도와 백분율로 제시하였으며, 연속형 변수로 수집된 영상 길이, 음성 정보량, 제목 길이, 썸네일 텍스트 수, 썸네일 선정성, 폭력성, 검색어 개수, 타 SNS 링크 수, 제작자 공신력, 채널의 평균 영상 업로드 주기, 개별 콘텐츠들의 게시 후 경과시간, 채널의 구독자 수와 운영 기간은 평균과

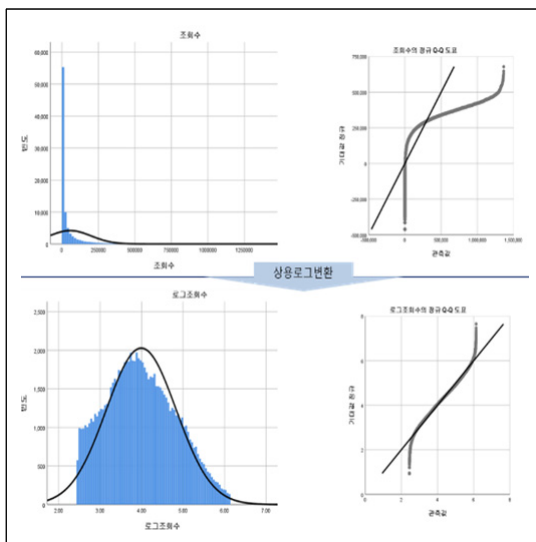


그림 3. 조회수 상용로그 변환 분포 및 Q-Q 도표

표준편차로 기술통계 값을 계산하였다.

이중 특히 조회수의 경우 총 표본 콘텐츠 수 89,834개의 평균값이 56,840.27이었으며 표준편차는 142,914.44였다. 그러나 왜도가 4.83, 첨도가 28.11으로 표본이 왼쪽으로 치우친 형태를 하고 있어 정규분포 곡선을 따르지 못하고 있었다. 자료의 정규성을 확보한 형태로 분석 모형에 투입하기 위해 상용로그로 치환하였다. 상용로그로 치환된 조회수의 평균값은 2.44, 표준편차는 0.84이었으며 왜도는 0.23, 첨도는 -0.63으로 도출되었다. 이와 같은 정보를 통해 상용로그로 치환한 조회수 값이 정규분포곡선에 가까워진 것을 확인함과 동시에 잔차항의 동분산성을 확보한 것을 확인하였다. <그림 3>은 조회수를 상용로그로 변환하였을 때의 분포와 Q-Q 도표이다. 분포가 정규곡선에 가까우며 Q-Q 도표가 직선에 가까운 값을 보이기 때문에 상용로그로 변환한 값이 정규성을 가지는 것을 확인할 수 있다.

### 2. 조회수에 영향을 미치는 소비 영향요인 분석

#### 1) 기초모형 분석: 연구문제 1

본 연구에서는 조회수에 대한 소비 영향요인들의 영향력을 검증하는 데 있어 콘텐츠 수준과 채널 수준의 영향요인이 유의미함을 우선적으로 확인하기 위해 1) 채널 수준의 오차항의 변량이 유의미한 수준에 있는지 확인하고, 2) 영상 수준의 오차항의 변량과 채널 수준의 오차항의 변량의 차이로 인해 발생한 분산값의 비율(ICC: Intra Class Correlation)을 분석하였다. 단, 종속변수인 조회수는 자료의 정규성과 계수 해석의 용이성을 위해서 로그로 변환한 값을 투입하였다. 로그로 변환한 조회수를 이용해 작성된 기초모형의 분석 결과는 <표 4>와 같다.

기초모형 분석 결과 채널 수준 오차항의 변량은 유의미한 수준에 있음을 확인하였다. 확률효과의 채널 수준 변량을 이용한 최대 우도비 검증(Maximum Likelihood Ratio: 이하 LR 테스트) 값이 45157.25( $p < .001$ )로 도출되었기 때문이다. 확률효과에서 채널 수준의 오차항의 변량이 유의미한 수준에 있다는 것은 귀무가설을 기각한다는 것을 의미한다. 기초모형에서의 귀무가설은 확률절편(random intercept)의 모형이 합동(pooled) OLS의 추정치보다 적절하지 않다는 것인데, 이 귀무가설은 기

표 3. 주요변수와 기술통계

변수명		평균	표준편차	빈도	%
조회수(왜도: 4.83, 첨도: 28.11)		56,840.27	142,914.44	-	-
로그 조회수(왜도: .23, 첨도: -.63)		2.44	.84	-	-
스포츠 요인	종목 유형	영역형 경쟁	-	29106	32.40%
		필드형 경쟁	-	20072	22.34%
		기록형 경쟁	-	16781	18.68%
		단련형	-	13809	15.37%
		투쟁형 경쟁	-	6642	7.39%
		네트형 경쟁	-	2886	3.21%
		표현 및 율동	-	538	0.54%
		아웃도어	-	36	0.03%
		기타	-	42	0.03%
	콘텐츠 유형	프로스포츠 정보 및 경기분석	-	35,361	39.37%
		레슨 및 교육	-	22,758	25.34%
		커뮤니티형 영상	-	13,425	14.91%
		하이라이트 및 경기영상	-	9,025	10.06%
		오락 및 도전 엔터테인먼트	-	7,793	8.67%
		기타(홍보, 이벤트 등)	-	1,550	1.63%
콘텐츠 크리에이티브 요인	영상 크리에이티브	영상길이(초)	803.96	1694.54	-
		음성 정보 량	33.79	57.70	-
	썸네일 크리에이티브	제목 길이(음절)	40.34	16.08	-
		텍스트 수(음절)	47.19	55.84	-
		선정성(5점 척도)	1.55	0.86	-
		폭력성 코딩(5점 척도)	1.59	0.71	-
접근성 요인		검색어 개수	17.62	16.59	-
		타 SNS 링크	3.51	6.06	-
채널 제작자 요인	성별	남성	-	79	63.7%
		여성	-	14	11.3%
		혼성	-	6	4.8%
		비공개	-	25	20.2%
	공신력	전문성	5.02	1.74	-
		신뢰성	5.15	1.47	-
	제작자 유형	개인	-	90	72.6%
		방송사	-	13	10.5%
		구단 및 협회	-	10	8.1%
		유튜브 전문 제작사	-	11	8.9%
채널 운영 요인	다종목 여부	단일종목	-	114	91.9%
		다종목	-	10	8.1%
	평균 업로드 주기(일 단위)		4.08	3.20	-
	영상 개수		912.24	1,789.31	-
통제 변수	콘텐츠	콘텐츠 게시 경과 시간(일)	848.63	603.71	-
		채널 구독자 수	242,746.77	361,956.63	-
	채널	채널 운영기간(일)	1,260.26	823.52	-

표 4. 조회수에 대한 기초 모형 분석 결과

고정효과	회귀계수(B)	표준오차(SE)	z
전체 평균( $\tau_{00}$ )	4.54***	.048	95.10
확률효과	분산성분(Var)	자유도	LR Test $\chi^2(1)$
채널 수준( $\mu_{0j}$ )	.529	123	45157.25***
영상 수준( $e_{ij}$ )	.425	-	-
model test			
ICC	.474		
BIC	178,818.3		
AIC	178,789.8		

주1. \*은  $P < .01$ , \*\*은  $p < .05$ 수준, \*\*\*은  $p < .001$  수준에서 유의미함.

각되어 다층모형의 추정치가 적절하다는 것을 의미한다. 이는 다시 말하자면 조회수의 크기가 채널에 따라 상이하기 때문에 채널 수준의 설명변수들의 분석이 적절하다는 것을 의미한다.

또한 영상 수준 오차항의 변량과 채널 수준 오차항의 변량의 차이로 인해 발생한 분산값의 비율 ICC값이 0.474로 도출되었는데, 이는 채널 수준에서 조회수를 설명하는 변량이 47.4%임을 의미한다. 결과적으로 채널 수준 설명변수들의 변량이 충분하다는 것은 다층모형을 이용해 조회수에 영향을 미치는 설명변수들의 분석이 적절함을 의미한다.

## 2) 1수준(콘텐츠) 소비 영향요인 다층모형 분석:

### 연구문제 2

영상 수준의 설명변수들의 영향력을 검증하기 위해 영상 수준에서 측정된 소비 영향요인들만을 고정효과(fixed effect) 변수로서 확률절편 모형(Random intercept Model)에 투입하였다. 분석한 결과는 <표 5>의 모형 1-1과 같다.

우선 영상 수준의 확률효과에 해당하는 변량  $e_{ij}$ 가 기초모형은 0.425인 반면 모형 1-1에서는 0.389로 줄어든 것을 확인할 수 있다. 이는 영상 수준의 설명변수들이 모형에 투입됨에 따라 설명의 확률절편의 분산추정치가 줄어들었기 때문이다. 이는 모형 1-1은 기초모형에 대비하여 설명변수에 의해 모형의 설명력이 높아졌음을 의미한다. 그리고 확률효과와 채널 수준 변량( $\mu_{0j}$ )을 이

용한 최대 우도비 검증(Maximum Likelihood Ratio: 이하 LR 테스트) 값은 29284.74( $p < .001$ )으로 도출되었는데, 이는 영상 수준 설명 변수에 따른 조회수 증감이 기초모형과 같이 모형 1-1에서 또한 채널별로 차이가 있다는 것을 의미한다. 즉 다층모형의 분석이 모형 1-1에서 또한 유의미하게 적용되었다는 것이다. 추가로 기초모형과 비교하여 설명변수가 추가된 모형 1-1의 설명력이 높아졌음은 AIC값과 BIC값의 비교를 통해서도 확인할 수 있다. 기초모형의 AIC(Akaike Information Criterion) 값은 178789.8, BIC 값은 178818.3이었는데 모형 1-1의 AIC값은 170961.5 BIC(Bayesian Information Criterion)값은 171196.6으로 더 낮게 도출되었고, 이를 통해 설명 변수가 추가된 모형 1-1이 기초모형에 대비하여 설명력이 높은 모형임을 확인할 수 있다.

해당 모형의 적절성이 검증되어 계수값 해석을 통해 콘텐츠 수준의 설명변수들의 영향력을 분석할 수 있다. 모형 1-1의 종속변수인 조회수는 로그로 변환되어 투입되었기 때문에 계수값을 백분율로 해석하였다. 분석 결과를 종합하면 다른 변수들이 통제된 상황에서 스포츠 요인인 종목 유형과, 스포츠 콘텐츠 유형에 따라 조회수가 상대적으로 유의미한 차이가 있는 것으로 분석되었다. 종목 유형의 경우 기준변수로 기록형 경쟁 스포츠를 투입하여 종목 유형 간의 차이를 분석하였다.

구체적으로 우선 종목 유형의 계수값을 해석해 보면 기록형 경쟁 스포츠 콘텐츠들의 조회수 보다 표현형 스포



표 5. 다층모형 분석결과(DV: 조회수)

$n(\text{영상}) = 89,894$ $n(\text{채널}) = 124$			모형 1-1		모형 1-2	
			<i>B</i>	<i>S.E</i>	<i>B</i>	<i>S.E</i>
1수준 고정 효과	종목 유형	영역형	-.2387***	.01	-.2359***	.02
		네트형	-.3285***	.03	-.3249***	.03
		필드형	.0082***	.02	.0096***	.02
		투쟁형	-.1596***	.03	-.1554***	.03
		단련형	-.2485***	.04	-.3177***	.04
	스포츠 요인	표현형	.1345***	.04	.1366***	.04
		오락 엔터	-.1222***	.01	-.1216***	.01
		하이라이트/경기	-.3021***	.02	-.2995***	.02
		프로스포츠분석	-.2878***	.02	-.2854***	.01
		커뮤니티	-.0725***	.02	-.0699***	.01
	콘텐츠 유형	기타	-.5147***	.02	-.5144***	.02
		제목길이	-.00003**	.00	.00004**	.00
		썸네일 선정성	.0213***	.00	.0125***	.00
		썸네일 폭력성	.0754***	.00	.0738***	.00
		썸네일 텍스트 수	.0003***	.00	.0003***	.00
	콘텐츠 크리에이 티브 요인	영상 길이	-.0000***	.00	-.0000***	.00
		영상 음성량	.0006***	.00	.0005***	.00
	접근성 요인	검색어 개수	.0041***	.00	.0041***	.00
		타 SNS 링크 개수	.0462***	.00	.0445***	.00
	통제	게시 경과시간	.0011***	.00	.0010***	.00
2수준 고정 효과	성별	비공개	-	-	-.1547***	.27
		여성	-	-	-.0569***	.23
		혼성	-	-	-.4605***	.31
	제작자 요인	전문성	-	-	-.0512***	.04
		신뢰성	-	-	.0141***	.03
	크리에이터 유형	방송사	-	-	-.5523***	.33
		협회 및 구단	-	-	-.7625***	.32
		유튜브 전문	-	-	.1380***	.16
	운영 요인	멀티스포츠	-	-	.2435***	.19
		영상 개수	-	-	.0007***	.00
		평균 업로드 주기	-	-	-.0008***	.01
	통제	채널 운영기간	-	-	.0005***	.00
		로그 채널 구독자수	-	-	.04013**	.00
확률 효과	절편		4.78***	.07	5.34***	.33
	채널 수준 변량( $\mu_{0j}$ )		.265***	.022	.122***	.014
	영상 수준 변량( $e_{ij}$ )		.389***	.020	.389***	.00
model test	ICC		.406		.238	
	AIC		170961.5		170894.1	
	BIC		171196.6		171151.5	
	LR Test $\chi^2(1)$		29284.74***		10960.41***	

츠 콘텐츠의 조회수가 약 13%( $b=.1315$ ,  $p<.001$ ) 높았다. 반면 기록형 스포츠 콘텐츠보다 필드형 경쟁 스포츠는 23.87%( $b=-.2387$ ,  $p<.001$ ), 네트형 경쟁 스포츠 32.85%( $b=-.3285$ ,  $p<.001$ ), 투쟁형 경쟁 스포츠 15.96%( $b=-.1596$ ,  $p<.001$ ), 단련형 스포츠 24.85%( $b=-.2485$ ,  $p<.001$ ) 만큼 낮은 조회수를 보였다. 한편 필드형 경쟁 스포츠의 경우 기록형 경쟁 스포츠와 유의미한 수준에서 조회수 차이를 발견하지 못하였다( $b=.0082$ ,  $p>.05$ ). 이처럼 종목 유형에 따라 서로 상이한 조회수를 보이는 것을 통해 종목 유형은 조회수에 유의미한 영향을 미치는 것을 확인할 수 있으며 기록형 경쟁 스포츠와 필드형 경쟁 스포츠 그리고 표현형 경쟁 스포츠가 상대적으로 높은 조회수를 획득하는데 다른 종목 유형들보다 유리한 종목 유형임을 확인할 수 있었다.

스포츠 콘텐츠 유형의 경우 기준변수로 레슨 및 정보형 영상을 투입하였고 이에 따라 계수값을 해석하였다. 결과에 따르면 레슨 및 정보형 영상이 모든 유형의 콘텐츠들보다 높은 조회수를 획득한 것으로 분석되었다. 구체적으로 레슨 및 정보형 콘텐츠는 오락 및 엔터테인먼트 유형 12.22%( $b=-.1222$ ,  $p<.001$ ), 하이라이트 및 경기 영상 30.21%( $b=-.3021$ ,  $p<.001$ ), 프로스포츠 분석 및 정보 영상 28.78%( $b=-.2878$ ,  $p<.001$ ), 커뮤니티형 영상 7.25%( $b=-.725$ ,  $p<.001$ ), 기타 유형 영상 51.47%( $b=-.5147$ ,  $p<.001$ )만큼 높은 조회수를 보이는 것으로 분석되었다. 결과적으로 콘텐츠 유형에 따라 서로 상이한 조회수를 보이는 것을 통해 콘텐츠 유형은 조회수에 유의미한 영향을 미치는 것으로 해석할 수 있으며, 구체적으로 레슨 및 정보형 영상이 조회수 획득에 타 유형보다 상대적으로 유리한 것으로 분석되었다.

콘텐츠 크리에이티브 요인의 경우 썸네일 크리에이티브 요인과 영상 크리에이티브 요인으로 나뉘어 분석되었다. 우선 썸네일 크리에이티브 요인 중 제목의 길이의 경우 1글자 길어질 경우 조회수가 0.003% 유의미하게 감소하는 것으로 나타났다( $b=-.00003$ ,  $p<.05$ ). 그리고 썸네일 선정성이 1단위 증가할 경우 조회수가 2.13% 유의미하게 상승하는 것으로 나타났으며( $b=.0213$ ,  $p<.001$ ), 썸네일 폭력성은 1단위 증가할 때 7.54%만큼 유의미한 증가가 나타났다( $b=.0754$ ,  $p<.001$ ). 썸네일에 포함된 텍스트의 개수는 1글자 증가할 때 0.03%의 유

미한 수준의 조회수 감소가 나타났다. 결과적으로 썸네일 크리에이티브 요인 중 제목의 길이와 썸네일에 포함된 텍스트 수는 조회수에 부(-)적 영향을 미치지만 나머지 썸네일의 선정성과 폭력성은 정(+)적인 영향을 미치는 것으로 분석되었다.

영상 크리에이티브 요인의 경우 영상의 길이와 영상에서 제공되는 음성정보의 양을 분석하였다. 구체적으로 영상의 길이의 경우 조회수에 유의미한 영향을 미치지 않는 것으로 분석되었다. 반면 영상에서 제공하는 음성정보의 양은 1단어가 증가할 때마다 조회수에 0.006%의 증가가 나타나는 것으로 분석되었다( $b=.00006$ ,  $p<.001$ ). 즉 스포츠 콘텐츠의 경우 영상 크리에이티브 요인 중 영상의 길이는 조회수에 영향을 미치는 요인이 아니며, 영상이 제공하는 정보의 양은 조회수에 정(+)적인 영향을 미치는 것으로 분석되었다.

접근성 요인의 경우 콘텐츠에 설정된 검색어(해쉬태그, Hash tags)의 개수와 콘텐츠가 페이스북과 트위터에 링크된 횟수를 분석하였다. 구체적으로 검색어 개수의 경우 검색어가 1개 늘어날 때마다 0.41%의 조회수 증가를 보였으며( $b=.0041$ ,  $p<.001$ ), 타 SNS 플랫폼 링크 개수가 1개 늘어날 때마다 4.62%의 유의미한 조회수 증가를 보였으며( $b=.0462$ ,  $p<.001$ ). 이를 통해 접근성이 증진될수록 조회수가 증가되는 것을 확인할 수 있었다.

마지막으로 콘텐츠 게시 경과시간을 변수로 투입하여 시간에 따른 조회수 획득의 차이를 통제하였는데, 분석 결과 시간의 경과에 따른 조회수 증가가 정(+)적으로 유의미한 관계가 있음을 발견하였다( $b=.0011$ ,  $p<.001$ ). 따라서 이를 통해 시간에 따른 조회수 획득 효과가 적절히 통제되었음을 확인할 수 있었다.

### 3) 2수준(콘텐츠, 채널) 소비 영향요인 분석:

#### 연구문제 3

콘텐츠 수준의 소비 영향요인들을 통제한 상태에서 채널 수준의 소비 영향요인들이 조회수에 미치는 영향력을 검증하기 위해 영상 수준에서 측정된 소비 영향요인과 함께 채널 수준에서 측정된 소비 영향요인들을 고정효과(fixed effect) 변수로서 확률절편 모형(Random intercept Model)에 투입하였다. 분석결과는 <표 5>의 모형 1-2와 같다.



모형 분석결과를 살펴보면 모형 1-1의 채널 수준의 확률효과 변량은 0.265이었지만, 모형 1-2의 채널 수준 확률효과 변량은 0.122이었다. 이를 통해 채널 수준의 확률효과가 감소했음을 확인할 수 있으며 동시에 채널 수준의 설명변수들이 투입됨에 따라 확률효과가 감소하여 고정변수들의 설명력이 높아졌음을 확인할 수 있다. 또한 모형의 설명력이 이전 모형에 대비하여 높아졌음은 AIC값과 BIC값을 통해서도 확인할 수 있는데, 모형 1-2의 AIC(170,894.1)값과 BIC(171,151.5)값이 모형 1-1의 AIC(170,961.5), BIC(171,196.6)값보다 낮아진 것을 통해 모형 1-2가 모형 1-1보다 설명력이 높은 모형인 것을 확인할 수 있다. 또한 채널 수준의 확률효과 변량( $\mu_{0j}$ )을 이용한 최대 우도비 검정 값이 10,960.41( $P < .001$ )의 유의미한 수준의 수치를 보이는 것으로 도출되었는데, 이를 통해 채널 수준 설명 변수에 따른 조회수 증감이 채널별로 차이가 있다는 것을 확인할 수 있다. 급내 상관계수 ICC값을 통해서도 채널 수준의 설명 변량과 영상 수준의 설명 변량의 비율을 확인할 수 있는데, 기초모형에서는 47.4%인 반면 모형 1-2에서는 23.8%로 수치가 현저히 떨어진 것을 확인할 수 있다. 이는 채널 수준의 설명변수가 적절히 모형에 투입되어 채널 수준의 확률효과가 줄어든 결과로 해석할 수 있다. 결과적으로 이러한 결과들은 모형 1-2가 가장 앞선 모형들보다 설명력이 높은 가장 정교한 모형이며, 채널 수준의 설명변수들에 의해서 조회수의 변화가 타당하게 설명된다는 것을 확인시켜준다.

이를 근거로 모형 1-2에 투입된 채널 수준의 소비 영향요인의 영향력을 계수값 해석을 통해 분석할 수 있다. 우선 모형 1-2에서 투입된 모든 콘텐츠 수준의 설명 변수들이 조회수에 미치는 영향이 모형 1-1과 동일한 수준에서 유의한 것으로 도출되었다. 이는 콘텐츠 수준의 변수들이 적절히 통제되었음을 의미한다. 한편 채널 수준의 소비 영향요인으로는 제작자 요인과 운영 요인 그리고 통제 요인을 분석하였고, 결과는 다음과 같다.

제작자 요인 중 제작자의 성별의 경우 남성을 기준으로 투입하였는데 비공개로 하였을 때가 남성일 경우보다 15.47% 낮은 조회수를 보였으며( $b = -.1547$ ,  $p < .001$ ), 혼성일 경우 46.06% 높은 조회수를 보였다( $b =$

$-.4606$ ,  $p < .001$ ). 반면 여성인 경우 남성인 경우와 유의미한 차이가 없는 것으로 분석되었다( $b = -.0569$ ,  $p > .05$ ). 제작자 요인 중 전문성( $b = -.0515$ ,  $p > .05$ )은 조회수에 유의미한 영향을 주지 않지만 신뢰성( $b = .0141$ ,  $p < .05$ )은 1단위 올라갈 때마다 0.141%의 조회수 상승효과를 주는 것을 분석되었다. 제작자 유형의 경우 제작자가 개인인 경우를 기준변수로 투입하였는데, 개인인 경우보다 방송사인 경우 55.23%( $b = -.5523$ ,  $p < .001$ ), 협회 및 구단인 경우 76.25%( $b = -.7605$ ,  $p < .001$ )의 더 낮은 조회수를 보였다. 반면 유튜브 전문 제작사에서 게시하는 경우는 조회수 획득에서 개인 채널과 유의미한 차이를 발견하지 못하였다( $b = .1380$ ,  $p > .01$ ).

한편 운영 요인 중 다양한 종목을 소재로 영상을 제작하는지 여부(멀티스포츠)에 따른 조회수 획득의 차이를 분석한 결과 다양한 종목을 소재로 할 때 단일종목으로 운영할 때보다 24.35% 조회수가 높은 것으로 확인되었다( $b = .2435$ ,  $p < .001$ ). 그리고 채널에 업로드된 영상의 개수는 1개 늘어날 때마다 조회수가 0.07% 높아졌고( $b = .0007$ ,  $p < .05$ ), 평균 업로드 주기가 하루 빨라질수록 조회수가 0.08%( $b = -.0008$ ,  $p < .001$ ) 높아지는 것을 분석을 통해 확인할 수 있었다.

마지막으로 모형 1-2에서는 채널이 보유한 구독자수와 채널의 운영 기간을 통제변수로 투입하여 채널의 규모와 운영시간에 따른 효과를 통제하고자 하였다. 분석 결과 구독자수가 1% 늘어날 때마다 0.4%( $b = .4013$ ,  $p < .001$ ) 늘어나는 것으로 분석되었고, 채널 운영 기간이 하루 늘어날 때마다 조회수가 0.05%( $b = .0005$ ,  $p < .001$ ) 늘어나는 것으로 분석되었다. 따라서 채널의 운영 기간과 채널 구독자수에 따른 조회수 상승효과가 적절히 통제되었음을 확인할 수 있다.

## 결론 및 논의

### 1. 연구 결론

유튜브 스포츠 콘텐츠 조회수에 영향을 미치는 소비 영향요인에 대한 주요 분석 결과를 요약하면 다음과 같

다. 첫째, 조회수에 영향을 미치는 소비 영향요인은 영상 수준과 채널 수준에 함께 존재 하며, 채널 별로 조회수가 상이한 형태로 분포한다. 둘째, 스포츠 종목과 콘텐츠 유형에 따라 조회수가 차이를 보인다. 특히 기록형, 필드형 표현형 스포츠 종목이 높은 조회수를 보였으며, 레슨 및 정보형 유형과 엔터테인먼트 유형의 콘텐츠가 높은 조회수를 획득하는 것으로 나타났다. 셋째, 영상에서 제공되는 정보의 양에 따라 조회수가 차이를 보였다. 영상의 길이는 중요한 예측 요인이 아니었지만 제공되는 정보의 양은 조회수에 긍정적인 영향을 미치는 것으로 나타났다. 넷째, 썸네일의 선정성과 폭력성이 높을수록 조회수가 높아진다. 이를 통해 자극적인 형태의 썸네일이 조회수를 높이는 데 유의미한 영향을 미치는 것으로 확인 되었다. 다섯째, 콘텐츠의 접근성이 높을수록 조회수가 높아진다. 이를 통해 콘텐츠가 쉽게, 많이 노출될수록 조회수가 높아지는 것을 확인할 수 있었다. 여섯째, 제작자의 성별과 유형에 따라 조회수가 차이를 보이며 제작자에게 형성된 신뢰성이 높을수록 조회수가 높아진다. 성별이 혼성인 경우 가장 높은 조회수 획득을 보였으며, 비공개인 경우와 여성인 경우보다는 남성인 경우가 조회수 획득에 유리한 것으로 분석되었다. 일곱 번째 채널의 운영 형태와 기간에 따라 조회수가 차이를 보였다. 개인이 운영하는 경우와 유튜브 전문 방송사(MCN)에서 운영하는 경우가 평균적으로 가장 높은 조회수를 획득하는 것으로 나타났으며, 콘텐츠 업로드 주기가 짧고, 채널을 오랜 기간 운영한 경우 높은 조회수를 얻는데 유리한 것으로 나타났다.

## 2. 주요 연구 결과 논의

앞선 분석 결과는 유튜브 플랫폼에 게시된 콘텐츠의 조회수에 영향을 미치는 소비 요인들의 유의한 영향력을 보여주었다. 구체적으로 1) 채널 수준과 영상 수준의 소비 영향요인들이 존재하며, 2) 조회수에 영향을 미치는 콘텐츠 수준의 특수한 요인이 있으며, 3) 콘텐츠 수준의 요인들을 통제했을 때 채널 수준의 특수한 소비 영향요인이 있었다. 이와 같은 결과는 아래 일곱 가지 관점에서 논의될 수 있다.

첫째, 영상 수준과 채널 수준의 소비 영향요인이 존재

한다는 것은 콘텐츠를 어떻게 제작하는가뿐만 아니라 누가 제작하는가 또한 중요하다는 것을 시사한다. 본 연구의 결과를 통해 영상 수준의 소비 영향요인들의 설명 변량이 채널 수준의 변량보다 높은 것을 확인할 수 있었는데, 이를 통해 일반적으로는 콘텐츠를 어떻게 제작하는가가 중요하다 할 수 있다. 그렇지만 채널 수준 소비 영향요인들의 설명 변량이 약 30% 이상을 보였고 이를 통해 누가 콘텐츠를 제작하였는가 또한 큰 영향을 미치는 것으로 판단할 수 있다. 특히, 본 연구의 모형에 투입된 채널 수준의 소비 영향요인 변수의 개수가 콘텐츠 수준의 변수 개수보다 적다는 것을 고려하였을 때 이와 같은 예상은 더욱 타당한 것으로 판단할 수 있다.

둘째, 스포츠 요인이 유의미한 소비 영향요인이었으며, 구체적으로 특정 주제와 콘텐츠 유형에 따라 성과지표가 차이를 보인다는 점이다. 이러한 결과는 스포츠 미디어의 특수한 소비 동기와 시청 즐거움이 유튜브 플랫폼에서도 매우 중요하게 작용하며, 기존의 스포츠 미디어 시장의 영향력이 유튜브 플랫폼상에 미치고 있음을 동시에 시사한다. 스포츠 요인인 종목 유형과 미디어 스포츠 성격 그리고 스포츠 콘텐츠 유형에 따라 서로 상이한 성과지표가 형성된 결과를 통해 이를 확인할 수 있다. 이는 스포츠 미디어의 시청 동기와 즐거움에 대해 밝힌 Raney(2006)의 선행연구의 결과를 지지한다고 할 수 있다. 기존의 텔레비전 스포츠 콘텐츠의 시청동기들이 유튜브 플랫폼 상에서도 적용된다는 것은, 플랫폼과 양식이 변화하여도 결국 스포츠라는 주제 속에서 본원적인 시청 즐거움이 형성된다는 것으로 해석할 수 있다.

셋째, 썸네일의 선정성과 폭력성이 높을수록 조회수가 높아진다는 결과는 기존 연구에서는 찾아볼 수 없는 유튜브 콘텐츠의 특수한 소비 영향요인을 검증한 것이라 할 수 있다. 썸네일은 콘텐츠를 대표하는 하나의 사진이자 광고이다. 기존의 미디어에서는 시청자의 선택을 받기 위해서 예고나 편성표를 활용하여 방송 프로그램을 홍보해왔다. 그러나 유튜브 플랫폼 상에서는 하나의 페이지에 수많은 콘텐츠들의 썸네일이 제공되고, 소비자는 썸네일을 기초 정보로 하여 콘텐츠를 선택한다. 각각의 콘텐츠는 하나의 상품이며, 상품은 광고를 통해 소비자의 선택을 유도한다는 점에서 유튜브에서의 콘텐츠 소비는 기존의 커뮤니케이션 행동이론보다는 마케팅 소비자는

행동의 관점으로 해석될 수 있다.

넷째, 콘텐츠의 접근성이 높을수록 조회수가 높아진다는 점 또한 기존의 미디어 시장에서의 소비 증진 방식과는 차이를 보인다는 것을 의미한다. 조회수에 검색어 개수와 타 SNS 링크 개수가 모두 정(+)의 유의미한 영향을 미친다는 결과는 스포츠 콘텐츠의 노출가능성이 성과 지표에 유의미한 영향을 미친다는 것을 증명한다. Duffett과 동료들(2019)은 이와 같은 콘텐츠 접근성은 성과가 높아질수록 더 높아지는 것으로 보고 하기도 하였는데, 한번 관심을 받은 콘텐츠들은 공유의 재공유를 거쳐 확산되어져 더욱 넓은 접근성을 형성하기 때문이다. 이를 통해 본 연구의 결과가 관련 선행연구의 연구결과를 지지하는 것을 확인할 수 있다.

다섯째, 논의점은 제작자의 성별과 유형, 신뢰성에 따라 조회수가 차이를 보였다는 것이다. 이는 유튜브 플랫폼 스포츠 콘텐츠 소비의 특수한 채널 효과가 있다는 점을 의미한다. Jenkins, Bugeja 그리고 Barber(2014)는 스포츠 지도자의 성별에 따른 역할이 사회적 역할 이론에 따라 학습자에게 상이하게 인지된다는 점에 주목하여 스포츠 학습의 효과가 상이함을 밝힌 바 있다. 스포츠 유튜브 콘텐츠에서 정보를 전달하는 화자는 스포츠와 관련된 정보를 전달하거나 즐거움을 제공하는 역할을 수행하는 데 이때 성별에 따른 차이가 있음은 이와 같은 맥락에서 해석될 수 있다. 또한 채널 유형에 따른 조회수 차이를 발견한 본 연구의 결과는 소비자가 인지한 채널의 제작 규모와 전문성에 따라 소비 행동의 차이를 보인다는 것을 채널 브랜드 후광효과를 통해 주장한 Welbourne과 Grant(2016)의 연구 결과와 일치하는 결과이다. 즉 유튜브 스포츠 콘텐츠 또한 보편적인 채널 후광효과가 있음을 확인할 수 있다. 그러나 Ohanian(1990), 유현주, 김현(2019) 등이 정보원 공신력 모델을 통해 주장한 화자의 공신력이 소비자의 행동에 미치는 긍정적인 영향력은 본 연구의 결과와는 완벽히 일치하지 않았다. 그 이유는 자료의 편향성과 관련된다. 본 연구에서는 스포츠와 관련된 유튜브 채널을 연구 대상으로 설정하였는데, 자료수집에 있어 구독자 수와 조회수를 기준으로 상위의 150개의 채널을 선정하였다. 상위의 유튜브 채널들은 대다수의 코너들에게 신뢰성과 전문성을 일정부분 확보되어 있었고, 때문에 내용 분석을 통해 작성

된 전문성의 변량이 충분하지 않았기 때문이다.

마지막으로, 채널의 운영 형태와 기간에 따라 성과지표가 차이를 보였다는 결과는, Coursaris와 Osch(2016)가 주장한 콘텐츠 업로드양, 업로드 주기와 같은 일종의 운영 지표들이 채널을 신뢰하게 하고 호의적인 태도를 형성하여 소비 행동을 증진한다는 선행 연구의 결과와 일치한다. 결국 평균 콘텐츠 업로드 주기가 짧고 많은 영상을 업로드하고, 운영 기간이 오래된 채널들의 콘텐츠들이 높은 성과를 보인다는 점은 유튜브 스포츠 채널 운영자의 성실함이 성과에 긍정적인 영향을 미친다는 점을 의미한다.

## 연구의 의의 및 제언

### 1. 연구의 의의 및 기여도

본 연구의 학술적 의의와 실무적 의의는 아래와 같다. 본 연구의 학술적 의의의 첫 번째는 선행 연구들에서 주로 이뤄진 스포츠 미디어 시청 동기와 즐거움과 같은 특수성을 바탕으로 유튜브 스포츠 콘텐츠 소비를 해석함과 동시에 큰 틀에서 커뮤니케이션 이론을 활용했다는 점이다. 특히 스포츠가 가진 고유 특성과 유튜브라는 새로운 커뮤니케이션이 만들어내는 새로운 현상을 해석함에 있어 기존의 관점을 수용하고 새로운 방식으로 해석하였다는 점은 해당 분야 연구의 연속성과 확장의 개념을 충족시킨다는 점에서 의의가 있다.

두 번째로 조회수에 미치는 영향요인(독립변수)을 스포츠 요인과 함께 콘텐츠 수준과 채널 수준으로 구분하여 동시에 검증했다는 점이다. 조회수나 댓글수와 같은 유튜브 플랫폼에서 콘텐츠가 달성한 성과에 영향을 미치는 요인을 연구한 연구들 중 콘텐츠 내용의 특수성에 기반하여 변수를 도출한 경우는 찾아보기 힘들다. 또한 다양한 영향요인을 체계적으로 구성한 경우 또한 없었다. 주로 한 가지 영향요인에 대하여 검증하는 경우가 많은데, 성과지표는 특정 한 가지 변수에 의해 영향을 받지 않는다. 본 연구는 특정 변수의 누락 효과를 효율적으로 통제하고 분석 결과의 타당성을 높이기 위해 콘텐츠 단

위와 채널 단위의 변인을 동시에 고려하여 변수로 설정하고 분석하였다. 이에 따라서 특정 변인에 한정된 분석을 진행해 온 선행 연구들의 한계를 보완하였다는 점에서 학술적 의의가 있다.

세 번째로 본 연구의 연구문제를 검증하기 위해 사용된 자료들이 웹상에 형성된 실제 데이터를 대상으로 하였다는 점이다. 기존 연구들에서는 주로 설문이나 실험을 통해서 소비 행동에 영향을 미치는 요인을 검증해 왔다. 그러나 본 연구에서는 웹 크롤링(Web Crawling) 기법을 활용하여 실제 웹상에 형성된 데이터를 수집하였다는 점과 10만 개에 가까운 대량의 자료를 수집했다는 점에서 기존의 연구들과는 차이를 보인다. 이는 일종의 2차 자료를 이용한 것인데, 이러한 방식은 설문이나 실험에서 흔히 발생할 수 있는 응답자의 기억과 인식 감정에 의존한 부정확한 응답에 의한 왜곡현상을 방지한다는 차원에서 보다 타당성있는 결과를 도출해 준다. 웹 크롤링을 이용한 빅데이터 연구가 필요함에도 불구하고 프로그램 코드작성 학습의 진입장벽이 높아 해당 영역에서 잘 사용되지 않고 있다는 점, 그리고 대량의 자료를 이용하여 결과의 신뢰성과 타당성을 높였다는 점은 체육학 연구에서 높은 학술적 의의를 가진다 할 수 있다.

마지막으로 본 연구는 체육학 선행 연구에서 자주 사용되지 않았던 다층모형을 이용하여보다 정교한 분석을 실시하였다는 점이다. 구체적으로 본 연구는 주로 다중 회귀분석이나 구조방정식, 분산분석 등과 같은 통계 방식 보다 고도화된 다층모형 분석을 이용하여 해당 분야 학술 연구 방법의 정교화에 기여하였다고 할 수 있다. 유튜브 스포츠 채널과 이에 배속(nested)되어 있는 콘텐츠로 자료를 구성하고, 해당 콘텐츠와 채널에 형성된 다양한 양적 질적 자료들을 변수화하여 체계적으로 다층모형에 투입하였는데, 이는 일반적으로 사용되는 회귀분석의 최소자승법 방식(ordinary least square)에서 발생하는 문제점들을 해결한다는 점에서 의미가 있다. 구체적으로 다층모형은 표본크기나 다변량 정규성 가정을 위반해 추정되지 않는 케이스들을 처리할 수 있으며, 불균형 자료를 추정하는 데서 오는 문제점 또한 베이지안(Bayesian) 추정 방법을 통해 해결할 수 있다.

한편 본 연구의 결과는 스포츠 미디어 시장의 새로운 경향을 파악하여 다양한 이해 당사자의 시장 전략에 필

요한 기초정보를 제공한다는 점에서 실무적 의미가 있다. 유튜브 스포츠 콘텐츠의 소비 영향요인이 콘텐츠 수준과 채널 수준에 각각 존재한다는 결과는 결국 플랫폼 광고주들이 유튜브 크리에이터와의 계약이나 광고 삽입을 의뢰할 때 종합적으로 고려해야 할 요인을 체계적으로 도출할 수 있다는 점에서 실무적 의미가 있다. 광고는 제품이나 회사가 가진 가치나 이미지가 매개되는 미디어, 콘텐츠와 일치성이 높아야 하는 것은 물론이며 광고가 얼마나 많은 사람에게 반복적으로 노출될 것인가에 따라 효과가 차이를 보인다. 즉, 유튜브 플랫폼에서 스포츠와 관련된 콘텐츠의 소비 영향요인은 유튜브 플랫폼을 활용한 광고나 판촉활동을 계획하는 광고주에게 효과적인 광고를 위해 고려해야 할 요인을 과학적인 근거를 통해 제공한다는 점에서 실무적인 의미가 크다.

## 2. 제언

본 연구는 총 다섯 가지 관점에서 한계점을 가지며, 각각의 한계점에 따른 추후 연구 방향을 아래와 같이 제시하였다.

첫째, 수집된 자료들의 게시 경과시간에 대한 통제가 좀 더 효과적으로 이뤄져야 한다는 점이다. 본 연구에서는 다층모형을 구성하기 위해 각각의 채널에 게시된 콘텐츠들을 분석 자료로 수집하였고, 이 과정에서 6개월이 지나지 않은 자료들은 게시 경과시간에 따른 성과 획득의 차이를 통제하기 위해 대상에서 제외하였다. 또한 게시 경과시간에 따른 성과 획득의 차이를 보다 효과적으로 통제하기 위해 콘텐츠 게시 경과시간을 통제변수로 설정하여 모형에 투입하기도 하였다. 선행 연구에서는 이와 같은 문제를 해결하기 위해 웹상에 게시된 콘텐츠의 활성수명(Active life span)을 분석하기도 한다. 후속 연구에서는 이와 같은 통제 작업이 원활히 이뤄져야 한다.

두 번째는 이론적으로 구성된 변수를 활용하여 더욱 정교한 다층모형을 테스트할 필요성이 있다는 점이다. 본 연구에서는 콘텐츠와 채널 수준에서 발생하는 다양한 소비 영향요인을 변수로 구성하여 확률 절편 모형(Random Intercept model)을 이용한 다층모형분석을 실시하였다. 그러나 변수 간의 관계 그리고 상호작용 가능성 등을 고려하여 확률 기울기 모형(Random Slope

model) 등을 구성하여 변수 간의 관계를 검증할 수도 있다.

세 번째는 본 연구의 범위의 문제를 들 수 있다. 본 연구는 유튜브 스포츠 채널이라는 다소 넓은 범주에서 연구를 진행하였다. 그러나 본 연구에서도 언급된 바와 같이 스포츠는 종목마다 서로 상이한 경기 방식을 가지고 있고, 미디어스포츠 내용의 전개 방식도 상이하다. 나아가 종목마다 문화적 정착 수준도 상이하기 때문에 세부적인 소비 영향요인의 측정이 필요하다. 구체성 있는 연구 대상의 설정과 이론적인 탐색을 통해서 상세한 관계를 측정하여 더욱 타당성 있는 연구 결과를 도출할 필요성이 있다.

마지막으로 소비 영향요인과 성과지표의 관계를 설명할 수 있는 세부 이론이 필요하다. 유튜브와 같은 OTT 플랫폼에서의 소비 영향요인과 성과지표의 관계를 설명한 이론을 이용하여 각각의 특수한 상황을 보다 정교하게 설명할 필요가 있다. 이러한 이론적 프레임워크는 연구 결과에 대한 타당한 설명과 동시에 보편적인 정보를 도출하는데 필수적인 과정이다. 본 연구는 급변하는 미디어 소비 환경에서 유튜브를 대표적인 예시로 설정하여 탐색적인 연구를 수행하였다. 이러한 환경에서의 소비자 행동을 설명하기 위해 전통적인 커뮤니케이션 이론과 선행 스포츠 미디어 소비 연구들을 인용하였다. 커뮤니케이션 이론과 선행 스포츠 미디어 연구들은 오랜 시간을 거듭하며 인용되고 검증되어져 왔기 때문이다. 그렇지만 오늘날 미디어 환경과 소비자에 적합한 형태의 이론이 개발되고 검증되어 더욱 타당한 가설과 연구 결과, 그리고 이에 따른 정교한 해석이 향후 연구에서 필요할 것이라 판단된다.

## 참고문헌

- 강명현. (2018). 방송 콘텐츠 가치평가 지표의 속성 및 시청률과의 관계 연구. **한국방송학보**, 32(3), 5-30.
- 김기한, 윤리라, 방신웅. (2012). 스포츠 전문 채널의 프로그램 유형과 종목에 따른 방송 비중, 편성의 다양성 및 시청률 분석. **한국스포츠산업경영학회지**, 17(5), 49-66.
- 김원제. (2005). **미디어스포츠사회학**, 서울: 커뮤니케이션북스.
- 녹스 인플루언서 (2021. 3). **TOP 100 스포츠 유튜브 채널**. 2021. 3 검색. <https://kr.noxinfluencer.com/youtube-channel-rank/top-100-all-sports-youtuber-sorted-by-subs-weekly>.
- 닐슨 (2020. 1. 5). 2019년 4분기 포털/SNS 분기보고서. **Portal/SNS Quarterly Report**. 2020. 10. 31 검색. [http://www.koreandclick.com/english/insights/qreports\\_view.html?id=138&code=portal&page=1](http://www.koreandclick.com/english/insights/qreports_view.html?id=138&code=portal&page=1).
- 박병언, 임규건. (2015). 일반영향요인과 댓글기반 콘텐츠 네트워크 분석을 통합한 유튜브(Youtube)상의 콘텐츠 확산 영향요인 연구. **지능정보연구**, 21(3), 19-36.
- 박보현. (2008). 2006년도 지상파 TV 스포츠 중계방송 프로그램 편성 분석. **한국콘텐츠학회논문지**, 8(6), 204-212.
- 이용관. (2019). 1인 미디어 크리에이터 활동 여건 분석. **문화경제연구**, 22(1), 59-80.
- 이주호, 김기한. (2019). 스포츠용품 브랜드 SNS 광고의 크리에이티브 전략에 대한 효과 분석, **한국광고홍보학보**, 21(1), 69-102.
- 유현주, 김현. (2019). 1인 미디어 크리에이터 콘텐츠의 특성이 소비자 태도 및 구매 의도에 미치는 영향: 크리에이터의 공신력 및 콘텐츠 진정성, 브랜드 진정성을 중심으로. **한국방송학회 학술대회 논문집**, 115-117.
- 정영남. (2008). **미디어스포츠**, 서울: 대한미디어
- Akrouf, S., Meriem, L., Yahia, B., & Eddine, M. N. (2013). Social Network Analysis and Information propagation: A case study using Flickr and YouTube networks. *International Journal of Future Computer and Communication*, 2(3), 246-252.
- Alavy, K., Gaskell, A., Leach, S., & Szymanski, S.

- (2010). On the edge of your seat: Demand for football on television and the uncertainty of outcome hypothesis. *International Journal of Sport Finance*, 5(2), 75-95.
- Berlo, D. (1960). *SMCR Communication Model. The Process of Communication*. San Francisco: Holt, Rinehart and Winston.
- Birrell, S., & Loy Jr, J. W. (1979). Media Sport: Hot and Cool1. *International Review of Sport Sociology*, 14(1), 5-19.
- Brynjolfsson, E., & Oh, J. (2012). The attention economy: measuring the value of free digital services on the Internet. *Proceedings of the International Conference on Information Systems*.
- Burgess, J., & Green, J. (2018). *YouTube: Online video and participatory culture*. John Wiley & Sons.
- Chaffey, D., & Smith, P. R. (2017). *Digital marketing excellence: planning, optimizing and integrating online marketing*. Taylor & Francis.
- Cheng, X., Dale, C., & Liu, J. (2007). Understanding the characteristics of internet short video sharing: YouTube as a case study. *arXiv preprint arXiv:0707.3670*.
- Cheng, X., Dale, C., & Liu, J. (2008). Statistics and social network of youtube videos. *2008 16th International Workshop on Quality of Service*.
- Coursaris, C. K., & Van Osch, W. (2016). Exploring the effects of source credibility on information adoption on YouTube. *HCI in Business, Government, and Organizations*.
- Cunningham, N., & Bright, L. F. (2012). The Tweet is in Your Court: Measuring Attitude Towards Athlete Endorsements in Social Media. *International Journal of Integrated Marketing Communications*, 4(2).
- Duffett, R., Petroșanu, D.-M., Negricea, I.-C., & Edu, T. (2019). Effect of YouTube marketing communication on converting brand liking into preference among millennials regarding brands in general and sustainable offers in particular. *South Africa and Romania. Sustainability*, 11(3), 604.
- Erdogan, B. Z. (1999). Celebrity endorsement: A literature review. *Journal of marketing management*, 15(4), 291-314.
- Gantz, W., & Wenner, L. A. (1995). Fanship and the television sports viewing experience. *Sociology of Sport Journal*, 12(1), 56-74.
- Gupta, H., Singh, S., & Sinha, P. (2017). Multimedia tool as a predictor for social media advertising - a YouTube way. *Multimedia tools and applications*, 76(18), 18557-18568.
- Hoiles, W., Krishnamurthy, V., & Aprem, A. (2016). PAC algorithms for detecting Nash equilibrium play in social networks: From Twitter to energy markets. *IEEE Access*, 4, 8147-8161.
- Hovland, C. I., Janis, I.L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.
- Jenkins, J. S., Bugeja, A. D., & Barber, L. K. (2014). More content or more policy? A closer look at syllabus detail, instructor gender, and perceptions of instructor effectiveness. *College Teaching*, 62(4), 129-135.
- Khan, M. L. (2017). Social media engagement:

- What motivates user participation and consumption on YouTube? *Computers in Human Behavior*, 66, 236-247.
- Manzerolle, V., & Wiseman, A. (2016). On the transactional ecosystems of digital media. *Communication and the Public*, 1(4), 393-408.
- McGuire, W. J. (1985). *Attitudes and attitude change*. The handbook of social psychology, 233-346.
- McLean, D., Hurd, A., & Brattain-Rogers, N. (2005). Philosophy of recreation and leisure. *Kraus' recreation and leisure in modern society*, 319-348.
- Ohanian, R. (1990). Construction and validation of a scale to measure celebrity endorsers' perceived expertise, trustworthiness, and attractiveness. *Journal of advertising*, 19(3), 39-52.
- Paek, H.-J., Hove, T., Jeong, H. J., & Kim, M. (2011). Peer or expert? The persuasive impact of YouTube public service announcement producers. *International Journal of Advertising*, 30(1), 161-188.
- Paul, R. J., & Weinbach, A. P. (2007). The uncertainty of outcome and scoring effects on Nielsen ratings for Monday Night Football. *Journal of Economics and Business*, 59(3), 199-211.
- Ramaprasad, J., & Hasegawa, K. (1992). Creative strategies in American and Japanese TV commercials: A comparison. *Journal of advertising research*, 32(1), 59-67.
- Raney, A. A. (2003). *Disposition-based theories of enjoyment*. In *Communication and emotion*. Routledge.
- Raney, A. A. (2006). *Why we watch and enjoy mediated sports*. *Handbook of sports and media*. Taylor & Francis.
- Raney, A. A. (2013). Reflections on communication and sport: On enjoyment and disposition. *Communication & Sport*, 1(1-2), 164-175.
- Shank, M. D., & Lyberger, M. R. (2014). *Sports marketing: A strategic perspective*. Routledge.
- Stauff, M. (2009). Sports on YouTube. *The YouTube Reader*, 236-251.
- Stewart, D. W., & Furse, D. H. (1986). *Effective television advertising: A study of 1000 commercials*. Lexington Books.
- Susarla, A., Oh, J.-H., & Tan, Y. (2012). Social networks and the diffusion of user-generated content: Evidence from YouTube. *Information Systems Research*, 23(1), 23-41.
- Welbourne, D. J., & Grant, W. J. (2016). Science communication on YouTube: Factors that affect channel and video popularity. *Public understanding of science*, 25(6), 706-718.
- Yoganarasimhan, H. (2012). Impact of social network structure on content propagation: A study using YouTube data. *Quantitative Marketing and Economics*, 10(1), 111-150.

## ABSTRACT

### Determinants of Consumption for YouTube Sport Contents: A Big Data Analysis using Multi-level Regression

Kim, Jongho (Seoul National University) ·  
Kim, Kihan (Seoul National University & SNU Institute of Sports Science)

This study analyzed the consumption determinants of Youtube sports video content. Using big data analytics, 124 Youtube sport channels and 89,834 video contents uploaded on the channels were empirically examined. Independent variables that can induce sport video consumption behavior were constructed at the video content level and Youtube channel level, and their effects on the number of view count was analyzed through Multi-Level Regression. In particular, the main independent variables of this study were constructed at two levels respectively. Specifically variables at the level of video contents included sport factors (type of sport, content types), content creative factors (videos, thumbnails), and accessibility factors (number of search words, number of other SNS links), the level of video channels included creator factors(credibility, professionalism, gender, types), and operating factors(number of contents posted, upload cycle, variety of subtitles). The main results of this study are as follows: First, at the content level, sport factors had a significant effect on view count. Second, the amount of video information, the form of thumbnails, and content accessibility had an effect on view count. In addition, with regards to the channel level, the gender and type of creators, perceived reliability, channel types and upload cycle had a significant influence on the number of view count. The results provide valuable information on what factors need to be prioritized to obtain high view count when creating content and operating a channel. This has practical implications in that it provides useful information to sports content creators. In addition, it is of academic significance in that the study contributes to the existing sport media research on Youtube by implementing the big data framework and analysis.

※ Key words : Youtube, Sports contents, Sports media, Big data, Multi level regression, Web Crawling

논문 투고일 : 2021. 04. 30.

논문 심사일 : 2021. 05. 15.

게재 확정일 : 2021. 06. 30.