『한국사회학』 제60집 제2호 (2026년), pp. 119~158 https://doi.org/10.21562/kjs.2026.05.60.2.119

연구논문

플랫폼 식별가능성과 논란 담론의 감성

디시인사이드 케이팝 아이돌 갤러리 분석*

조 성 권** · 김 민 석*** · 김 재 훈**** · 구 정 우*****

케이팝 아이돌이 글로벌 문화 브랜드로 부상하면서 한국 사회는 이들에게 높은 규범적 기대를 부여했다. 그 결과 아이돌의 사회적 논란은 개인적 차원의 사건에 그치지 않고 공적 논의의 대상으로 전환되곤 한다. 본 연구는 이러한 배경에서 온라인 공간이 제공하는 참여 방식의 차이가 논란 담론의 감성적 톤에 어떠한 차이를 만들어내는지를 실증적으로 분석한다. 이를 위해 디시인사이드의 고정닉/유동닉 구조에 주목하여, 발화자의 식별가능성 수준이 게시물의 감성 방향과 강도에 어떠한 관련을 갖는지를 검토하였다. 구체적으로 디시인사이드의 2세대 이후 케이팝 아이돌 관련 57개 갤러리에서 논란 관련 게시물을 선별하고, 감성분석을 통해 게시물 단위의 감성 방향과 강도를 점수화한 뒤 다중회귀분석을 수행하였다. 분석 결과, 식별가능성이 낮은 유동닉 게시물은 고정닉 게시물에 비해 전반적으로 더 낮은 감성 점수를 보였으며, 이러한 경향은 가치·규범적 차원의 갈등을 야기할 수 있는 논란 유형에서 더욱 두드러졌다. 이는 사회적 논란에 대한 온라인 반응이 플랫폼이 제공하는 참여 구조와 그에 따라 달라지는 발화 비용과도 연결될 수 있음을 시사한다. 본 연구는 온라인 공간에서의 식별가능성을 플랫폼이 제도화한 참여 조건으로 파악함으로써, 디지털 공론장에서 논란 담론의 감성이 사회적 차원으로 조직될 수 있다는 경험적 틀을 제시한다.

주제어: 케이팝 아이돌, 사회적 논란, 식별가능성, 디시인사이드, 감정분석

* 이 논문은 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구이며(NRF-2024S1A5C3A01043455), 대한민국 과학기술정보통신부와 정보통신기획평가원의 지원을 받아 수행된 연구임(IITP-2025-No.RS-2024-00436936).

** 주저자, 성균관대학교 BK21 소셜이노베이션 융합전공 박사후연구원(crusian@skku.edu)
*** 성균관대학교 BK21 소셜이노베이션 융합전공 박사과정(gms80175064@gmail.com)
**** 한양대학교 인공지능학과 석박통합과정(jaehoonkimm@hanyang.ac.kr)
***** 교신저자, 성균관대학교 사회학과 교수(jkoo@skku.edu)

Ⅰ. 서론

케이팝 아이돌이 글로벌 문화 브랜드를 대표하는 상징적 존재로 부상함에 따라, 이들을 둘러싼 사회적 논란 역시 단순한 연예계 이슈를 넘어 사회적 공론의 대상이 되고 있다. 한국리서치(2024)의 조사에 따르면, 연예인의 생활 중 범죄 혹은 비윤리적인 행동의 경우 알 권리라는 응답이 70% 이상일 정도로 논란이 되는 이슈에 있어서의 관심도 높다. 이렇듯 대중은 아이돌을 단순한 오락이나 유희의 제공자가 아니라 일정한 공적 책임과 도덕적 기대가 부과되는 존재로 인식하는 경향을 보인다. 또한 음주나 연애와 같은 사적 행위나 사소한 일탈에 대해서도 낮은 관용을 드러낸다(Kang and Park, 2022). 아이돌 관련 논란은 소셜미디어와 온라인 커뮤니티를 경유하며 빠르게 확산되고, 당사자에 대한 비판과 평가, 해명 요구와 조롱, 지지 철회와 같은 반응을 동반한다. 이처럼 아이돌 논란은 특정 사건에 대한 대중의 일시적 반응에 그치기보다, 규범적 기대와 공적 평가가 온라인 공간에서 어떻게 조직되고 증폭되는가를 보여주는 사례라고 할 수 있다(Aitchison and Meckled-Garcia, 2021; Forestal, 2024; Hier, 2008; Ivaturi and Chua, 2019; Rost, Stahel, and Frey, 2016; Williams, 2007).

기존 연구들은 이러한 논란의 양상을 설명하는 과정에서 아티스트의 성별, 국적, 인지도 혹은 논란에 대응하는 소비자의 속성에 주로 주목해 왔다(Ahn and Lin, 2019; Lee, Chang, and Einwiller, 2020; Rifon, Jiang, and Wu, 2023). 이러한 접근은 어떤 대상이 더 비판받기 쉬운지, 혹은 어떤 소비자가 더 강하게 반응하는지를 설명하는 데 유용하다. 그러나 최근 아이돌 논란의 주요 공론장이 온라인 공간으로 이동하고 있음에도, 온라인 플랫폼의 구조적 조건이 이러한 반응의 형식과 수위를 어떻게 달리 만드는지는 충분히 설명하지 못했다. 온라인 공간에서의 논란 담론은 동일한 사건이라 하더라도 반복적 상호작용에 따른 평판 비용, 플랫폼 내부의 규범과 위계, 발화자의 식별가능성에 따라 상이하게 조직될 수 있다(Cho and Kwon, 2015; Friedman and Resnick, 2001; Moore, Fredheim, Wyss, and Beste, 2021). 따라서 디지털 공간에서 나타나는 비판과 평가의 감성적 반응은 개인의 일탈적 공격성이나 특정 사건의 속성으로만 환원하기보다, 플랫폼 환경 아래에서 이루어지는 집합적 상호작용의 산물로 이해할 필요가 있다(Rösner and Krämer, 2016;

Vilanova, Beria, Costa, and Koller, 2017).

본 연구는 이러한 문제의식에서 출발하여, 케이팝 아이돌의 사회적 논란에 대한 온라인 반응이 플랫폼의 식별가능성 구조에 따라 어떻게 달라지는지를 실증적으로 분석한다. 여기서 식별가능성은 단순히 실명 여부를 뜻하는 것이 아니라, 동일한 참여자가 반복적으로 인식되고 발화의 기록이 누적될 가능성을 포함하는 개념이다. 이 관점에서 보면, 온라인 논란 담론의 감성적 반응은 사건의 심각성이나 개인의 성향에 의해서만 결정되기보다, 플랫폼이 제공하는 식별가능성의 차이와도 체계적으로 연결될 수 있다.

이를 위해 본 연구는 국내 최대 온라인 커뮤니티 가운데 하나인 디시인사이드(DCinside)의 아이돌 관련 갤러리를 분석의 장으로 삼는다. 디시인사이드의 아이돌 갤러리는 아이돌 논란에 대한 온라인 반응이 집중적으로 표출되는 공간인 동시에, 고정닉과 유동닉이라는 서로 다른 참여 방식이 공존한다는 점에서 식별가능성의 차이를 경험적으로 비교하기에 적합하다. 고정닉은 동일한 닉네임과 계정을 기반으로 활동 기록이 누적되기 때문에 반복적으로 식별될 가능성이 높은 반면, 유동닉은 로그인 없이 일시적으로 참여하는 방식이어서 발화의 지속성과 추적 가능성이 상대적으로 낮다. 본 연구는 고정닉과 유동닉을 단순한 이용자 선택의 차이가 아니라, 발화에 수반되는 평판 비용과 책임의 정도가 달라지는 식별가능성의 차이로 이해한다.

선행연구에 따르면 사용자는 자신의 활동이 누적되고 추적될 수 있는 조건과 그러한 추적이 어려운 조건에서 서로 다른 태도와 행동을 보일 가능성이 높다(송준모·강정한, 2018; 황병화·이은규, 2025). 그런 점에서 디시인사이드의 고정닉/유동닉 구조는 플랫폼의 식별가능성 차이가 집단 규범 동조와 감성 표현의 수위를 어떻게 강화하거나 제약하는지를 관찰하기에 적합하다(강옥희, 2023; 권지현·김명혜, 2014; 송경재, 2005; 유지연, 2020; 황병화·이은규, 2025). 즉, 디시인사이드는 동일한 커뮤니티 내부에서 식별가능성의 차이가 감성 표현의 형식과 수위에 어떤 차이를 만드는지를 비교할 수 있는 사례이다.

구체적으로 본 연구는 케이팝의 생산·유통·수용이 유튜브와 소셜미디어, 온라인 커뮤니티를 중심으로 재조직되기 시작한 동시에, 온라인 공간에서 아이돌 논란이 공론화되는 양상이 보다 뚜렷하게 나타나는 2세대 이후 케이팝 아이돌을 대상으로 분석을 진행한다. 이를 위해 디시인사이드 57개 갤러리에서 웹크롤링을 통해 수집한 714,778건의 게시물을 분석의 출발점으로 삼았다. 먼저 KcELECTRA 기반 딥러닝 텍스트 분류기를 적용하여 사회적 논란과 관련된 게시물 122,277건을 선별하고,

이를 최종 분석 데이터셋으로 구축하였다. 또한 해당 게시물에 KOTE(Korean Online That-gul Emotions)를 적용하여 각 게시물의 세부 감정 확률을 산출하고, 이를 바탕으로 감성 방향과 강도를 요약한 게시물 단위 감성 점수를 구성하였다(Jeon, Lee, and Kim, 2022). 이후 게시물 작성자의 참여 방식(고정닉/유동닉), 즉 식별가능성 수준을 핵심 독립변수로 설정하고, 관련 통제변수를 포함한 다중회귀분석을 통해 식별가능성의 차이가 게시물의 감성 점수와 어떠한 관련을 갖는지를 검토하였다. 특히 본 연구는 평균적인 차이를 확인하는 데 그치지 않고, 식별가능성의 효과가 사회적 논란의 유형에 따라 상이하게 나타나는지도 함께 살펴본다.

본 연구의 의의는 크게 세 가지로 제시할 수 있다. 첫째, 케이팝 아이돌 논란에 대한 사회적 대응을 개인의 감정 표출이 아니라 온라인 공론장에서 논란 담론의 감성적 반응이 조직되는 과정으로 재구성한다. 둘째, 고정닉과 유동닉의 차이를 단순한 이용자 선택이 아니라 플랫폼이 제도화한 식별가능성의 차이로 파악함으로써, 발화에 수반되는 평판 비용과 책임 구조가 온라인 감성 표현에 미치는 영향을 경험적으로 검토한다. 셋째, 대규모 텍스트 자료와 딥러닝 기반 감성분석을 결합하여 아이돌 논란 연구를 보다 체계적인 경험적 자료 위에서 확장한다.

Ⅱ. 선행연구 검토

1. 케이팝 아이돌 논란의 도덕적 기대와 공적 평가

연예인 논란이 폭발적 비난과 제재 요구로 이어지는 현상은 현대 대중문화에서 반복적으로 관찰되어 왔다. 특히 유명인의 일탈은 개인적 사건에 그치지 않고, 사회적으로 용인되는 행위의 경계를 둘러싼 공적 평가의 대상으로 재구성되곤 한다. 도덕적 공황 연구는 이러한 과정을 사회질서 위반의 상징을 둘러싼 집단적 불안과 비난의 동원으로 설명해 왔으며(Cohen, 2019; Critcher, 2008; 2009), 최근 연구는 이러한 대중적 반응이 온라인 환경에서 댓글, 공유, 캡처, 재유포와 같은 일상적 상호작용을 통해 더 빠르게 확산될 수 있음을 보여준다(Haugh, 2022; Lageson and Maruna, 2018). 또한 도덕적 일탈을 저지른 연예인에 대해 한국 소비자가 보이는 비난과 지지 철회 양상은 규범적 기대가 실제 소비와 평가의 차원에서도 강하게 작동함을 시사한다(Kang and Park, 2022; Lee and Abidin, 2024).

케이팝 아이돌은 일반적인 연예인 범주 안에서도 특히 높은 도덕적 기대와 이미지 관리의 대상이 되어 왔다. 이들은 한국 사회에서 단순한 오락 제공자를 넘어 건전성, 규율, 모범성의 기대를 부여받는 존재로 자리매김해 왔으며(오세정, 2012), 이들의 일탈이나 논란은 사적인 실수라기보다 집단의 체면을 훼손하거나 공적 이미지를 손상시키는 사건으로 읽히곤 한다(김수정·김수아, 2015). 또한 아이돌은 개인적 재능이나 스타성만으로 소비되는 존재라기보다, 훈련된 자기관리, 팀워크, 팬과의 지속적 소통, 공적 이미지의 안정적 유지가 함께 요구되는 문화적 상품으로 구성되어 왔다. 이 때문에 아이돌을 둘러싼 평가는 작품이나 퍼포먼스의 수준에만 머무르지 않고, 일상적 태도와 사적 행위, 집단 내부의 관계, 팬과 대중에 대한 책임감까지 포괄하는 경향이 강하다.

한편 연예인이 야기하는 논란에 대중이 어떻게 반응하는지와 관련하여, 기존 연구는 주로 아티스트의 성별, 국적, 인지도 혹은 논란에 반응하는 소비자의 속성에 주목해 왔다(Ahn and Lin, 2019; Lee et al., 2020; Rifon et al., 2023). 이러한 접근은 어떤 대상이 더 비판받기 쉬운지, 혹은 어떤 조건에서 대중의 반응이 달라지는지를 설명하는 데 유용하다. 그러나 논란이 실제로 유포되고 확산되는 온라인 공간의 조건, 그리고 그 안에서 반응이 어떤 감성적 형식으로 표출되는지에 대해서는 상대적으로 충분히 설명하지 못했다. 온라인 공간에서 아이돌 논란은 팬, 비판적 팬, 관전자, 안티팬 등 다양한 행위자가 서로 다른 위치에서 개입하는 담론의 장으로 형성되며, 그 반응 역시 단일한 방향으로 수렴하지 않는다.

이 점에서 최근 팬덤 연구는 중요한 시사점을 제공한다. 온라인 플랫폼 환경에서 팬 활동이 일상화되고 분화되면서, 열성적 지지자뿐 아니라 비판적 팬, 관전자, 안티팬 등 다양한 행위자가 동일한 대상에 관여하게 되었기 때문이다(Abd-Rahim, 2019; Duffett, 2013; King-O'Riain, 2021; Lee, Tak, Kwak, and Lim, 2020). 특히 논란 국면에서는 감시, 조롱, 비난, 해명 요구, 정보 재유포와 같은 부정적 참여 양식이 강화될 수 있으며, 이는 다크 팬덤 또는 안티팬덤 논의와도 접점을 가진다(Aburime, 2022; Click, 2019; Gray, 2003; 2007; Wang and Ge, 2023). 이러한 맥락에서 케이팝 아이돌 논란은 단일한 집단의 반응이라기보다, 서로 다른 위치의 참여자들이 개입하는 다층적 담론 공간으로 이해할 수 있다.

요컨대 케이팝 아이돌 논란은 높은 도덕적 기대가 집중되는 사건인 동시에, 팬과 비팬, 관전자와 안티팬이 함께 개입하는 온라인 담론의 장이다. 이 때문에 논란에 대한 반응을 특정 집단의 속성이나 개인의 성향만으로 설명하기는 어렵다. 오히려

중요한 것은 이러한 반응이 어떤 온라인 환경에서, 어떠한 참여 조건 아래에서, 어떠한 감성적 형식으로 표출되는가이다.

2. 온라인 플랫폼의 식별가능성과 부정적 감성 표현

온라인에서의 논란은 댓글, 공유, 캡처, 재유포를 거치며 빠르게 확산되는 경향이 있다. 또한 현재 진행 중인 논란뿐 아니라 과거의 사건이 폭로나 제보를 통해 새롭게 드러나거나, 한동안 잊혔던 이슈가 다시 소환되기도 한다. 이 과정에서 부정적 평가는 반복적으로 가시화되고, 특정 사건에 대한 해석과 감성이 집단적으로 강화될 수 있다. 일부 연구는 이러한 현상이 온라인 공간의 높은 확산성과 기록 가능성 속에서 공적 비난의 형태로 발전할 수 있음을 지적해 왔다(Favarel-Garrigues, Tanner, and Trottier, 2020; Haugh, 2022; Trottier, 2017).

온라인 공간에서 부정적 표현이 강화되는 대표적 요인 가운데 하나는 익명성이다. 익명적 환경에서는 발화자가 실명이나 지속적 정체성과 분리된 상태에서 참여할 수 있기 때문에, 자신의 발언이 이후 관계나 평판에 미칠 영향을 상대적으로 덜 고려하게 된다(최희재·이준기·김정원, 2009; Kim, Ellithorpe, and Burt, 2023; Zimmerman and Ybarra, 2016). 이러한 경향은 익명성이 비난, 조롱, 모욕과 같은 고강도 표현을 보다 쉽게 허용한다고 보는 '온라인 탈억제(online disinhibition)' 논의와도 맞닿아 있다(Christopherson, 2007; Lapidot-Lefler and Barak, 2012; Suler, 2004). 온라인 탈억제 논의는 디지털 환경에서 이용자들이 왜 더 쉽게 적대적 반응을 보이거나 비난과 욕설을 표출하는지를 설명하는 데 유용하다.

그러나 익명성의 효과를 개인의 절제력 약화나 공격성 증가로만 이해하는 것은 충분하지 않다. 온라인 공간은 익명적 발화가 가능하다는 점에서 개인의 책임 부담을 낮추지만, 동시에 반복적 상호작용을 통해 커뮤니티나 플랫폼 고유의 규범과 기대가 형성되는 공간이기도 하다. 따라서 온라인 행태를 분석할 때 중요한 것은 참여자의 반응을 단순한 개인 감정의 표출로만 보는 것이 아니라, 자신이 속한 공간의 규범에 동조하고 특정한 감성 표현을 반복하는 사회적 행위로 이해하는 것이다. 관련 연구는 개인에 대한 식별성이 낮아질수록 참여자가 익명 환경 속 집단 정체성에 더 강하게 귀속되고, 그 결과 집단 규범에 대한 동조가 강화될 수 있음을 보여준다(신지명·김경규·박주연, 2011; Postmes, Spears, Sakhel, and de Groot, 2001; Reicher, Spears, and Postmes, 1995).

본 연구는 온라인 논란의 감성적 확산을 설명하기 위해, 익명성 일반보다 플랫폼이 제공하는 식별가능성의 차이에 주목한다. 많은 온라인 커뮤니티는 이용자의 발화가 얼마나 기록되고, 반복적으로 인식되며, 과거 활동과 연결될 수 있는지를 서로 다르게 설계한다. 이러한 구조는 발화자가 자신의 발언이 이후 관계 속에서 어떻게 기억되고 재소환될 수 있는지, 그리고 그로 인해 어떤 평판 비용을 감수해야 하는지를 판단하는 데 영향을 준다(Gillespie, 2010; Treem and Leonardi, 2013). 식별가능성이 높은 환경에서는 발화가 기록되고 추적될 가능성이 크기 때문에 자기검열이나 완곡화가 나타나기 쉽다. 반면 식별가능성이 낮은 환경에서는 발화의 사회적 비용이 상대적으로 줄어들어, 논란 국면에서 공유되는 감성이 더 직접적이고 강한 형태로 표출될 가능성이 있다(Friedman and Resnick, 2001; Moore et al., 2021). 본 연구는 동일한 논란이라도 참여 방식에 따라 게시물의 감성 톤이 달라질 수 있다는 점에 주목하며, 이를 디시인사이드의 고정닉/유동닉 구조를 통해 검토한다.

3. 디시인사이드의 참여 구조와 식별가능성의 차이

앞선 논의를 종합하면, 온라인 논란 담론의 감성적 확산을 이해하기 위해서는 동일한 사건이라도 어떤 식별가능성 구조 아래에서 발화가 이루어지는지를 관찰할 수 있는 구체적 장이 필요하다. 디시인사이드는 이러한 연구 목적에 부합하는 사례이다. 디시인사이드는 동일한 커뮤니티 내부에 ‘고정닉’과 ‘유동닉’이라는 서로 다른 참여 방식을 유지하고 있으며, 두 방식은 발화자의 식별가능성과 기록의 누적 정도에서 차이를 보인다. 고정닉은 회원 가입 후 로그인을 통해 특정 닉네임을 지속적으로 사용하며 활동 기록이 누적되는 방식이다. 반면 유동닉은 로그인 절차 없이 게시물이나 댓글을 작성하는 방식으로, 갤러리 설정에 따라 ‘ㅇㅇ’ 또는 해당 갤러리에서 사전에 설정한 기본 닉네임으로 표시된다. <Figure 1>은 디시인사이드 갤러리에서 게시글 작성자의 ID가 어떻게 표시되는지를 보여준다. 고정닉의 경우 동일한 닉네임을 통해 작성자의 활동을 반복적으로 인식할 수 있는 반면, 유동닉의 경우 동일한 표기가 여러 이용자에게 공유될 수 있기 때문에 개별 작성자의 발화를 지속적으로 연결해 파악하기 어렵다. 이처럼 동일한 커뮤니티 안에서 서로 다른 수준의 식별가능성이 병존한다는 점은, 디시인사이드를 참여 방식의 차이가 감성 표현에 미치는 영향을 살펴보기에 적절한 장으로 만든다.

〈Figure 1〉 The Example of Fixed/Ephemeral Case

번호	제목	글쓴이	작성일	조회	추천
설문	논란된 스타들 이미지 세탁 그만 해줬으면 하는 프로그램은?	운영자	26/01/26	-	-
이슈	[디시人터뷰] 라이징 스타를 넘어 믿고 보는 배우로, 이효나	운영자	26/01/27	-	-
AD	설 황금연휴 여행 특가	운영자	26/01/23	-	-
공지	■■■■ DC 몬스타엑스 갤러리 가이드 ■■■■ [30]	DC몬엑갤	16.05.08	34405	127
공지	몬스타 엑스 갤러리 이용 안내 [45]	운영자	15.06.02	23217	90
345011	드코준비했긔?	ㅇㅇ (106.101)	10:14	23	0
345010	우깡이랑 아직 사겨? [15]	몬갤러 (118.235)	01.28	655	0
345007	어제 생일 라이브 [5]	몬갤러 (118.235)	01.27	443	4
345006	개동장 그지야 관심을 받고프면 [14]	ㅇㅇ (121.133)	01.27	516	19
345002	아아악! [3]	ㅇㅇ (106.101)	01.26	451	1
345001	소원대로 30곡 [10]	루시콧물 (125.176)	01.26	439	1
345000	엠콜곡 첫곡? [1]	ㅇㅇ (106.101)	01.26	272	0
344999	일등으로 축하해야징 [8]	ㅇㅇ (106.101)	01.26	433	10
344998	쏭그지는 친구생일 축하해줄거면 [5]	몬갤러 (118.235)	01.25	501	13
344997	골디 지금 봤는데 [7]	ㅇㅇ (118.235)	01.25	468	0
344996	보키는 알아서 개동장 얘기 안하는데 [9]	몬갤러 (223.38)	01.24	684	16
344994	아재들 ㄹㅇ 감없네 [4]	ㅇㅇ (211.36)	01.24	638	3
344993	됐고 개동장에서 개팔아먹는 사생녀야 [14]	ㅇㅇ (1.219)	01.23	937	19
344992	입장할때 포카주시긔 [8]	ㅇㅇ (106.101)	01.23	341	3

Source: Dcinside MONSTA X Gallery

디시인사이드의 이러한 구조는 디시인사이드만의 예외적인 현상이라기보다, 온라인 커뮤니티가 개방성과 회원 관리라는 두 요구를 절충하는 과정에서 형성된 제도적 배열로 이해할 수 있다. 다수의 온라인 플랫폼은 낮은 진입장벽을 통해 즉각적 참여를 허용하는 한편, 반복 참여자에 대한 관리와 추적 가능성도 함께 유지한다. 선행연구 역시 이러한 익명성 또는 식별가능성의 차이가 발언권, 상호작용 방식, 커뮤니티 내부의 위계와 규범 형성에 영향을 줄 수 있음을 지적해 왔다(권지현·김명혜, 2014; 오관석, 2009; 황병화·이은규, 2025; Treem and Leonardi, 2013). 특히 유사한 이분화 구조를 다룬 연구들은 익명 수준의 차이가 참여자의 발화 방식과 상호작용 패턴을 달리 만들 수 있음을 보여준다(송준모·강정한, 2018; Friedman and Resnick, 2001).

이러한 논의가 시사하는 바는 온라인에서 관찰되는 감성 표현의 차이를 개인의 고유한 공격성으로 환원하기보다, 플랫폼이 제공하는 식별가능성과 그에 따른 사회적 비용 구조의 차이로 설명할 필요가 있다는 점이다. 디시인사이드에서 고정닉은 반복적 상호작용과 누적된 평판이 존재하므로, 동일한 논란에 대해서도 발화가 이후 관계와 평판에 미칠 영향을 고려할 유인이 상대적으로 크다. 반면 유동닉은 계정의 지속성이 낮고 식별가능성이 약하므로, 발화에 수반되는 평판 손실이나 관계

적 제재를 상대적으로 덜 부담한다. 따라서 동일한 논란이라도 유동닉 게시물에서는 감성 표현이 보다 직접적이고 강한 형태로 나타날 가능성이 있다.

다만 이러한 효과가 모든 논란에서 동일한 크기로 나타난다고 보기는 어렵다. 논란의 성격에 따라 유동닉의 낮은 식별가능성이 더 강하게 작동하는 조건이 있을 수 있기 때문이다. 예컨대 정치사회적 발언처럼 규범 해석이 쉽게 갈리는 사안에서는 사실 여부의 판단보다 그 행위를 둘러싼 해석 경쟁이 먼저 발생하기 쉽다. 이 과정에서 해당 아티스트의 발언이나 행위가 특정 정치적 입장 또는 상대 진영의 표식으로 간주되면, 온라인 커뮤니티에서 빠르게 쟁점화되고 찬반 논쟁이 격화될 가능성이 크다(Park, Lee, Ryu, and Hahn, 2015; Samosir and Wee, 2023). 또한 다국적 구성원을 기반으로 하는 케이팝 그룹의 특성상, 특정 멤버의 한국 문화나 사회에 대한 무지가 한국 사회에 대한 무례함이나 무시, 혹은 차별적 태도로 해석될 가능성도 존재한다(Ahn and Lin, 2019). 이러한 유형의 논란에서는 발화의 책임과 평판 비용이 낮은 유동닉 환경에서 더 직접적인 감성 표현이 나타날 수 있다.

집단 정체성이 전면화되는 논란 유형에서도 식별가능성의 차이는 더 크게 작동할 수 있다. 팬덤 사이의 갈등, 혹은 지지하는 아이돌의 성적표·순위·성과를 둘러싼 정당성 시비처럼 내집단과 외집단의 경계가 비교적 분명한 사안에서는 사실 판단보다 '어느 편에 설 것인가'가 먼저 조직되기 쉽다(Abd-Rahim, 2019). 이 경우 낮은 식별가능성은 개인적 책임 부담을 낮추는 동시에, 집단이 공유하는 규범과 적대적 언어에 대한 동조를 강화할 수 있다(Postmes et al., 2001; Reicher et al., 1995). 특히 유동닉은 반복적 관계와 평판의 부담이 상대적으로 작기 때문에, 집단 경계가 뚜렷한 논란에서 비판이나 조롱, 부정적 감성 표현을 보다 직접적으로 드러낼 가능성이 있다. 이러한 논의를 바탕으로 본 연구는 다음과 같은 가설을 제시한다.

H1. 유동닉 게시물은 고정닉 게시물에 비해 논란 담론에서 더 낮은 감성 점수를 보일 것이다.

H2. 유동닉 게시물의 감성 점수 하락 효과는 규범 해석이나 집단 정체성이 전면화되는 논란 유형에서 더 크게 나타날 것이다.

Ⅲ. 자료와 방법론

1. 분석 자료

본 연구는 참여자의 식별가능성이 발화의 내용과 감성적 표현에 어떠한 차이를 만들어내는지를 검토하고자 한다. 이를 위한 분석의 장으로 디시인사이드를 선택하였다. 디시인사이드는 2023년 기준 도메인 트래픽 순위 4위, 약 2억 3,411만 접속을 기록하였고, 일평균 방문자 수는 3,031,113명, 하루 평균 게시물 수는 882,260건에 이르는 것으로 알려져 있다(조선비즈, 2023). 국내외 포털 사이트를 제외하면 국내에서 가장 큰 규모의 온라인 커뮤니티 가운데 하나이다. 디시인사이드에는 특정 주제나 대상을 중심으로 구성된 '갤러리'라는 하위 커뮤니티가 존재하며, 참여자는 여기서 정보 공유, 잡담, 지지, 비판, 조롱 등 다양한 유형의 발화를 생산한다. 실제로 디시인사이드 갤러리는 이용자들 사이의 소속감과 유대감을 형성하는 공간인 동시에, 신념이나 성향의 차이에 따라 위계 형성과 갈등, 충돌이 빈번하게 나타나는 장으로 기능해 왔다(강옥희, 2023; 박성희·박정윤, 2006; 송경재, 2005; 유지연, 2020; 장민지, 2016).

본 연구에서 다루는 아이돌의 '사회적 논란'은 해당 아이돌이 논란의 가해자이거나, 최소한 논란을 촉발하거나 유발한 당사자로 지목되는 사건으로 한정하였다. 다시 말해, 당사자가 단지 사건의 주변 인물로 언급되는 경우, 외부 환경 요인에 의해 수동적으로 연루된 사례, 또는 당사자가 피해자인 경우는 분석 대상에서 제외하였다. 이처럼 사건의 책임이나 원인이 비교적 분명하게 아이돌에게 귀속되는 사례를 중심으로 범위를 한정한 것은, 비교 가능한 담론 환경을 구성하고 표본의 이질성을 줄이기 위한 것이다.

분석을 위한 자료 수집 범위는 아이돌의 세대 구분법을 제시한 신희연·한혜원의 연구(2025)와 한국국제문화교류진흥원의 기고문(2003)을 참고하여 아이돌 세대를 통제변수로 포함하고 2세대 이후 케이팝 아이돌만을 분석 대상으로 선정했다. 케이팝 세대 구분은 공식적 제도 분류라기보다 데뷔 시기, 미디어 유통 환경, 팬덤 조직 방식, 해외 진출 양상 등을 종합적으로 반영한 산업적·담론적 구분에 가깝다. 본 연구에서 2세대 이후 아이돌을 표집 기준으로 삼은 이유는 이 시기부터 케이팝 아이돌을 둘러싼 대중적 평판이 온라인 커뮤니티와 플랫폼 이용자들의 지속적인 게시, 댓글, 공유, 재해석 과정을 통해 빠르게 확산되고 재구성되기 시작했기 때문이다.

아이돌 관련 논란 역시 이러한 환경 속에서 특정 사건에 대한 정보가 빠르게 순환하고, 이용자들의 평가와 감성적 반응이 온라인 담론으로 축적되며, 다시 언론 보도나 플랫폼 알고리즘을 통해 증폭되는 양상을 보인다(Nakamura, Nishimoto, Sakata, and Asatani, 2025). 따라서 2세대 이후 아이돌은 온라인 커뮤니티 기반 여론 형성과 감성적 평가의 구조를 분석하려는 본 연구의 문제의식에 적합한 대상이라고 할 수 있다.

구체적으로 케이팝 세대론에서 논의되는 미디어 유통 환경, 팬덤 조직 방식, 해외 진출 양상 등을 함께 고려하여 2세대, 2.5세대, 3세대, 4세대로 구분하였다. 2.5세대는 2세대와 3세대 사이의 과도기적 집단으로, 2010년대 초반을 전후하여 2세대의 제작·팬덤 구조를 공유하면서도 SNS와 온라인 커뮤니티 기반 팬덤 확산이 본격화된 사례를 포함한다. 4세대는 2018년 이후 글로벌 플랫폼과 알고리즘 기반 확산, 실시간 소통, 초국적 팬덤 동시성이 더욱 강화된 시기의 집단으로 정의하였다. 본 연구는 이러한 기준에 따라 전체 케이팝 아이돌 관련 73개 갤러리 가운데 2세대 이후 아이돌과 관련된 57개 갤러리를 최종 분석 대상으로 선정하였다(부록의 <Appendix 1> 참조).

선정된 갤러리를 바탕으로 각 아이돌 개인 또는 그룹과 관련된 논란의 발생일과 유형을 조사한 뒤, 전체 사건을 12개 유형으로 분류하였다. 또한 갤러리마다 차이는 있었으나, 대체로 논란이 발생한 시점부터 약 15일 동안 관련 이슈에 대한 언급이 지속되는 경향을 확인하고 해당 기간을 게시물 수집 범위로 설정하였다. 게시물 수집은 Python의 BeautifulSoup와 Selenium 라이브러리를 활용하여 제작한 웹 크롤러를 통해 이루어졌으며, 57개 갤러리에서 총 714,778건의 텍스트를 확보하였다. 동시에 이후 다중회귀모형에서 아이돌 수준의 변수를 통제하기 위해 성별, 활동 유형, 세대 구분, 소속사 상장 여부 등 기본 정보를 별도로 수집하여 보조 데이터베이스를 구축하였다.

2. 딥러닝 기반 텍스트 분류 및 감성분석

앞서 구축한 코퍼스는 논란 발생 시점을 중심으로 일정 기간 게시물을 포괄적으로 수집한 자료이므로, 잡담, 팬 활동 기록, 일반 정보 공유 등 논란과 직접 관련이 없는 글이 함께 포함되어 있다. 이러한 게시물의 혼재는 이후 분석에서 종속변수와 핵심 설명변수 사이의 관계를 약화시킬 수 있으므로, 본 연구는 분석 단위를 ‘논란 사건을 실질적으로 다룬 게시물’로 정의하고 이를 식별하기 위한 분류 절차를 도입

〈Table 1〉 The Variable Lists

Panel A.	mean	std	min	max
Sex(d) (male=0, female=1)	0.3	0.46	0	1
Foreigner (%)	0.04	0.13	0	1
Type(d) (solo=0, group=1)	0.72	0.45	0	1
Stock market(d) (listed=1, not-listed=0)	0.38	0.49	0	1
Anonymous(d) (fixed=0, ephemeral=1)	0.7	0.4	0	1
Panel B. Idol Generation	n	%		
2nd Generation	15,505	12.68		
2.5th Generation	25,016	20.46		
3rd Generation	59,925	49.01		
4th Generation	21,831	17.85		

하였다. 57개 갤러리에서 수집한 714,778건의 게시물을 모두 사람이 직접 분류하는 것은 현실적으로 어렵고, 은어·초성·비속어·표기 변형이 빈번한 커뮤니티 텍스트를 일관되게 판정하는 데에도 한계가 있기 때문에 딥러닝 기반 자동 분류 기법을 적용하였다.

자동 분류 모델로는 KcELECTRA를 사용하였다(신미르·신유현, 2024; 이영준·박하명·강승식, 2023; Clark, Luong, Le, and Manning, 2020; Yoo, Lee, Song, and Jeong, 2023). 본 연구가 KcELECTRA 기반 지도학습 분류기를 사용한 이유는 논란 관련/비관련 판별이 개방형 생성 과제가 아니라, 대규모 코퍼스에 반복 적용되는 폐쇄형 이진 분류 과제이기 때문이다. 이러한 과제에서는 한국어 온라인 댓글 말뭉치에 기반해 학습된 분류 모델이 예측의 일관성과 계산 효율성 측면에서 적합하며, 축약·비문·속어·표기 변형이 빈번한 커뮤니티 텍스트를 안정적으로 처리하는 데에도 유리하다고 판단하였다.

분류기를 전체 코퍼스에 적용하기에 앞서, 본 연구는 디시인사이드에서 무작위로 추출한 텍스트 데이터 1,021개에 대해 7명의 코더가 논란 관련/비관련 여부를 판정한 휴먼코딩 자료를 구축하였다. 전체 표본은 층화 분할을 통해 훈련 세트 70%(715개), 검증 세트 15%(153개), 평가 세트 15%(153개)로 나누었다. 휴먼코딩 기준 논란 관련 게시물은 285개였다. 훈련 세트는 모델 학습에, 검증 세트는 학습 과정에서의 성능 점검과 최적 모델 선택에, 평가 세트는 최종 일반화 성능의 독립적 검증에 활용하였다. 평가 세트 기준 정확도는 0.778, macro F1은 0.732, weighted F1은 0.781로 나타났으며, 논란 관련 클래스의 precision은 0.596, recall은 0.651, F1은 0.622였다.[1] 이후 학습된 분류기를 전체 714,778건의 게시물에 적용한 결과, 최종

적으로 122,277건(17.4%)이 논란 관련 게시물로 식별되었다. 디시인사이드 갤러리에서는 논란 발생과 무관한 다양한 게시물이 지속적으로 생산되므로, 논란 관련 게시물이 전체 게시물 가운데 상대적으로 적은 비중을 차지하는 것은 자연스러운 결과로 볼 수 있다.

다음으로 논란 게시물의 감성적 특성을 계량화하기 위해 감성분석을 수행하고, 그 결과를 종속변수로 활용하였다. 본 연구가 분석하는 온라인 논란 담론은 단순한 긍·부정 이분법으로 환원되기보다 분노, 혐오, 실망, 불안, 경멸 등 복수의 감정이 중첩되어 나타나는 경우가 많다. 이에 따라 본 연구는 한국어 온라인 댓글 환경에 맞추어 구축된 KOTE(Korean Online That-gul Emotions) 데이터셋과 모델을 활용하였다(Jeon et al., 2022). KOTE는 한국어 온라인 댓글 50,000개를 뉴스, 커뮤니티, 소셜미디어, 이커머스 등 12개 플랫폼에서 수집하고, 43개 감정 범주와 무감정을 포함한 총 44개 범주에 대해 다중라벨을 부여한 자료이다. 또한 한 문서에 복수의 감정이 공존할 수 있다는 점을 반영하여, 각 텍스트에 대해 개별 감정 범주의 등장 확률을 산출한다. 이러한 특성은 온라인 논란 게시물에서 나타나는 세부 감정의 차이를 단일 극성 분류보다 더 세밀하게 포착할 수 있다는 장점을 가진다.

본 연구의 관심은 개별 감정 범주의 정밀한 판별 자체가 아니라, 논란 관련 게시물이 전반적으로 어느 방향의 감성을 띠는지를 비교하는 데 있다. 이에 따라 본 연구는 각 게시물에 대해 44개 감정 범주의 예측 확률 p_k를 산출한 뒤, 기본적으로 $p_k \geq 0.4$인 범주만을 활성 감정으로 간주하였다. 이후 KOTE 연구에서 구분한 감정 분류표에 따라 세부 감정을 긍정·부정·중립 정서 그룹으로 재분류하고(부록의 <Appendix 2> 참조), 부정 감성에는 음의 부호를, 긍정 감성에는 양의 부호를 부여한 다음, 이들 값의 평균치를 게시물 단위 감성 점수로 산출하였다. 이렇게 구성한 감성 점수를 종속변수로 설정하고, 아이돌 특성, 논란 유형, 게시물 작성자의 참여 방식(고정닉/유동닉)을 포함하는 회귀모형을 통해 논란 담론의 감성 방향과 강도의 차이를 검토하였다.

추가적으로 이러한 점수 구성이 임계치나 산식에 민감할 가능성을 점검하기 위해 강건성 검정을 수행하였다. 먼저 활성 감정의 기준을 $p_k \geq 0.5$로 상향한 대안 점

1) 논란 관련 클래스의 precision은 0.596, recall은 0.651로 나타났다. 이는 논란 관련으로 분류된 게시물에 일정한 위양성 사례가 포함될 수 있고, 실제 논란 관련 게시물 중 일부가 분석 표본에서 제외되었을 가능성을 의미한다. 다만 이러한 오분류는 논란과 직접 관련이 약한 게시물을 분석 표본에 포함하거나 일부 관련 게시물을 누락시킴으로써, 본 연구가 관심을 두는 감성 점수 차이를 과대평가하기보다 희석시키는 방향으로 작용했을 가능성이 크다.

수를 구성하여 동일한 분석을 반복하였다. 또한 임계치를 적용하지 않고 극성별 확률을 합산하여 구성한 감성 점수에 대해서도 동일한 분석을 수행하였다. 나아가 KOTE 기반 감성 점수의 외적 강건성을 점검하기 위해 별도의 두 가지 감성분석 모델을 추가로 활용하였다. 두 외부 모델의 출력값 역시 확률가중평균 방식으로 게시물 단위 감성 점수로 재구성한 뒤 동일한 회귀모형을 추정하였으며, 결과는 부록 <Appendix 5>에 제시하였다.

Ⅳ. 분석 결과

1. 논란 유형별 감성 점수와 세부 감정 구성

먼저 <Table 2>는 저자들이 분류한 12개 논란 유형별로 고정닉과 유동닉 게시물의 평균 감성 점수 및 평균차에 대한 t-test 결과를 제시한다. 이를 통해 각 논란 유형이 전반적으로 어떠한 감성적 방향을 보이는지, 그리고 동일한 논란 유형 안에서도 참여 방식에 따라 평균 감성 점수의 차이가 통계적으로 유의하게 나타나는지를 확인할 수 있다.

분석 결과, 논란 유형에 따라 평균 감성 점수의 수준은 상당한 차이를 보였다. 먼저 군 복무는 고정닉과 유동닉 모두에서 전반적으로 낮은 감성 점수를 나타냈다. 군 복무 범주의 평균 감성 점수는 고정닉 -0.388, 유동닉 -0.464로 모두 낮았으며, 두 집단 간 차이도 통계적으로 유의하였다(평균차=-0.077, t=3.484). 다만 이 차이는 다른 유형과 비교할 때 상대적으로 크지 않았다. 이는 병역 의무가 한국 사회에서 오랫동안 공정성과 책임의 문제로 강하게 규범화되어 왔으며, 위반 또는 회피 의혹이 제기될 경우 고정닉과 유동닉 모두에서 비판적 반응이 비교적 쉽게 형성되는 범주이기 때문으로 볼 수 있다. 다시 말해 군 복무는 식별가능성의 차이가 전혀 작동하지 않는 범주라기보다, 이미 비판의 방향이 강하게 공유된 상태에서 참여 방식에 따른 추가적 차이가 제한적으로 나타나는 범주로 해석할 수 있다.

반면 마약 및 불법 도박을 포함하는 위법행위 범주는 법적 중대성에 비해 평균 감성 점수가 상대적으로 낮지 않았고, 식별가능성에 따른 차이도 유의하지 않았다. 이는 사건의 법적 심각성이 온라인 공간에서의 감성 톤과 항상 비례하는 것은 아님을 보여준다. 특히 온라인 커뮤니티에서는 어떤 사건이 즉각적인 분노와 비판의 대

〈Table 2〉 Statistics of Sentiment Analysis Depending on Controversies

Type	Fixed Accounts (N)	Ephemeral Accounts (N)	Mean (Fixed)	Mean (Ephemeral)	Mean Diff. (Ephemeral -Fixed)	t-value
Fandom Conflict (팬덤 갈등)	730	5131	-0.125	-0.495	-0.370***	15.859
Political/ Social Statements (정치사회적 발언)	800	5435	-0.207	-0.458	-0.251***	12.144
Attitude (태도 논란)	5001	11231	-0.14	-0.376	-0.235***	24.604
Agency/ Member Conflict (소속사/멤버 불화)	2906	25227	-0.159	-0.391	-0.232***	20.168
Intellectual Property (지식재산권)	3002	7421	-0.142	-0.358	-0.217***	17.543
School Violence/ Assault (학교폭력/폭행)	1788	1185	-0.055	-0.272	-0.216***	10.083
Dating/Romance (연애/이성 문제)	5178	23991	-0.17	-0.327	-0.158***	18.345
Sexual Misconduct (성 관련 사건)	1904	6363	-0.208	-0.306	-0.098***	6.44
Military Service (군 복무)	592	8441	-0.388	-0.464	-0.077***	3.484
Traffic Offenses (도로교통 위반)	1719	2102	-0.362	-0.423	-0.061***	3.899
Public Order (공공질서 위반)	308	1043	-0.355	-0.345	0.01	-0.285
Drugs/Gambling (위법행위)	362	367	-0.128	-0.067	0.062	-1.399

상으로 수렴하기보다 희화화나 밈적 소비의 형식으로 재맥락화될 수 있다(Phillips, 2015; Sykes and Matza, 2017). 따라서 특정 사건이 얼마나 중대한가보다, 그것이 커뮤니티 내부에서 어떤 방식으로 해석되고 소비되는가가 실제 게시물의 감성 방향을 가르는 중요한 조건이 될 수 있다.

고정닉과 유동닉 간 평균 차이가 상대적으로 크게 나타난 범주는 팬덤 갈등, 정치사회적 발언, 태도 논란, 소속사/멤버 불화, 지식재산권, 학교폭력/폭행 등이었다. 이들 범주에서는 모두 고정닉보다 유동닉의 감성 점수가 유의하게 낮게 나타났다. 특히 팬덤 갈등, 정치사회적 발언, 소속사/멤버 불화는 사건의 사실관계만이 아니라 가치 판단이나 집단 정체성의 대립이 함께 전면화되기 쉽다는 점에서 식별가능성에

따른 차이를 해석하는 데 중요하다. 정치사회적 발언은 특정 입장 표명이 곧바로 규범적 정당성의 문제로 전환되기 쉽고, 팬덤 갈등은 타 팬덤과의 경쟁이나 경계 짓기가 직접적으로 표출되는 영역이며, 소속사/멤버 불화는 팬들이 기대해 온 팀 정체성이나 공동체 서사를 흔드는 사건으로 받아들여질 수 있다. 이처럼 논란 유형에 따라 사건의 사회적 의미가 다르게 구성되기 때문에, 참여 방식의 차이 역시 모든 범주에서 동일하게 나타나지 않고 특정 유형에서 더 크게 드러날 수 있다.

<Table 2>에서 확인되는 감성 점수의 차이는 실제 게시물에서는 발화의 어조, 비판의 직접성, 집단 경계의 구성 방식의 차이로 나타난다. 특히 정치사회적 발언, 팬덤 갈등, 소속사/멤버 불화는 고정닉과 유동닉의 평균 차이가 상대적으로 크게 나타난 범주로, 아래 사례는 이러한 차이가 실제 담론에서 어떠한 방식으로 드러나는지를 보여준다. 먼저 정치사회적 발언은 특정 발언이나 이슈가 곧바로 규범적 정당성의 문제로 전환되기 쉬운 범주이다. 이 경우 고정닉은 발화가 누적되고 추적될 가능성을 고려하여 표현을 조절할 유인이 상대적으로 큰 반면, 유동닉은 낮은 평판 비용 아래에서 보다 직접적이고 단정적인 평가를 표출할 수 있다(Friedman and Resnick, 2001; Reicher et al., 1995). 실제 사례에서도 유동닉 게시물은 대상에 대한 평가를 짧고 직접적인 어조로 제시하고 있다.

“안중근 모른다치고 김두한을 긴또깡으로 부른거는 ㅋㅋ □□ 이거생각하면 ㅋㅋ”	정치사회적 발언, AOA, 유동닉, 감성점수: -0.23

소속사/멤버 불화 역시 평균 감성 점수의 차이가 실제 담론에서 어떻게 나타나는지를 보여주는 범주이다. 케이팝 아이돌은 팀워크와 공동체적 정체성을 중시하는 존재로 기대되기 때문에, 멤버 불화는 팬들이 기대해 온 집단적 가치와 서사를 훼손하는 사건으로 받아들여질 수 있다(김수정·김수아, 2015). 이 범주에서 유동닉 게시물은 특정 인물에 대한 적대감과 책임 귀속을 강한 감정 표현과 결합해 드러내는 반면, 고정닉 게시물은 사태의 향방이나 관계 변화, 혹은 향후 집단 내부의 정렬을 가늠하는 방식으로 발화하는 경향을 보였다. 이는 동일한 사건을 둘러싼 반응이라도 참여 방식에 따라 담론의 감성적 결이 달라질 수 있음을 시사한다.

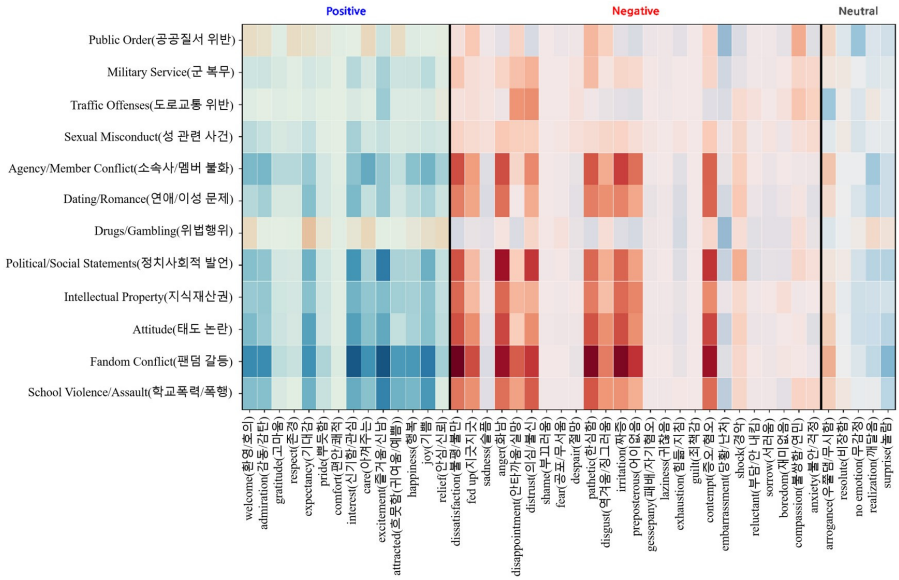
“륵한 □□새끼 때문에 홧병이 가시질 않는다 좀 잊을만하면 다시 떠올라서 또 빡침 이 □□새끼 인생 조지는 꼴은 봐야되는데”	불화, EXO, 유동닉, 감성 점수: -0.86
“5명이 끝까지 함께하겠지 다만 데스피 잔류냐 소속사 이적이냐의 차이일뿐 명 부모들 분위기는 양보안할거 같다”	불화, 카라, 고정닉, 감성 점수: -0.08

이처럼 정치사회적 발언에서는 규범적 판단의 직접성이, 팬덤 갈등에서는 집단 경계의 전면화가, 소속사/멤버 불화에서는 책임 귀속과 적대감의 표출이 상대적으로 두드러지게 나타났다. 다시 말해 <Table 2>에서 나타나는 차이는 단순한 수치상의 차이에 그치지 않고, 동일한 논란을 둘러싼 평가와 비판이 식별가능성에 따라 서로 다른 감성적 형식으로 조직될 수 있음을 보여준다.

<Figure 2>는 이러한 평균 감성 점수의 차이가 세부 감정 범주의 구성에서 어떻게 나타나는지를 보여준다. KOTE는 각 게시물에 대해 복수의 감정 범주가 동시에 존재할 수 있음을 전제로 감정별 확률을 산출하므로, <Figure 2>는 고정닉과 유동닉 게시물의 차이가 단순히 점수의 높고 낮음에 그치지 않고 어떤 감정 라벨의 상대적 분포 차이로 구성되는지를 확인하게 해준다. 분석 결과, 유동닉 게시물에서는 *불평/불만, 의심/불신, 화남/분노, 증오/혐오, 짜증, 한심함*과 같은 부정적 감정 라벨이 상대적으로 높게 나타나는 경향이 확인되었다. 이는 유동닉 게시물이 고정닉 게시물보다 평균적으로 더 낮은 감성 점수를 보인다는 <Table 2>의 결과와 방향을 같이한다.

다만 <Figure 2>의 결과는 유동닉 게시물이 항상 단일한 방식으로 “부정적”이라는 뜻이라기보다, 논란 유형에 따라 감정 구성의 차이가 다르게 나타난다는 점을 보여준다. 정치사회적 발언에서는 규범적 판단과 입장 표명이 전면화되면서 유동닉 게시물에서 불신, 분노, 혐오와 같은 감정 라벨이 상대적으로 강하게 나타났다. 이는 정치사회적 발언이 사실 판단을 넘어 발화자의 태도나 정체성에 대한 평가로 전환되기 쉽고, 낮은 식별가능성 아래에서 이러한 평가가 보다 직접적으로 표출될 수 있음을 시사한다. 팬덤 갈등에서도 유동닉 게시물은 상대 집단을 향한 조롱이나 경계 짓기와 결합된 부정적 감정 라벨을 더 많이 보였다. 이는 팬덤 갈등이 단순한 사건 평가라기보다 내집단과 외집단의 위치를 확인하고 경쟁하는 담론으로 구성되기 쉽기 때문이다.

〈Figure 2〉 Heatmap of Differences in Sentiment Distributions Across Controversy Categories(Fixed vs. Ephemeral)



한편 팬덤 갈등에서 고정닉 게시물이 *기대감, 환영/호의*와 같은 상대적으로 긍정적인 감정 라벨을 더 많이 보인 점은 갈등 자체에 대한 긍정적 평가로 해석하기보다, 내집단 방어와 결속의 표현으로 이해할 필요가 있다. 고정닉은 갤러리 내부에서 반복적으로 등장하며 관계망과 평판을 축적한 참여자이므로, 갈등 국면에서도 자신의 집단을 지지하거나 상황 반전을 기대하는 방식으로 감정을 표출할 가능성이 있다. 따라서 <Figure 2>는 동일한 논란 상황에서도 유동닉은 직접적 비판과 부정적 감정 표출에, 고정닉은 내부 결속과 관계적 정렬에 더 가까운 감정 구성을 보일 수 있음을 보여준다. 이는 본 연구가 제시한 식별가능성의 차이가 단순히 감성 점수의 평균 차이만이 아니라, 논란 담론이 어떤 감정적 구성으로 조직되는가와도 연결되어 있음을 시사한다.

2. 온라인 식별가능성에 따른 감성 점수의 효과

<Table 3>은 게시물의 감성 점수를 종속변수로 하여 참여 방식(유동닉/고정닉), 즉 식별가능성 수준과 아이돌 단위의 통제변수를 포함한 다중회귀분석을 수행한 결과이다. 본 연구는 (1) 아이돌 특성과 참여 방식을 포함한 기본 모형, (2) 논란 유형

고정효과를 추가한 모형, (3) 참여 방식과 논란 유형의 상호작용항을 포함한 모형, (4) 아이돌 고정효과를 포함한 모형을 순차적으로 추정하였다. 특히 최종 모형인 Model 4는 갤러리 분위기, 팬덤 문화, 고정된 이미지와 인지도처럼 아이돌별로 관측하기 어려운 이질성을 통제함으로써, 동일 아이돌 내부에서 나타나는 감성 점수의 차이를 보다 보수적으로 추정한 것이다. 또한 감성 점수 구성 방식에 따른 민감도를 점검하기 위해, 임계치를 0.5로 설정한 대안 점수와 임계치를 적용하지 않고 각 극성별 확률을 합산해 구성한 대안 점수에 대해서도 동일한 분석을 수행하였으며, 그 결과는 부록 <Appendix 3>와 <Appendix 4>에 제시하였다.

〈Table 3〉 Regression of Sentiment Scores in Controversy-Related Discussions

	(1)		(2)		(3)		(4)	
	Coef	S.E	Coef	S.E	Coef	S.E	Coef	S.E
Sex(d)	0.002	0.034	0.024	0.039	0.022	0.037	0.004	0.013
Type(d)	-0.080*	0.035	-0.056	0.039	-0.063	0.036	-0.122***	0.008
Foreigner(d)	0.111	0.106	0.124	0.121	0.117	0.123	-0.068*	0.031
Stock market(d)	0.105***	0.024	0.084**	0.027	0.090**	0.028	0.061***	0.010
Generation(ref. 2nd)								
2.5th generation	0.059	0.035	0.124*	0.049	0.120**	0.046	0.011	0.016
3rd generation	-0.056	0.045	-0.022	0.055	-0.023	0.053	-0.053***	0.014
4th generation	-0.054	0.047	-0.036	0.053	-0.040	0.053	0.017	0.013
Anonymous(d)	-0.184***	0.028	-0.162***	0.028	-0.133**	0.047	-0.127*	0.054
Controversy FE		No		Yes		Yes		Yes
anonym×Controversy		No		No		Yes		Yes
Artist FE		No		No		No		Yes
Constant		-0.152***		-0.173***		-0.192***		0.046
Log likelihood		-95884		-95247		-95090		-93360
Adjusted R^2		0.03		0.04		0.05		0.07

N=122,277, (d): dummy variable

Note: *p<.01, **p<.05, ***p<.001; Models 1 and 2 estimate the average association between anonymity and sentiment, whereas Models 3 and 4 examine whether this association varies across controversy categories through interaction terms.

분석 결과, 모든 모형에서 참여 방식 변수(Anonymous)는 음(-)의 계수를 보였고 통계적으로 유의하였다. 이는 전반적으로 유동닉 게시물이 고정닉 게시물보다 더 낮은 감성 점수와 관련되어 있음을 의미한다. 다만 Model 1과 Model 2의 계수는 논란 유형과의 상호작용을 포함하지 않은 모형에서 추정된 값이므로, 유동닉과 감성 점수 사이의 평균적 관계를 보여주는 것으로 해석할 수 있다. 반면 Model 3과

Model 4는 참여 방식과 논란 유형의 상호작용항을 포함하므로, 이들 모형에서 Anonymous의 단일 계수는 기준 범주에서의 효과를 의미한다. 따라서 상호작용 모형의 실질적 해석은 논란 유형별 총효과를 제시한 <Figure 3>을 중심으로 이루어질 필요가 있다.

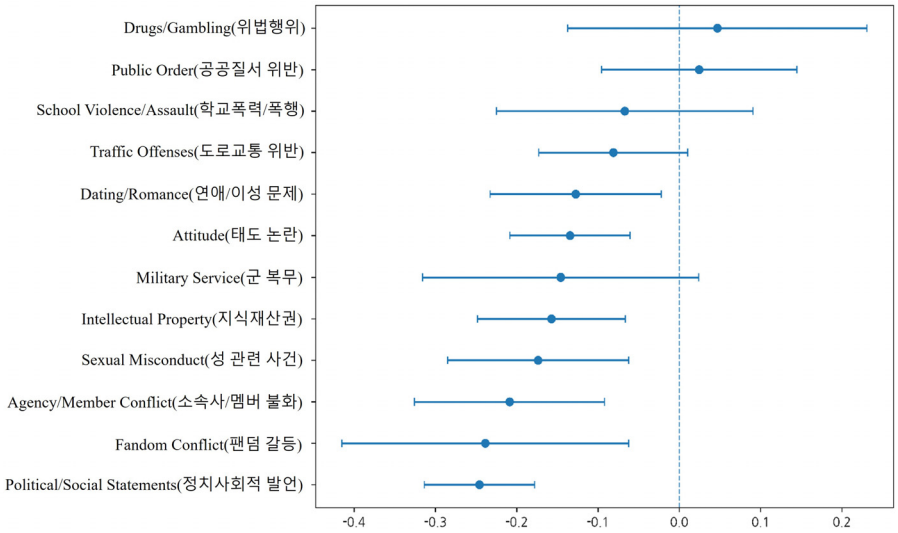
통제변수와의 관계를 살펴보면, 성별은 전반적으로 유의하지 않았고, 솔로/그룹 여부, 외국인 비율, 상장 여부, 세대 변수는 일부 모형에서만 유의하게 나타났다. 이는 아이돌의 기본 특성이 감성 점수와 일정한 관련을 가질 수 있음을 시사하지만, 그 효과가 일관적이지는 않았음을 보여준다. 특히 세대 변수는 모형에 따라 방향과 유의성이 달라졌는데, 이는 세대 자체의 고정된 효과라기보다 각 세대에서 어떤 사건이 대표적으로 논쟁화되었는지가 감성 점수의 편차에 더 중요한 영향을 미쳤을 가능성을 시사한다. 모형의 adjusted R^2는 Model 1의 0.03에서 Model 4의 0.07로 상승하였다. 이는 추가 변수와 고정효과를 포함한 모형이 감성 점수의 일부 변동을 더 포착하고 있음을 보여주지만, 본 연구의 해석은 설명력의 크기 자체보다 유동닉 계수의 방향성과 논란 유형별 조건부 효과에 초점을 둔다.

그렇다면 논란 유형에 따라 유동닉 효과의 실질적 크기와 방향은 어떻게 달라지는가? <Figure 3>은 아이돌 고정효과를 포함한 Model 4를 기준으로, 논란 유형별 유동닉 효과를 95% 신뢰구간과 함께 시각화한 결과이다. 즉, 상호작용 모형의 추정치를 바탕으로 각 논란 유형에서 유동닉 게시물이 고정닉 게시물에 비해 감성 점수에서 어떤 차이를 보이는지를 재구성한 것이다.

분석 결과, 유동닉 효과는 논란 유형마다 상이한 패턴을 보였다. 특히 정치사회적 발언, 팬덤 갈등, 소속사/멤버 불화에서 유동닉 게시물은 고정닉 게시물보다 더 낮은 감성 점수를 보였으며, 그 차이도 상대적으로 크게 나타났다. 여러 범주 가운데서도 정치사회적 발언은 -0.246, 팬덤 갈등은 -0.239, 소속사/멤버 불화는 -0.209로 나타나, 이 세 범주에서 유동닉 효과가 가장 크게 관찰되었다. 반면 군 복무, 도로교통, 학교폭력/폭행, 공공질서 위반, 위법행위 범주는 효과의 방향은 음(-)으로 나타났지만 통계적으로 유의하지 않았다. 이는 유동닉 효과가 일률적인 것이 아니라, 논란 유형에 따라 크기와 유의성이 달라지는 조건부 패턴을 보인다는 점을 뜻한다.

이 결과는 앞서 <Table 2>에서 확인한 유형별 평균 차이와 대체로 같은 방향을 보이지만, 동시에 중요한 차이를 드러낸다. <Table 2>는 통제변수를 포함하지 않은 기초적 평균 비교로서 팬덤 갈등, 정치사회적 발언, 소속사/멤버 불화 등에서 고정

〈Figure 3〉 Ephemeral Effects by Controversy Type(Based on Model 4; 95% CIs)



닉과 유동닉 간 차이가 크게 나타남을 보여주었다. <Figure 3>은 여기에 아이돌 특성, 논란 유형, 아이돌 고정효과를 통제한 이후에도 정치사회적 발언, 팬덤 갈등, 소속사/멤버 불화에서 유동닉 효과가 상대적으로 크게 유지된다는 점을 보여준다. 반면 <Table 2>에서 비교적 큰 평균 차이를 보였던 일부 유형은 회귀모형에서 통계적 유의성이 약화되었는데, 이는 단순 평균 차이의 일부가 아이돌별 이질성이나 사건 분포의 차이와 관련되어 있었을 가능성을 시사한다. 따라서 <Figure 3>의 결과는 단순한 기술통계 차이를 넘어, 동일 플랫폼 내부에서 식별가능성의 차이가 어떤 논란 유형에서 더 안정적으로 작동하는지를 보여준다는 점에서 의미가 있다.

이러한 결과는 식별가능성이 낮은 참여 방식의 효과가 모든 논란에서 동일하게 작동하기보다, 정치화되거나 집단화되기 쉬운 갈등에서 상대적으로 크게 나타날 수 있음을 시사한다. 정치사회적 발언은 규범 해석이 분기되고 찬반의 경계가 뚜렷하게 형성되기 쉬운 범주이며, 팬덤 갈등과 소속사/멤버 불화는 집단 정체성과 결속 규범이 전면화되기 쉬운 범주이다(Friedman and Resnick, 2001; Gillespie, 2010). 이러한 사건에서는 고정닉 참여자가 발화의 누적성과 추적 가능성을 고려하여 표현을 조절할 유인이 상대적으로 큰 반면, 유동닉 참여자는 낮은 평판 비용 아래에서 직접적인 감성 표현을 보일 가능성이 있다. 반대로 군 복무처럼 이미 비판의 규범적 방향이 비교적 강하게 공유된 범주에서는 고정닉과 유동닉 모두 유사한 감성 톤으로 수렴하

여, 식별가능성의 추가 효과가 상대적으로 작게 나타났을 가능성이 있다.

요컨대 <Figure 3>의 결과는 유동닉 효과를 익명성의 자동적 결과로 이해하기보다, 갈등의 정치화 또는 집단화 정도와 결합할 때 그 크기가 더 커질 수 있는 조건부 효과로 해석할 필요가 있음을 보여준다. 이는 온라인 논란에 대한 반응이 개인의 성향이나 사건의 속성만으로 설명되기보다, 동일 플랫폼 내부에서 제도화된 식별가능성의 차이와 결합하여 서로 다른 감성 톤으로 조직될 수 있음을 시사한다(Cho and Kwon, 2015; Haugh, 2022; Trottier, 2017).

V. 결론 및 함의

본 연구는 케이팝 아이돌을 둘러싼 사회적 논란이 온라인 커뮤니티에서 어떠한 감성적 형식으로 조직되는지를, 플랫폼이 제공하는 식별가능성의 차이에 주목하여 분석하였다. 한국 사회에서 아이돌은 높은 규범적 기대와 공적 평가의 대상이 되기 쉽고, 이들을 둘러싼 논란은 온라인 공간을 경유하며 빠르게 확산된다(김수정·김수아, 2015; 오세정, 2012; Kang and Park, 2022; Lee and Abidin, 2024). 본 연구는 이러한 현상을 사건 자체의 성격이나 개별 이용자의 성향만으로 설명하기보다, 동일한 논란이라도 온라인 공간에서의 참여 방식이 다를 때 게시물의 감성 톤이 어떻게 달라지는지에 주목하였다. 이를 위해 디시인사이드의 고정닉/유동닉 구분을 활용하여, 참여 방식의 차이가 논란 담론의 감성 점수와 어떠한 관련을 갖는지를 분석하였다.

분석 결과, 여러 통제변수와 아이돌 고정효과를 단계적으로 적용한 모형에서도 유동닉으로 작성된 게시물은 고정닉 게시물에 비해 유의하게 낮은 감성 점수를 보였다. 이는 온라인 공간에서 관찰되는 감성적 반응이 사건 자체의 내용뿐 아니라, 플랫폼이 설계한 식별가능성의 수준과 그에 따라 달라지는 평판 비용 구조와도 체계적으로 관련될 가능성이 있다는 점을 보여준다(Cho and Kwon, 2015; Friedman and Resnick, 2001; Moore et al., 2021; Treem and Leonardi, 2013). 본 연구는 익명성을 개인의 일탈적 공격성으로 환원하지 않고, 사회적 상호작용의 비용과 책임을 재구성하는 플랫폼의 참여 조건으로 개념화함으로써, 온라인 논란 담론의 감성적 분화가 어떠한 제도적 환경 속에서 나타나는지를 설명하는 경험적 근거를 제시하였다.

또한 본 연구는 유동닉 효과가 모든 논란 유형에서 동일하게 나타나지 않음을 확인하였다. 정치사회적 발언, 팬덤 갈등, 소속사/멤버 불화와 같은 범주에서는 유동닉 게시물이 더 낮은 감성 점수를 보였고, 그 차이도 상대적으로 크게 나타났다. 반면 군 복무처럼 비판의 방향이 비교적 강하게 공유된 범주에서는 고정닉과 유동닉이 유사한 감성 톤으로 수렴하여, 식별가능성의 추가 효과가 상대적으로 제한되었다. 이러한 결과는 온라인 논란의 감성적 반응이 사건의 객관적 심각성에 일률적으로 비례하기보다, 논란이 정치화되거나 집단화되기 쉬운 방식으로 구성될 때 식별가능성의 효과가 더 크게 나타날 수 있음을 시사한다(Haugh, 2022; Postmes et al., 2001; Reicher et al., 1995; Trottier, 2017). 즉, 유동닉 효과는 익명성의 자동적 결과라기보다, 갈등의 성격과 결합하여 확대되거나 약화되는 조건부 효과로 이해할 필요가 있다.

이러한 결과는 최근의 케이팝 팬덤 문화와 온라인 논란 담론을 이해하는 데 중요한 함의를 제공한다. 본 연구는 논란에 대한 부정적 감성 표현을 예외적 일탈이나 특정 소수 집단의 문제로만 보지 않는다. 오히려 동일한 플랫폼 내부에서도 참여 방식과 식별가능성의 차이에 따라 감성 표현의 수위와 방향이 달라질 수 있음을 보여주었다. 이는 케이팝 논란을 설명할 때 아이돌의 특성이나 사건 자체뿐 아니라, 플랫폼 환경과 감성적 상호작용의 구조를 함께 고려할 필요가 있음을 뜻한다(Abd-Rahim, 2019; Duffett, 2013; King-O'Riain, 2021). 나아가 플랫폼은 단순한 의견 표현의 장을 넘어, 감성 표현의 강도와 방향을 매개하는 구조적 환경이라는 점에서 이론적 의미를 갖는다(Gillespie, 2010; 2018). 특히 본 연구는 한국 사회에서 아이돌을 둘러싼 도덕적 기대와 공적 평가가 온라인 담론 속에서 어떠한 감성적 형태로 표면화되는지를 경험적으로 보여준다는 점에서 의의를 갖는다.

다만 본 연구의 결과를 해석할 때에는 부정적 감성 점수의 의미를 신중하게 이해할 필요가 있다. 논란에 대한 부정적 감성은 반드시 사회적으로 해로운 표현을 뜻하지 않는다. 논란의 원인을 제공한 당사자에 대한 비판, 책임 요구, 실망의 표현은 온라인 공론장에서 정당한 평가의 일부일 수 있다. 본 연구가 포착한 것은 특정 게시물이 정당한 비판인지, 과도한 조롱인지, 혹은 혐오적 표현인지를 개별적으로 판정한 결과가 아니라, 참여 방식에 따라 게시물의 전반적 감성 방향과 강도가 어떻게 달라지는가이다. 따라서 본 연구의 결과는 부정적 감성 자체를 문제시하기보다, 식별가능성이 낮은 조건에서 더 직접적이고 강한 감성 표현이 조직될 가능성을 보여주는 것으로 이해하는 것이 적절하다. 다만 이러한 감성적 조건이 반복적 조롱, 표적화,

혐오적 표현과 결합할 경우 온라인 갈등이 심화될 가능성이 있다는 점은 후속 연구와 플랫폼 운영 차원에서 주의 깊게 검토될 필요가 있다(Favarel-Garrigues et al., 2020; Haugh, 2022; Lageson and Maruna, 2018).

이 점에서 정책적 함의 역시 도출될 수 있다. 온라인 커뮤니티의 과열된 논란을 완화하기 위한 논의는 단순히 개인의 책임만을 강조하기보다, 플랫폼이 참여 비용과 노출 구조를 어떻게 설계하는가의 문제와 함께 검토될 필요가 있다. 예컨대 특정 논란이 급격히 확산되는 시점에는 반복적 공격 표현에 대한 점검, 과열 징후 탐지, 일시적 노출 완화, 신고 및 경고 장치의 정교화 등이 검토될 수 있다. 일부 온라인 커뮤니티가 특정인에 대한 비하 표현이나 반복적 공격 표현을 제한하는 장치를 도입하는 것처럼, 플랫폼 차원의 개입은 이용자에게 발화의 사회적 책임을 환기하고 과도한 공격적 표현을 완화하는 데 기여할 수 있다. 다만 이러한 개입은 표현의 자유와 커뮤니티 자율성의 문제와 긴밀하게 연결되므로, 실제 제도 설계는 과잉 규제의 위험을 피하면서도 반복적 표적화와 공격적 확산을 완화하는 방향에서 신중하게 논의되어야 한다(Gillespie, 2010; Gorwa, Binns, and Katzenbach, 2020; Sander, 2019; Seering, 2020).

이러한 함의에도 불구하고 본 연구에는 몇 가지 제한이 있다. 첫째, 디시인사이드라는 단일 커뮤니티에 집중하였기 때문에, 디시인사이드가 한국 온라인 여론 형성에서 큰 영향력을 갖는 공간이라는 점에서 연구의 의미는 분명하지만, 플랫폼 고유의 문화와 상호작용 규범 역시 결과에 반영되었을 가능성이 있다. 둘째, 논란 관련 게시물 식별과 감성분석 과정에는 방법론적 한계가 존재한다. 논란 관련 게시물 식별에 사용한 휴먼코딩 표본은 1,021개 규모로 제한적이며, 평가 세트 기준 분류 성능 역시 일정한 오분류 가능성을 내포한다. 특히 학습 자료에서 논란 관련 게시물의 비중이 상대적으로 낮았기 때문에 클래스 불균형에 따른 성능 차이가 나타났고, 논란 관련 클래스의 precision과 recall 역시 각각 0.596과 0.651에 머물렀다. 따라서 본 연구의 결과는 이러한 보수적 조건하에서도 유동닉 게시물의 감성 점수가 일관되게 낮게 나타났다는 점에서 해석할 필요가 있다. 후속 연구에서는 더 큰 규모의 라벨링 자료와 클래스 불균형을 보정한 학습 전략을 활용하여 논란 관련 게시물 식별의 안정성을 추가로 검증할 필요가 있다. 셋째, 본 연구는 논란 유형과 참여 방식의 차이에 따른 전반적 패턴을 보여주는 데에는 강점을 가지지만, 구체적인 상호작용 맥락과 의미 형성 과정까지 충분히 포착하기 위해서는 후속 질적 연구가 병행될 필요가 있다.

향후 연구에서는 플랫폼 간 비교, 댓글·대댓글 수준의 상호작용 분석, 그리고 특정 논란이 어떻게 집단적 감성으로 조직되고 재유포되는지에 대한 질적 추적을 통해 이러한 논의를 더욱 확장할 수 있을 것이다. 특히 다른 형태의 온라인 커뮤니티나 소셜 네트워크 플랫폼과의 비교를 통해, 유사한 플랫폼 조건에서 부정적 감성 표현의 확산과 규범적 비판의 조직 방식이 어떻게 달라지는지를 확인하는 작업이 중요하다. 이러한 연구가 축적될 때, 온라인 논란 담론의 감성적·규범적 조직 방식에 대한 경험적·이론적 이해도 한층 심화될 수 있을 것이다.

참고문헌

강옥희. 2023. “덕질과 새로운 콘텐츠의 생산과 수용.”『한국엔터테인먼트산업학회논문지』 17(7): 295-304.

Kang, Okhee. 2023. “Deokjil(Fanaticism) and Production and Acceptance of New Content.” *Journal of Korea Entertainment Industry Association* 17(7): 295-304.

권지현·김명혜. 2014. “디시인사이드 ‘주군의 태양’ 갤러리 구성원 유형화를 통한 인터넷 팬 커뮤니티 권력관계 특성연구.”『지역과 커뮤니케이션』18(3): 5-48.

Kwon, Jihyun, and Myoung Hye Kim. 2014. “Characteristics of Power Relations in Internet Fan Community - A Case Study of DCinside Gallery ‘Master’s Sun’ -.” *Locality & Communication* 18(3): 5-48.

김수정·김수아. 2015. “‘집단적 도덕주의’ 에토스: 혼종적 케이팝의 한국적 문화정체성.”『언론과 사회』23(3): 5-52.

Kim, Sujeong, and Sooah Kim. 2015. “The Ethos of Collective Moralism: The Korean Cultural Identity of K-Pop.” *Journalism and Society* 23(3): 5-52.

박성희·박정윤. 2006. “온라인 팬덤 커뮤니케이션 특성 연구: 디시인사이드 스타크래프트 갤러리의 임요환 관련 게시물을 중심으로.”『사회과학연구논총』15: 45-62.

Park, Sunghee, and Jungyun Park. 2006. “On the Characteristics of Online Fandom Communication: The Case of Im Yo Hwan, the Professional Gamer of Starcraft Galaxy in DCInside.” *Ewha Journal of Social Sciences* 15: 45-62.

송경재. 2005. “네트워크 시대의 인터넷 정치참여-탄핵정국 디시인사이드 정치토론 게시판을 중심으로.”『담론 201』8(3): 123-160.

Song, Kyong Jae. 2005. “Political Participation in the IT Based Network Age.” *Discourse 201* 8(3): 123-160.

송준모·강정한. 2018. “메갈리아의 두 딸들: 익명성 수준에 따른 온라인 커뮤니티의 정체성 분화.”『한국사회학』 52(4): 161-206.

Song, Junmo, and Jeonghan Kang. 2018. “Two Daughters of Megalia: Identity Differentiation in Response to the Anonymity Level in Online Community.” *Korean Journal of Sociology* 52(4): 161-206.

신미르·신유현. 2024. “적은 양의 음성 및 텍스트 데이터를 활용한 멀티 모달 기반의 효율적인 감정 분류 기법.”『정보처리학회 논문지』 13(4): 174-180.

Shin, Mirr, and Youhyun Shin. 2024. “Efficient Emotion Classification Method Based on Multimodal Approach Using Limited Speech and Text Data.” *The Transactions of the Korea Information Processing Society* 13(4): 174-180.

신지명·김경규·박주연. 2011. “사회적 익명성이 커뮤니티 품질에 미치는 영향: 표현욕구, 자기통제, 유대감의 매개효과를 중심으로.”『한국전자거래학회지』 16(4): 257-281.

Shin, Ji-Myoung, Kyung-Kyu Kim, and Joo-Yeon Park. 2011. “The Effect of Social Anonymity on Community Quality: Mediated by Self Expression, Self Control, and Community Ties.” *The Journal of Society for e-Business Studies* 16(4): 257-281.

신희연·한혜원. 2025. “K-pop 4세대 여성 아이돌의 음악과 서사의 상관관계 연구: 에스파와 뉴진스를 중심으로.”『글로벌문화콘텐츠』 63: 205-226.

Shin, Heeyeon, and Hyewon Han. 2025. “A Study on the Correlation between Music and Narrative Structure of 4th Generation Female Idols in K-pop: Focusing on aespa and NewJeans.” *Global Cultural Contents* 63: 205-226.

오관석. 2009. “디지털 군중(Digital mobs)의 시민성 부재에 관한 연구.”『사회과학연구』 25(1): 53-78.

Oh, Kwan-suk. 2009. “A Study on the Absence of Citizenship of Digital Mobs.” *Social Science Research* 25(1): 53-78.

오세정. 2012. “K-POP의 선호 요인: 미주 지역의 K-POP 소비자를 대상으로.”『주관성연구』 24: 205-223.

Oh, Se Jung. 2012. “Attractiveness factors in K-POP: Focused on American Consumer of K-Pop in U. S. A.” *Journal of Korean Society for the Scientific Study of Subjectivity* 24: 205-223.

유지연. 2020. “허구와 현실을 오가는 은밀한 공동체로서의 온라인 팬덤 사례 연구.”『문화콘텐츠연구』 18: 77-118.

Yoo, Jiyun. 2020. “A Case Study of Online Fandom as a Clandestine Imagined Community Between Reality and Fiction.” *The Journal of Culture Contents* 18: 77-118.

이영준·박하명·강승식. 2023.12. “KcELECTRA를 활용한 혐오성 댓글 분류.” 한국정보과학회 2023 한국소프트웨어종합학술대회.

Lee, Youngjun, Hamyoung Park, and Seungshik Kang. 2023.12. “Classifying Hateful Comments with KcELECTRA.” Korean Institute of Information Scientists and Engineers.

장민지. 2016. “디지털 네이티브 여/성주체(Digital Native Fe/male Subject)의 운동 전략.”『미디어, 젠더 & 문화』31(3): 219-255.

Jang, Min-Gi. 2016. “Social Movement Strategy of Digital Native Fe/male Subject: Focused on ‘Megalia’.” *Media, Gender & Culture* 31(3): 219-255.

조선비즈. 2023. 5. 25. “구글 4월 이용자 수, ‘네이버+다음’ 제쳤다... 韓 인터넷 시장 내줄 위기.” https://biz.chosun.com/it-science/ict/2023/05/23/WHGI3WC4ONAAJDSVZKMCL6PCBM/?utm_source=naver&utm_medium=original&utm_campaign=biz.

ChosunBiz. 2023. 5. 25. “Google’s April User Count Surpasses the Combined Users of Naver and Daum... Korea Risks Losing Its Internet Market.” https://biz.chosun.com/it-science/ict/2023/05/23/WHGI3WC4ONAAJDSVZKMCL6PCBM/?utm_source=naver&utm_medium=original&utm_campaign=biz.

최희재·이준기·김정원. 2009. “온라인 커뮤니티에서 기술적, 사회적 익명성이 자기통제에 미치는 영향에 관한 연구.”『인터넷정보학회논문지』10(1): 63-74.

Choi, Heejai, Zoonky Lee, and Jungwon Kim. 2009. “An Empirical Study on the Effect of Technical and Social Anonymity on the Self-Control Behavior.” *Journal of Internet Computing and Services* 10(1): 63-74.

한국국제문화교류진흥원. 2023. “4세대? 5세대? 아이돌 세대론은 어떻게 만들어지는가?” (음악평론가 김영대 기고문). https://www.archivecenter.net/hallyuresearch/archive/collection/ArchiveCollectionView.do?con_id=3267.

Korea Foundation for International Cultural Exchange. 2023. “Fourth Generation? Fifth Generation? How Are Idol Generations Defined?” (Essay by music critic Kim Young-dae). https://www.archivecenter.net/hallyuresearch/archive/collection/ArchiveCollectionView.do?con_id=3267.

한국리서치. 2024. 5. 28. “[기획] 연예인의 역할 및 대중의 알 권리에 대한 인식.” https://hrcopinion.co.kr/archives/29982.

Hankook Research. 2024. 5. 28. “[Planning] Perceptions of the Role of Celebrities and the Public’s Right to Know.” https://hrcopinion.co.kr/archives/29982.

황병화·이은규. 2025. "인터넷 커뮤니티에서의 익명성과 사용자의 단어사용 사이의 관계 분석 연구."『한국통신학회논문지』50(1): 105-115.

Hwang, Byeonghwa, and Eun-Kyu Lee. 2025. "Analysis of the Impact of Anonymity on Members' Use of Words in Online Communities." *The Journal of Korean Institute of Communications and Information Sciences* 50(1): 105-115.

Abd-Rahim, Atiqah. 2019. "Online Fandom: Social Identity and Social Hierarchy of Hallyu Fans." *Journal for Undergraduate Ethnography* 9(1): 65-81.

Aburime, Samantha. 2022. "Hate Narratives, Conditioned Language and Networked Harassment: A New Breed of Anti-Shipper and Anti-Fan – Antis." *Journal of Fandom Studies* 10(2-3): 135-155.

Ahn, Ji-Hyun, and Tien-wen Lin. 2019. "The Politics of Apology: The 'Tzuyu Scandal' and Transnational Dynamics of K-Pop." *International Communication Gazette* 81(2): 158-175.

Aitchison, Guy, and Saladin Meckled-Garcia. 2021. "Against Online Public Shaming: Ethical Problems with Mass Social Media." *Social Theory and Practice* 47(1): 1-31.

Cho, Daegon, and Kyounghee H. Kwon. 2015. "The Impacts of Identity Verification and Disclosure of Social Cues on Flaming in Online User Comments." *Computers in Human Behavior* 51: 363-372.

Christopherson, Kimberly M. 2007. "The Positive and Negative Implications of Anonymity in Internet Social Interactions: "On the Internet, Nobody Knows You're a Dog"." *Computers in Human Behavior* 23(6): 3038-3056.

Clark, Kevin, Minh-Thang Luong, Quoc V. Le, and Christopher D. Manning. 2020. "Electra: Pre-Training Text Encoders as Discriminators Rather Than Generators." Arxiv Preprint Arxiv:2003.10555.

Click, Melissa A. 2019. *Anti-Fandom: Dislike and Hate in the Digital Age*. New York, NY: Nyu Press.

Cohen, Stanley. 2019. "Folk Devils and Moral Panics: The Creation of Mods and Rockers (1972-2002)." pp.461-482 in *Crime and Media*, edited By Chris Greer. Oxfordshire, UK: Routledge.

Critcher, Chas. 2008. "Moral Panic Analysis: Past, Present and Future." *Sociology Compass* 2(4): 1127-1144.

_____. 2009. "Widening the Focus: Moral Panics as Moral Regulation." *British Journal of Criminology* 49(1): 17-34.

Duffett, Mark. 2013. *Understanding Fandom: An Introduction to the Study of Media Fan Culture*. London, UK: Bloomsbury.

Favarel-Garrigues, Gilles, Samuel Tanner, and Daniel Trottier. 2020. "Introducing Digital Vigilantism." *Global Crime* 21(3-4): 189-195.

Forestal, Jennifer. 2024. "Social Media, Social Control, and the Politics of Public Shaming." *American Political Science Review* 118(4): 1704-1718.

Friedman, Eric J., and Paul Resnick. 2001. "The Social Cost of Cheap Pseudonyms." *Journal of Economics & Management Strategy* 10(2): 173-199.

Gillespie, Tarleton. 2010. "The Politics of "Platforms"." *New Media & Society* 12(3): 347-364.

_____. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.

Gorwa, Robert, Reuben Binns, and Christian Katzenbach. 2020. "Algorithmic Content Moderation: Technical and Political Challenges in the Automation of Platform Governance." *Big Data & Society* 7(1). https://doi.org/10.1177/2053951719897945

Gray, Jonathan C. 2003. "New Audiences, New Textualities: Anti-Fans and Non-Fans." *International Journal of Cultural Studies* 6(1): 64-81.

_____. 2007. "Antifandom and the Moral Text." In *Fandom: Identities and Communities in a Mediated World*, edited By Jonathan C. Gray, Lee Harrington, and Cornel Sandvoss. New York, NY: New York University Press.

Haugh, Michael. 2022. "(Online) Public Denunciation, Public Incivilities and Offence." *Language & Communication* 87: 44-59.

Hier, Sean P. 2008. "Thinking Beyond Moral Panic: Risk, Responsibility, and the Politics of Moralization." *Theoretical Criminology* 12(2): 173-190.

Ivaturi, Koteswara, and Cecil Chua. 2019. "Framing Norms in Online Communities." *Information & Management* 56(1): 15-27.

Jeon, Duyoung, Junho Lee, and Cheongtag Kim. 2022. "User Guide for KOTE: Korean Online Comments Emotions Dataset." Arxiv Preprint Arxiv:2205.05300.

Kang, In-Hye, and Taehoon Park. 2022. "Cultural Differences in Consumer Responses to Celebrities Acting Immorally: A Comparison of The United States and South Korea." *Journal of Business Ethics* 180: 373-389.

Kim, Mikay, Morgan Elizabeth Ellithorpe, and Sybil Alexandra Burt. 2023. "Anonymity and Its Role in Digital Aggression: A Systematic Review." *Aggression and Violent Behavior* 72: 101856.

King-O'Riain, Rebecca Chiyoko. 2021. ""They Were Having So Much Fun, So Genuinely...": K-Pop Fan Online Affect and Corroborated Authenticity." *New Media & Society* 23(9): 2820-2838.

Lageson, Sarah E., and Shadd Maruna. 2018. "Digital Degradation: Stigma Management in the Internet Age." *Punishment & Society* 20(1): 113-133.

Lapidot-Lefler, Noam, and Azy Barak. 2012. "Effects of Anonymity, Invisibility, and Lack of Eye-Contact on Toxic Online Disinhibition." *Computers in Human Behavior* 28(2): 434-443.

Lee, Hosun, Dae Ryun Chang, and Sabine Einwiller. 2020. "A Study on the Dynamics Between the Moral Reasoning Process and Celebrity Image and Their Impact on Consumers' Support for Celebrity Comebacks After a Transgression." *Journal of Product & Brand Management* 29(6): 729-743.

Lee, Jin, and Crystal Abidin. 2024. "Backdoor Advertising Scandals, Yingyeo Culture, and Cancel Culture Among Youtube Influencers in South Korea." *New Media & Society* 26(1): 405-425.

Lee, Shin Haeng, Jin-Young Tak, Eun-Joo Kwak, Tae Yun Lim. 2020. "Fandom, Social Media, and Identity Work: The Emergence of Virtual Community Through the Pronoun "We"." *Psychology of Popular Media* 9(4): 436-446.

Moore, Alfred, Rolf Fredheim, Dominik Wyss, and Simon Beste. 2021. "Deliberation and Identity Rules: The Effect of Anonymity, Pseudonyms and Real-Name Requirements on the Cognitive Complexity of Online News Comments." *Political Studies* 69(1): 45-65.

Nakamura, Ryota, Keita Nishimoto, Ichiro Sakata, and Kimitaka Asatani. 2025. "Unveiling the Listener Structure Underlying K-Pop's Global Success: A Large-Scale Listening Data Analysis." Arxiv Preprint Arxiv:2509.06606.

Park, Sungjin, Jihye Lee, Seungjin Ryu, and Kyu S. Hahn. 2015. "The Network of Celebrity Politics: Political Implications of Celebrity Following on Twitter." *The Annals of the American Academy of Political and Social Science* 659(1): 246-258.

Phillips, Whitney. 2015. *This Is Why We Can't Have Nice Things: Mapping the Relationship Between Online Trolling and Mainstream Culture*. Cambridge, MA: MIT Press.

Postmes, Tom, Russell Spears, Khaled Sakhel, and Daphne de Groot. 2001. "Social Influence in Computer-Mediated Communication: The Effects of Anonymity on Group Behavior." *Personality and Social Psychology Bulletin* 27(10):

1243-1254.

Reicher, Stephen D., Russell Spears, and Tom Postmes. 1995. "A Social Identity Model of Deindividuation Phenomena." *European Review of Social Psychology* 6(1): 161-198.

Rifon, Nora J., Mengtian Jiang, and Shuang Wu. 2023. "Consumer Response to Celebrity Transgression: Investigating the Effects of Celebrity Gender and Past Transgressive and Philanthropic Behaviors Using Real Celebrities." *Journal of Product & Brand Management* 32(4): 517-529.

Rösner, Leonie, and Nicole C. Krämer. 2016. "Verbal venting in the social web: Effects of anonymity and group norms on aggressive language use in online comments." *Social Media + Society* 2(3): 2056305116664220.

Rost, Katja, Lea Stahel, and Bruno S. Frey. 2016. "Digital Social Norm Enforcement: Online Firestorms in Social Media." *PLoS ONE* 11(6): e0155923.

Samosir, Nora, and Lionel Wee. 2023. *Sociolinguistics of the Korean Wave: Hallyu and Soft Power*. Oxfordshire, UK: Routledge.

Sander, Barrie. 2019. "Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation." *Fordham International Law Journal* 43(4): 939-1006.

Seering, Joseph. 2020. "Reconsidering Self-Moderation: The Role of Research in Supporting Community-Based Models for Online Content Moderation." *Proceedings of the ACM on Human-Computer Interaction* 4(CSCW2): 107. https://doi.org/10.1145/3415178

Suler, John. 2004. "The Online Disinhibition Effect." *Cyberpsychology, Behavior, and Social Networking* 7(3): 321-326.

Sykes, Gresham M., and David Matza. 2017. "Techniques of Neutralization: A Theory of Delinquency." pp.33-41 in *Delinquency and Drift Revisited, Volume 21*, edited by Thomas G. Blomberg, Francis T. Cullen, Christoffer Carlsson, Cheryl Lero Jonson. Oxfordshire, UK: Routledge.

Treem, Jeffrey W., and Paul M. Leonardi. 2013. "Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association." *Annals of the International Communication Association* 36(1): 143-189.

Trottier, Daniel. 2017. "Digital Vigilantism as Weaponisation of Visibility." *Philosophy & Technology* 30: 55-72.

Vilanova, Felipe, Francielle Machado Beria, Ângelo Brandelli Costa, and Silvia Helena Koller. 2017. "Deindividuation: From Le Bon to the Social Identity Model of

Deindividuation Effects.” *Cogent Psychology* 4(1): 1308104.

Wang, Erika Ningxin, and Liang Ge. 2023. “Fan Conflicts and State Power in China: Internalised Heteronormativity, Censorship Sensibilities, and Fandom Police.” *Asian Studies Review* 47(2): 355-373.

Williams, Matthew. 2007. “Policing and Cybersociety: The Maturation of Regulation Within An Online Community.” *Policing & Society* 17(1): 59-82.

Yoo, SoYeop, HaYoung Lee, JeIn Song, and OkRan Jeong. 2023. “A Korean Emotion-Factor Dataset for Extracting Emotion and Factors in Korean Conversations.” *Scientific Reports* 13: 18547.

Zimmerman, Adam G., and Gabriel J. Ybarra. 2016. “Online Aggression: The Influences of Anonymity and Social Modeling.” *Psychology of Popular Media Culture* 5(2): 181-193.

조성권 박사후연구원은 성균관대학교에서 사회학과 박사 학위를 수여받고 현재 성균관대 BK21 소셜이노베이션 융합전공 박사후연구원으로 재직 중이다. 주된 관심 분야는 텍스트 분석을 활용한 저널리즘 및 인권, 문화 연구이다.

김민석 박사과정생은 성균관대학교 사회학과 박사과정에 재학 중이며, 현재 <K-pop 팬덤의 역동성과 새로운 공동체 가능성 연구> 소속으로 활동 중이다. 주된 관심 분야는 온라인 커뮤니티에서의 혐오표현에 대한 사회학적 분석이다.

김재훈 석박통합과정생은 한양대학교 인공지능학과 석박통합과정에 재학 중이며, 현재 <K-pop 팬덤의 역동성과 새로운 공동체 가능성 연구> 소속으로 활동 중이다. 주된 관심 분야는 인간-컴퓨터 상호작용, 소셜 컴퓨팅, 자연어처리이다.

구정우 교수는 스탠포드대학교에서 박사학위를 수여받고, 현재 성균관대 사회학과 교수 및 BK21 소셜이노베이션 융합전공 부단장으로 역임 중이다. 또한, 성균관대 ‘인권과개발센터’ 소장 및 <K-pop 팬덤의 역동성과 새로운 공동체 가능성 연구>의 책임교수로 활동 중이며, 주된 관심 분야는 인권, 기업과 사회적 책임, 혐오 표현 연구이다.

[2026.3.4 접수; 2026.4.30 수정; 2026.5.21. 게재확정]

부록

〈Appendix 1〉 Descriptive Statistics of K-Pop Idol Galleries

Artist	gen	Posts	Ephemeral ratio	Fixed ratio	Artist	gen	Posts	Ephemeral ratio	Fixed ratio
Kang Daniel (강다니엘)	3	4967	97.69	2.31	JBJ	3	2003	76.36	23.64
MONSTA X (몬스타엑스)	3	1996	95.45	4.55	TWICE (트와이스)	3	2215	75.43	24.57
BTOB (비투비)	2.5	1861	93.95	6.05	IVE (아이브)	4	563	74.13	25.87
JYJ	2	3753	92.82	7.18	WJSN (우주소녀)	3	2278	73.03	26.97
EXO (엑소)	3	5770	92.22	7.78	FTISLAND (FT아일랜드)	2	2270	71.48	28.52
BTS (방탄소년단)	3	4515	91.89	8.11	KARA (카라)	2	2891	70.26	29.74
Wanna One (워너원)	3	3381	91.12	8.88	Girls' Generation (소녀시대)	2	2724	70.2	29.8
Jay Park (박재범)	2	1383	90.32	9.68	Suzy (수지)	2.5	2599	69.68	30.32
WINNER (위너)	3	1357	89.13	10.87	Taeyeon (태연)	2	1567	69.4	30.6
AOA (에이오에이)	2.5	2883	88.53	11.47	LE SSERAFIM (르세라핌)	4	1802	69.09	30.91
INFINITE (인피니트)	2.5	2306	88.07	11.93	MAMAMOO (마마무)	3	2453	67.47	32.53
SHINee (샤이니)	2	2021	86.97	13.03	ITZY (있지)	4	916	65.72	34.28
SEVENTEEN (세븐틴)	3	1571	86.84	13.16	GFRIEND (여자친구)	3	2240	64.59	35.41
B.A.P	2.5	1731	86.5	13.5	2am	2	1184	64.01	35.99
2NE1	2	2634	86.23	13.77	Onew (온유)	2	1206	63.91	36.09
MBLAQ (엠블랙)	2	1502	84.37	15.63	Girl's Day (걸스데이)	2.5	2349	61.63	38.37
NCT (엔시티)	3	739	82.86	17.14	Block B (블락비)	2.5	1351	59.9	40.1
EXO-K	3	2183	82.44	17.56	(G)I-DLE ((여자)아이들)	4	2230	59.57	40.43
Seungri (승리)	2	1801	82.29	17.71	Wendy (웬디)	3	920	58.78	41.22

BLACKPINK (블랙핑크)	3	2542	81.91	18.09	ZE:A (제국의아이들)	2.5	1248	57.33	42.67
SS501	2	1831	79.78	20.22	Lovelyz (러블리즈)	3	1878	56.66	43.34
N.Flying (엔플라잉)	3	335	79.58	20.42	IU (아이유)	2	3651	51.82	48.18
ASTRO (아스트로)	3	439	78.63	21.37	aespa (에스파)	4	1754	51.66	48.34
Apink (에이핑크)	2.5	1378	77.94	22.06	Jessica (제시카)	2	1553	41.12	58.88
Roy Kim (로이킴)	3	2818	77.88	22.12	T-ARA (티아라)	2	3658	38.29	61.71
NU'EST (뉴이스트)	2.5	1528	77.82	22.18	NewJeans (뉴진스)	4	2197	35.74	64.26
Seulgi (슬기)	3	563	77.32	22.68	LOONA (이달의 소녀)	4	1390	34.23	65.77
Kang Seung Yoon (강승윤)	3	935	77.12	22.88	CNBLUE (씨엔블루)	2.5	2479	26.39	73.61
Park Yoo-chun (박유천)	2	3667	77.07	22.93					

〈Appendix 2〉 List of Emotion Categories in KOTE Classification

Valence	Interpretation	Example words in the cluster
Negative	dissatisfaction (불평/불만)	dissatisfied, oppose, criticize, complaint
	embarrassment (당황/난처)	embarrassed, disconcerted, awkward, untoward
	irritation (짜증)	irritated, pissedoff, ridiculous
	sadness (슬픔)	sad, miss, lonely, tear
	despair (절망)	frustrated, joys&sorrows, hurt, grief, letdown
	shame (부끄러움)	ashamed, humiliated
	boredom (재미없음)	bored, tedium, trite, dull
	disappointment (안타까움/실망)	disappointed, sorry, upset, deplorable, regretful
	disgust (역겨움/징그러움)	disgusted, repulsive, dirty
	shock (경악)	shocked, flabbergasted, passout, freakedout
	reluctant (부담/안 내킴)	unwilling, denial, pressure, giveup
	fear (공포/무서움)	fear, anxious, tense, pressed
	contempt (증오/혐오)	contempt, hatred, scorn, vilifying
	guilt (죄책감)	guilt, blamed, repentance, remorse
	anxiety (불안/걱정)	apprehensive, worry, threatened
	distrust (의심/불신)	suspicious, doubtful, lie
	anger (화남)	anger, rage, obsessed, fury
	gessepany (패배/자기혐오)	failure, miserably, extorted
	laziness (귀찮음)	bothered, dawdling
	sorrow (서러움)	sorrowful, mirthless, weary, sobbing, upset
	fed up (지긋지긋)	fedup, struggle, arduous, sickandtired
	preposterous (어이없음)	dumbfounded, stunned, sttufy, enervated, WTF
	compassion (불쌍함/연민)	pity, sadly, chockedup, heartrending
	pathetic (한심함)	pathetic, belittled, stupid, impudence
	exhaustion (힘듦/지침)	tired, peak, exhausted
Positive	admiration (감동/감탄)	admiring, great, praise, compliment
	happiness (행복)	happy, affection, valuable, hope, luck
	joy (기쁨)	delight, ecstasy, love
	gratitude (고마움)	praiseworthy, commendable, favor, blessing
	excitement (즐거움/신남)	excited, funny
	care (아껴주는)	caring, adore, dear
	expectancy (기대감)	new, achieve, together, harmonious, vitality
	comfort (편안/쾌적)	comfortable, ease, cozy, cool, warm
	welcome (환영/호의)	welcome, approval, kindness, enthusiastic
	interest (신기함/관심)	interested, curious
	relief (안심/신뢰)	relief, trust, intimate, close
	respect (존경)	respect, loyal, veneration, follow, obedience
	attracted (흐뭇함(귀여움/예쁨))	handsome, pretty, sweet, thrilled, cute, aegyo
	pride (뿌듯함)	successful, victory, worthwhile, accomplish

Neutral	arrogance (우쭐댐/무시함)	arrogance, pompous, ignore, bragging, boast
	surprise (놀람)	astonished, startled
	realization (깨달음)	realize, enlightened, wakened, conviction, belief
	resolute (비장함)	resolute, determination

Note: The emotion categories were translated into Korean from the English classifications reported in Jeon, Lee, and Kim (2022).

〈Appendix 3〉 Regression Analysis Based on a 0.5 Threshold for Sentiment Scores

	(1)		(2)		(3)		(4)	
	Coef	S.E	Coef	S.E	Coef	S.E	Coef	S.E
Sex(d)	0.000	0.036	0.026	0.041	0.025	0.040	0.004	0.014
Type(d)	-0.088*	0.039	-0.062	0.042	-0.070	0.039	-0.062	0.009
Foreigner(d)	0.124	0.114	0.136	0.130	0.128	0.132	-0.075*	0.035
Stock market(d)	0.114***	0.026	0.092**	0.029	0.098**	0.030	0.067***	0.011
Generation(ref. 2nd)								
2.5th generation	0.064	0.038	0.134**	0.052	0.129**	0.049	0.011	0.017
3rd generation	-0.060	0.048	-0.023	0.059	-0.025	0.057	-0.058***	0.015
4th generation	-0.060	0.051	-0.038	0.056	-0.043	0.057	0.016	0.015
Anonymous(d)	-0.197***	0.031	-0.173***	0.031	-0.140**	0.052	-0.135*	0.060
Controversy FE	No		Yes		Yes		Yes	
Anonym×Controversy	No		No		Yes		Yes	
Artist FE	No		No		No		Yes	
Constant	-0.167***		-0.192***		-0.214***		0.045	
Log likelihood	-106940		-106323		-106167		-104506	
Adjusted R^2	0.03		0.04		0.04		0.07	

N=122,277, (d): dummy variable

Note: *p<.01, **p<.05, ***p<.001; Models 1 and 2 estimate the average association between anonymity and sentiment, whereas Models 3 and 4 examine whether this association varies across controversy categories through interaction terms.

〈Appendix 4〉 Regression Analysis Based on Non Threshold for Sentiment Scores

	(1)		(2)		(3)		(4)	
	Coef	S.E	Coef	S.E	Coef	S.E	Coef	S.E
Sex(d)	0.001	0.019	0.014	0.021	0.013	0.020	0.003	0.007
Type(d)	-0.042*	0.019	-0.031	0.020	-0.034	0.018	-0.066***	0.005
Foreigner(d)	0.059	0.057	0.058	0.066	0.054	0.067	-0.042*	0.019
Stock market(d)	0.055***	0.012	0.047**	0.015	0.050**	0.016	0.041***	0.005
Generation(ref. 2nd)								
2.5th generation	0.032	0.020	0.068**	0.025	0.066**	0.024	0.012	0.009
3rd generation	-0.023	0.025	-0.005	0.030	-0.005	0.029	-0.017*	0.008
4th generation	-0.030	0.026	-0.016	0.028	-0.017	0.028	0.013	0.008
Anonymous(d)								
Controversy FE	No		Yes		Yes		Yes	
Anonym×Controversy	No		No		Yes		Yes	
Artist FE	No		No		No		Yes	
Constant	-0.012		-0.029		-0.036		0.084***	
Log likelihood	-13417		-12671		-12504		-10561	
Adjusted R^2	0.04		0.05		0.05		0.08	

N=122,277, (d): dummy variable

Note: *p<.01, **p<.05, ***p<.001; Models 1 and 2 estimate the average association between anonymity and sentiment, whereas Models 3 and 4 examine whether this association varies across controversy categories through interaction terms.

〈Appendix 5〉 External Robustness Check

	(1)		(2)	
	Coef	S.E	Coef	S.E
Sex(d)	-0.023	0.013	0.002	0.015
Type(d)	-0.122***	0.010	-0.138***	0.010
Foreigner(d)	-0.152***	0.024	-0.098*	0.038
Stock market(d)	0.096***	0.009	0.081***	0.012
Generation(ref. 2nd)				
2.5th generation	-0.014	0.015	0.027	0.019
3rd generation	0.061***	0.016	-0.040**	0.013
4th generation	0.061***	0.016	0.027	0.014
Anonymous(d)	-0.093**	0.032	-0.173***	0.034
Controversy FE	Yes		Yes	
Anonymous×Controversy	Yes		Yes	
Artist FE	Yes		Yes	
Constant	-0.486***		0.322***	
Log likelihood	-135652		-141746	
Adjusted R^2	0.03		0.03	

N=122,277, (d): dummy variable

Note: For both Models 1 and 2, only the full specifications including controversy fixed effects, Anonymous×Controversy interaction terms, and artist fixed effects are reported.

본 연구에서 종속변수는 게시물 단위 감성 점수이므로, 본문에서 확인된 회귀분석 결과가 특정 감성분석 모델이나 특정 점수 산식에 과도하게 의존하지 않는지를 점검하는 것은 연구의 타당성을 확보하는 데 중요한 의미를 갖는다. 이에 따라 본 연구는 KOTE 기반 감성 점수의 내부 강건성 검정에 더해, 별도의 극성 기반 감성 모형을 활용한 외부 강건성 검정을 추가로 수행하였다. 구체적으로 huggingface에서 한국어 게시물에 적용가능한 두 개의 감성분석 모델, 즉 모델 1: multilingual-sentiment-analysis와 모델 2: koelectra-base-v3-generalized-sentiment-analysis를 동일한 논란 게시물 집합에 적용하여 게시물 단위 감성 점수를 새롭게 구성하고, 본문과 동일한 회귀모형을 반복 추정하였다.

첫 번째 모델인 multilingual-sentiment-analysis는 Very Negative, Negative, Neutral, Positive, Very Positive의 5점 극성 범주에 대한 예측 확률을 산출한다. 두 번째 모델인 koelectra-base-v3-generalized-sentiment-analysis는 긍정/부정의 이진 확률을 출력하므로, 부정=-1, 긍정=1의 가중치를 적용하여 동일하게 확률가중평균 방식의 게시물 단위 감성 점수를 산출하였다. 이러한 방식은 서로 다른 구조의 감

성모형 출력을 공통된 연속형 감성 지표로 변환함으로써, 본문의 회귀분석과 직접 비교 가능한 형태의 종속변수를 구성하기 위한 것이다.

그 결과, 두 외부 감성모형을 사용한 경우, 참여 방식 변수인 유동닉은 모두 음(-)의 회귀계수를 나타냈다. 즉, 감성 점수의 세부 산식이나 사용한 모델이 달라지더라도, 식별가능성이 낮은 참여 방식이 상대적으로 더 낮은 감성 점수와 관련된다는 본문의 핵심 결과는 일관되게 유지되었다. 이는 본문의 결과가 특정 KOTE 산식 하나에만 의존한 것이 아니라, 서로 다른 극성 기반 감성모형을 적용하더라도 비교적 안정적으로 재현되는 패턴임을 보여준다. 다시 말해, 본 연구의 실증적 결론은 단일 모델의 선택에 의해 우연히 도출된 결과라기보다, 게시물의 감성 방향을 연속형 점수로 요약하는 서로 다른 모델링 접근에서도 대체로 반복적으로 확인되는 경향이라고 해석할 수 있다.

Platform Identifiability and the Sentiment of Controversy Discourse: An Analysis of K-Pop Idol Galleries on DCInside

Seong-Gwon Cho
Sungkyunkwan University
Min-seok Gim
Sungkyunkwan University
Jae-hoon Kim
Hanyang University
Jeong-woo Koo
Sungkyunkwan University

As K-pop idols have emerged as symbolic representatives of global cultural brands, controversies surrounding them have increasingly moved beyond entertainment news to become objects of rapid circulation and public evaluation in online communities. In South Korea, idols are often subject to high normative expectations, and controversies involving them tend to generate strong sentiment-driven responses among fans and broader publics. This study examines how platform-based differences in participation modes shape the sentiment of controversy discourse. Focusing on DCInside, one of South Korea's largest online communities, this study analyzes its distinctive structure in which registered users with persistent "fixed nicknames" and unregistered users with temporary "ephemeral nicknames" coexist. This structure makes it possible to examine whether differences in user identifiability are associated with the direction and intensity of sentiment in controversy-related posts. The study collected 714,778 posts from 57 K-pop idol galleries covering second-generation and later idols. A KcELECTRA-based classifier identified 122,277 posts directly related to social controversies, and the KOTE model was used to construct post-level sentiment scores from fine-grained emotion probabilities. Regression analyses controlling for idol characteristics, controversy types, and idol fixed effects show that posts written under lower identifiability conditions have significantly lower sentiment scores. This pattern is especially pronounced in controversies involving political or social statements, fandom conflicts, and agency/member disputes. These findings suggest that online controversy responses are shaped not only by event characteristics but also by platform identifiability structures and reputational costs.

Keywords: K-pop idols, social controversy, platform identifiability, DCInside, sentiment analysis