

Chọn mẫu trong phân tích nội dung định giới trên Tiktok hiện nay

Sampling in the analysis of gender-bias content on TikTok today

PGS, TS. BÙI THU HƯƠNG

Học viện Báo chí và Tuyên truyền;

Email: huongbui75@gmail.com

NGUYỄN THỊ HƯƠNG GIANG

Học viện Báo chí và Tuyên truyền;

Email: nguyenhuonggiang122005@gmail.com

Ngày nhận bài: 06/5/2026

Ngày phản biện: 12/5/2026

Ngày duyệt đăng: 20/6/2026

Chủ đề của bài viết thuộc lĩnh vực: Khoa học xã hội và nhân văn

Tóm tắt: TikTok đã và đang trở thành nền tảng truyền thông phổ biến và môi trường xã hội giới quan trọng đối với giới trẻ. Các nội dung chứa định kiến giới được lan truyền trên Tiktok cũng nhanh chóng trở thành mối quan tâm của nhiều nhà nghiên cứu. Tuy nhiên, với đặc thù của nền tảng mạng xã hội này, mỗi nhà phân tích nội dung lại sử dụng các chiến lược chọn mẫu và xác định cỡ mẫu khác nhau. Dựa trên cơ sở tổng hợp và phân tích một số nghiên cứu thực nghiệm tiểu biểu có liên quan trong khoảng thời gian từ 2022 đến 2025, bài viết giới thiệu và luận giải hai chiến lược chọn mẫu phổ biến hiện nay. Đó là chiến lược chọn mẫu dựa trên hashtag và chọn mẫu có chủ đích theo tài khoản. Đồng thời, bài viết cũng thảo luận về ngưỡng cỡ mẫu tối thiểu cho nghiên cứu phân tích nội dung định lượng và định tính.

Từ khoá: TikTok; định kiến giới; phân tích nội dung; chọn mẫu, cỡ mẫu.

Abstract: TikTok has become one of the most popular communication platforms and an important social environment for young people. Content containing gender stereotypes disseminated on TikTok has consequently attracted increasing attention from researchers. However, owing to the unique characteristics of this social media platform, content analysts employ different sampling strategies and sample size determination methods. Based on a review and analysis of selected empirical studies published between 2022 and 2025, this article introduces and discusses two of the most commonly used sampling strategies: hashtag-based sampling and purposive account-based sampling. The article also discusses minimum sample size thresholds for both quantitative and qualitative content analysis.

Keywords: TikTok; gender stereotypes; content analysis; sampling; sample size.

1. Đặt vấn đề

Với hơn một tỷ người dùng hoạt động thường xuyên tính đến năm 2025, TikTok có thể được xem như một không gian và môi trường xã hội hoá về giới quan trọng, và có ảnh hưởng đặc biệt đối thế hệ Z (1997-2012) và thế hệ Y/Millennials (1981-1996) trên phạm vi toàn cầu. Đặc biệt, cơ chế đề xuất nội dung thông qua trang For You Page (FYP) được cá nhân hóa bằng thuật toán học máy đặc thù của Tiktok đã và đang giúp tạo ra một môi trường thông tin không chỉ phản ánh mà còn tích cực tái sản xuất

và khuếch đại các khuôn mẫu quy định những đặc điểm, vị trí, vai trò phù hợp của nam và nữ trong tất cả các mối quan hệ xã hội (Yin & Abdullah, 2024).

Các quan tâm nghiên cứu gần đây về vấn đề này đã ghi nhận nhiều hình thức định kiến giới phổ biến trên Tiktok, từ xu hướng kỳ thị và phân biệt đối xử nhắm vào phụ nữ (Yin & Abdullah, 2024), các thông điệp nam tính độc hại trong cộng đồng những người có tính cách vượt trội và khác biệt sigma (Tanner & Gillardin, 2025), đến việc chuẩn hóa các vai trò giới truyền thống thông qua các xu hướng nội dung như

#StayAtHomeGirlfriends (Tirocchi & Taddeo, 2024). Song song đó, không ít nghiên cứu đã quan sát được sự phản kháng và tái thương thuyết các khuôn mẫu giới của chính người dùng nền tảng (Krippendorff, 2025).

Tuy nhiên, một điểm đáng chú ý là giữa các công trình nghiên cứu về định kiến giới trên TikTok không có sự nhất quán trong cách thức lựa chọn mẫu và xác định cỡ mẫu. Một số nghiên cứu dựa trên các mẫu có quy mô nhỏ, được lựa chọn thuận tiện (Tirocchi & Taddeo, 2024), trong khi những nghiên cứu khác phân tích hàng nghìn video nhưng thiếu biện giải rõ ràng về chiến lược chọn mẫu và tính đại diện của mẫu (Tanner & Gillardin, 2025). Sự thiếu chuẩn hóa phương pháp này không chỉ gây khó khăn cho việc so sánh và tổng hợp kết quả mà còn tạo ra những điểm mù nghiêm trọng giới hạn những hiểu biết về cơ chế lan truyền và tác động của định kiến giới trên nền tảng kỹ thuật số. Xuất phát từ khoảng trống đó, bài báo này bàn luận về một số chiến lược chọn mẫu và cỡ mẫu phù hợp đang được sử dụng cho các nghiên cứu phân tích nội dung về định kiến giới trên TikTok hiện nay.

2. Phương pháp nghiên cứu

Bài viết sử dụng kết quả phân tích nội dung định tính 10 bài báo khoa học đăng trên các tạp chí có uy tín trong khoảng thời gian từ 2022 đến 2025. Các bài viết được tìm kiếm thông qua nền tảng Google Scholar và Semantic Scholar sử dụng một số nhóm từ khoá cơ bản như: (1) phân tích nội dung/thông điệp định kiến giới trên Tiktok; (2) chiến lược chọn mẫu và cỡ mẫu cần thiết trong phân tích nội dung nói chung và trên Tiktok nói riêng. Các bài viết được tuyển chọn để phân tích nội dung phải thỏa mãn một số tiêu chí như: (1) Là kết quả tổng quan và/hoặc kết quả của một nghiên cứu thực nghiệm liên quan đến phân tích nội dung truyền thông và Tiktok trong vòng 5 năm trở lại đây; (2) Phải tường minh về phương pháp nghiên cứu; và (3) Phải có thể tải được toàn văn.

Cuối cùng, mỗi bài trong số 10 bài báo khoa học được lựa chọn đều được đọc kỹ, các câu và phân đoạn quan trọng được phân tích và mã hoá, các mã tương đồng được nhóm và xâu chuỗi thành các chủ đề thể hiện các quan điểm chung của các tác giả về phương pháp phân tích nội dung định kiến giới trên

Tiktok, các chiến lược chọn mẫu và cỡ mẫu phù hợp. Sau khi rà soát và đối chiếu với bối cảnh cụ thể của từng bài viết, các chủ đề được tinh chỉnh, sắp xếp và cấu trúc lại để phù hợp với mục đích nghiên cứu riêng của nhóm tác giả bài viết này. Kết quả nghiên cứu được trình bày trong các phần kế tiếp, theo hai chủ đề chính là: (1) Cách thức chọn mẫu trong phân tích nội dung định kiến giới trên Tiktok và (2) Xác định cỡ mẫu trong phân tích nội dung định kiến giới trên Tiktok.

3. Kết quả nghiên cứu và thảo luận

3.1. Cách thức chọn mẫu trong phân tích nội dung định kiến giới trên Tiktok

Có thể thấy trong nghiên cứu về định kiến giới trên truyền thông số, đặc biệt là trên TikTok, đơn vị phân tích có thể là các video đơn lẻ, bình luận, hashtag, hoặc một chuỗi tương tác (Tirocchi & Taddeo, 2024) hoặc kết hợp phân tích video với phân tích bình luận nhằm nắm bắt cả chiều sản xuất lẫn tiếp nhận của thông điệp (Werdhaningtyas et al., 2025).

Đặc biệt, không thể biết chính xác số lượng và truy cập toàn bộ nội dung video về định kiến giới được đăng tải trên TikTok tại bất kỳ thời điểm nào. Giao diện API nghiên cứu chính thức của TikTok (TikTok Research API) có giới hạn truy vấn tối đa 100 bản ghi mỗi lần gọi và chỉ cung cấp siêu dữ liệu chứ không phải nội dung video đầy đủ (Marynak et al., 2022). Điều này khiến hầu hết các nghiên cứu phân tích nội dung TikTok về định kiến giới phải dựa vào các chiến lược lấy mẫu phi xác suất và có chủ đích. Do vậy, các chiến lược chọn mẫu cũng được thiết kế để phản ánh thực tế này.

3.1.1. Chọn mẫu dựa trên hashtag

Chiến lược lấy mẫu phổ biến nhất trong nghiên cứu TikTok về định kiến giới là chọn mẫu sử dụng hashtag. Trên mạng xã hội, hashtag là các một từ/cụm từ viết liền đặt sau dấu thăng (#) được sử dụng để phân loại, nhóm các nội dung và kết nối người dùng có cùng quan tâm). Theo đó, nhà nghiên cứu xác định một hay một danh mục các hashtag liên quan đến chủ đề và thu thập các video sử dụng những hashtag đó. Chiến lược này có ưu điểm quan trọng là cho phép kiểm soát tương đối về tính liên quan chủ đề của mẫu, đồng thời tận dụng thực tế rằng người tạo nội dung trên TikTok thường gán hashtag

để định vị nội dung trong một cộng đồng hoặc xu hướng cụ thể (Marynak et al., 2022).

Tanner và Gillardin (2025) áp dụng chiến lược này một cách có hệ thống trong nghiên cứu về nam tính độc hại, thu thập toàn bộ video có chứa ba hashtag #sigma, #sigmamale và #sigmamalegrindset tương ứng với khoảng 320 video mỗi hashtag trước khi chạm ngưỡng tối đa hiển thị của nền tảng, tạo thành kho dữ liệu ban đầu khoảng 953 video trước khi chọn lọc. Trong nghiên cứu về hình ảnh của phụ nữ trên mạng xã hội, Tirocchi và Taddeo (2024) sử dụng hashtag #StayAtHomeGirlfriends và thu nhận được hơn 328,3 triệu lượt xem trong một năm, cho thấy tiềm năng phủ sóng rộng của dữ liệu thu thập theo chiến lược này (tr. 4).

Tuy nhiên, chọn mẫu dựa trên hashtag có những hạn chế nội tại đáng lưu ý. Về bản chất, hashtag do người tạo nội dung lựa chọn, gán cho các bài đăng và chỉ phản ánh ý kiến và cách hiểu của cá nhân họ về chủ đề, và do vậy, hashtag được gán cho bài đăng không nhất thiết phản ánh trung thực thông điệp thực sự của bài đăng. Tương tự, nhiều bài đăng mang định kiến giới nhưng lại không có gán các hashtag mà các nhà nghiên cứu muốn tìm kiếm. Do vậy, trong một số trường hợp, các nhà nghiên cứu có thể thu thập được được nhiều bài đăng không liên quan và bỏ sót một lượng lớn các bài đăng khác có nội dung về định kiến giới do không được gán các hashtag liên quan trực tiếp (Krippendorff, 2025). Đây cũng chính là một trong những hạn chế của nghiên cứu mà nhóm nghiên cứu cần làm rõ khi thiết kế và báo cáo kết quả nghiên cứu.

Về quy trình thực tiễn, chọn mẫu dựa trên hashtag hiệu quả nhất cần kết hợp ba nhiệm vụ cơ bản.

Thứ nhất, nhóm nghiên cứu cần xây dựng một danh sách hashtag đa tầng bao gồm các hashtag trực tiếp đến hashtag liên quan để giảm thiểu tối đa nguy cơ bỏ sót các nội dung tốt. Ví dụ danh mục hashtag ban đầu về định kiến giới không chỉ là những hashtag trực tiếp như #dinhkiengioi hay #MeToo mà còn có thể mở rộng hơn như #congaiphainhuthenao, #congaingoan, #girlmath, #phunuvietnam, #phunuhiendai, #danongdichthuc, #danonglamviectha, #sigma, #heforshe, v.v. Danh mục hashtag có thể được mở rộng trong quá trình làm việc trực tiếp

với các bài đăng.

Danh mục hashtag đa tầng này được phát triển và tổng hợp từ:

(1) tổng quan tài liệu và thao tác hoá các khái niệm công cụ;

(2) theo dõi và tìm kiếm trực tiếp trên Tiktok;

(3) gợi ý của Tiktok. Sau khi đã có một danh mục hashtag toàn diện, nhóm nghiên cứu có thể sử dụng và tìm kiếm các bài đăng liên quan trong phạm vi thời gian xác định của nghiên cứu và sau đó có thể sử dụng các tiêu chí lọc để loại và chọn được các bài đăng đưa vào mẫu phân tích nội dung.

Tuy nhiên, nhóm nghiên cứu cũng có thể tiến hành lọc luôn và chỉ chọn 10 đến 50 hashtag quan trọng nhất để tìm kiếm các bài đăng dựa vào một số tiêu chí như:

(1) mức độ phổ biến hay tổng số lượt xem và số bài đăng sử dụng hashtag đó;

(2) mức độ bao phủ của chủ đề (phù hợp với mục đích nghiên cứu, thuộc nhóm các chỉ báo và các thuộc tính mô tả đối tượng nghiên cứu);

(3) tính liên quan (không được phép quá chung, liên quan đến giới và định kiến giới, phải sử dụng tiếng Việt).

Ngoài ra, nhóm nghiên cứu có thể sử dụng công cụ Trung tâm sáng tạo Tiktok (TikTok Creative Center) miễn phí để tìm ra những hashtag và các bài đăng thịnh hành nhất (có thể có hoặc không chứa các hashtag thịnh hành) có liên quan đến chủ đề nghiên cứu trong một khoảng thời gian xác định. Từ danh mục đó, nhóm có thể tuyển chọn các bài đăng phù hợp nhất với mục đích nghiên cứu để phân tích (Krippendorff, 2025).

Thứ hai, nhóm nghiên cứu cần sử dụng tài khoản được đăng ký riêng cho nghiên cứu và truy cập sử dụng hàng ngày để có thể thu thập được các đề xuất nội dung mà nhóm quan tâm và sử dụng để mở rộng các tìm kiếm tiếp theo. Một số nhà nghiên cứu có thể sử dụng luôn tài khoản tạo lập này như một địa bàn nghiên cứu và từ đó chỉ tập trung thu thập và phân tích các đề xuất nội dung của Tiktok trong Trang cho bạn trong một khoảng thời gian nào đó. Theo đó, nhà nghiên cứu tạo ra các tài khoản được kiểm soát (tài khoản con rôi) với hành vi tương tác được lập trình sẵn với các nội dung liên quan đến giới và định kiến

giới để quan sát và ghi lại cách thức thuật toán TikTok xác định và đề xuất nội dung trong hệ sinh thái thực tế về các chủ đề này. Trên cơ sở đó, nhóm nghiên cứu không chỉ có thể trả lời được câu hỏi về thông điệp nào về định kiến giới đang được lan truyền và củng cố mà còn loại thông điệp nào về định kiến giới nhận được lợi thế hiển thị từ hệ thống gợi ý của TikTok. Đây là chiều kích quan trọng bởi vì thuật toán Trang của bạn (FYP) không chỉ phản ánh nhu cầu người dùng mà còn tích cực định hình những thông điệp định kiến giới có khả năng và cần được tiếp cận rộng rãi nhất. Hạn chế chính là tính không thể tái lập của kết quả, vì cùng một chuỗi hành vi được lập trình có thể tạo ra các đề xuất khác nhau tùy thuộc vào thời điểm và điều kiện thuật toán (Tanner & Gillardin, 2025).

Thứ ba, nhóm nghiên cứu cần chuẩn bị một danh mục các tiêu chí loại trừ để lựa chọn bài đăng phù hợp cho phân tích nội dung. Một số tiêu chí cơ bản bao gồm:

(1) Về phạm vi (loại các bài đăng ngoài khoảng thời gian xác định, các bài đăng lặp lại, các bài đăng được đặt ở chế độ riêng tư hoặc bị xoá trước khi mã hoá, các video phải được phép tải và lưu, v.v.);

(2) Mức độ phổ biến (loại các bài có mức độ tương tác thấp hơn 10.000 lượt xem hoặc thấp hơn 1000 lượt thích);

(3) Mức độ liên quan (loại bỏ nội dung rác, nội dung quảng cáo, nội dung không phù hợp với hashtag, nội dung sử dụng ngôn ngữ khác với ngôn ngữ của nhà nghiên cứu) (Krippendorff, 2025).

3.1.2. Chọn mẫu có chủ đích theo tài khoản

Chọn mẫu có chủ đích là phương pháp trong đó nhà nghiên cứu chủ động lựa chọn các đơn vị nghiên cứu dựa trên mức độ phù hợp với mục tiêu nghiên cứu. Khi áp dụng cho TikTok, đơn vị chọn mẫu đầu tiên không phải là các bài đăng hay video mà là tài khoản. Sau khi chọn được các tài khoản phù hợp, nhà nghiên cứu mới lấy các video từ các tài khoản đó làm dữ liệu phân tích. Chọn mẫu theo tài khoản có một số lợi thế như:

(1) Phần lớn nội dung TikTok gắn với người sáng tạo nội dung, do vậy, phân tích theo tài khoản không thường xuyên hoặc thường xuyên đăng tải về các chủ đề giới và định kiến giới nói riêng giúp nhận diện

khuynh hướng diễn ngôn ổn định về chủ đề quan tâm nghiên cứu;

(2) Thuật toán gợi ý của TikTok không cho phép tiếp cận tổng thể dữ liệu một cách ngẫu nhiên, do vậy phân tích nội dung các bài đăng phù hợp với mục đích nghiên cứu từ một hay một vài tài khoản tiêu biểu có thể đảm bảo thu thập được các dữ liệu có hệ thống và kiểm soát được;

(3) Các tài khoản tiêu biểu có lượng người theo dõi lớn, có số lượng video lan truyền rộng rãi sẽ có tác động xã hội lớn hơn nhiều so với các bài đăng/video đơn lẻ được lựa chọn vào mẫu theo các cách khác.

Theo đó, tiêu chí lựa chọn tài khoản thường bao gồm: (1) số lượng người theo dõi đủ lớn; (2) tần suất đăng nội dung liên quan đến chủ đề giới và định kiến giới, và (3) giới tính của người tạo nội dung (Bueno-Fernandes & Campos-Pellanda, 2022). Khung lấy mẫu để lựa chọn các tài khoản tiêu biểu phù hợp có thể là Danh sách các nhà sáng tạo nội dung đạt các giải thưởng của chính TikTok Việt Nam (TikTok Award Việt Nam) năm gần nhất. Từ khung lấy mẫu này, nhóm nghiên cứu có thể chọn ra 5-10 TikToker đạt giải có nội dung đề cập đến định kiến giới và sau đó chọn các bài đăng phù hợp. Nhà phân tích có thể hoặc phân tích toàn bộ hoặc một mẫu có hệ thống các nội dung bài đăng từ các tài khoản đã tuyển chọn.

Ưu điểm của chiến lược chọn mẫu này là cho phép phân tích theo chiều dọc, tức là nhà nghiên cứu có thể theo dõi sự biến đổi của thông điệp về định kiến giới theo thời gian hoặc trong bối cảnh phát triển sự nghiệp của người sáng tạo nội dung. Hạn chế chính là nguy cơ thiên lệch lựa chọn: các tài khoản có nhiều người theo dõi có thể giúp nhà nghiên cứu hiểu biết sâu về một hiện tượng cụ thể chứ không nhất thiết đại diện cho toàn bộ cách thức định kiến giới được sản xuất và tiêu thụ trên nền tảng (Tirocchi & Taddeo, 2024).

3.2. Xác định cỡ mẫu trong phân tích nội dung định kiến giới trên TikTok

3.2.1. Cỡ mẫu trong phân tích nội dung định lượng

Cần phải lưu ý rằng cho đến nay chưa có một hệ thống tiêu chí thống nhất nào giúp cho các nhà phân tích nội dung lựa chọn cỡ mẫu phù hợp. Do vậy, trên

thực tế, các nhà phân tích nội dung định lượng thường tham khảo và học hỏi từ các nhà nghiên cứu đi trước hoặc sử dụng các phương pháp giúp xác định cỡ mẫu hiện có. Một phương pháp tính cỡ mẫu được nhiều nhà nghiên cứu lựa chọn đó là phương pháp xác định cỡ mẫu sử dụng các công thức tính độ sai số chuẩn và khoảng tin cậy. Sau khi xác định được các thông số này, tùy theo trường hợp biết/không biết quy mô quần thể cần nghiên cứu, nhà nghiên cứu có thể đưa vào các công thức có sẵn để tính hoặc sử dụng các trang mạng giúp tính cỡ mẫu miễn phí và các bảng gợi ý cỡ mẫu sẵn có theo các quy mô quần thể nghiên cứu khác nhau (Bùi Thu Hương, 2017, tr. 77-80).

Trong thực tiễn phân tích nội dung TikTok về định kiến giới theo tiếp cận định lượng, để có thể đo lường tần suất, phân bố và mô hình của các biểu hiện định kiến giới, phạm vi cỡ mẫu dao động khá rộng. Nghiên cứu năm 2025 của Tanner và Gillardin phân tích 195 video từ kho dữ liệu 953 video ban đầu, bằng cách chọn lọc các video có chứa nội dung phù hợp với hệ thống danh mục đã xác định của họ. Trong nghiên cứu về lạm dụng nicotine trên TikTok, Marynard và cộng sự (2022) đã thu thập 100 000 bài đăng cho hashtag *#nicotineaddiction* để đảm bảo vừa đủ bảo hòa nội dung chủ đề vừa thực hiện được phân tích thống kê mô tả.

Như trên đã phân tích, nhà nghiên cứu cần lấy mẫu vượt mức từ 20-30% so với cỡ mẫu mục tiêu để bù đắp cho việc loại trừ các đơn vị không đáp ứng tiêu chí hoặc không còn khả dụng do bị xóa khỏi nền tảng (Hennink & Kaiser, 2022). Nguyên tắc này đặc biệt quan trọng trong nghiên cứu nội dung nhạy cảm như định kiến giới, vì các nền tảng có xu hướng gỡ bỏ nội dung vi phạm chính sách nhanh hơn các nội dung khác.

3.2.2. Cỡ mẫu trong phân tích nội dung định tính

Phân tích nội dung định tính có thể chia thành ba loại cơ bản: (1) Phân tích nội dung theo chủ đề; (2) Phân tích diễn ngôn phê phán; (3) Phân tích ký hiệu học. Định hướng bởi các câu hỏi *“như thế nào”* và *“tại sao”*, các nhà phân tích nội dung định tính thường xác định cỡ mẫu theo tiêu chí bảo hòa dữ liệu (Hennink & Kaiser, 2022). Bảo hòa dữ liệu là điểm mà tại đó các dữ liệu thu thập thêm không còn mang

đến cho nhà nghiên cứu những hiểu biết, chủ đề hay phạm trù mới nào liên quan đến câu hỏi nghiên cứu. Có bốn loại bảo hòa phổ biến. Thứ nhất là bảo hòa lý thuyết. Đây là khái niệm do Glaser và Strauss giới thiệu khi bàn về phương pháp lý thuyết nền từ năm 2017. Loại bảo hòa này đạt được khi dữ liệu bổ sung không còn góp phần tinh chỉnh hoặc mở rộng khung lý thuyết đang được phát triển. Thứ hai là bảo hòa dữ liệu/chủ đề đề cao độ rộng của tập dữ liệu cả về phạm vi và sự đa dạng của các nguồn dữ liệu khác nhau được thu thập. Bảo hòa dữ liệu đạt được khi không có khuôn mẫu hay phạm trù mới nào xuất hiện thêm trong quá trình phân tích, đảm bảo bao quát toàn diện hiện tượng đang được nghiên cứu. Thứ ba là bảo hòa mã. Loại này tập trung vào mức độ chi tiết của phân tích và theo đó cho biết điểm mà không cần thêm bất cứ mã nào để mô tả dữ liệu đầy đủ hơn. Cuối cùng, bảo hòa ý nghĩa mở rộng ra ngoài việc xác định các chủ đề để đảm bảo sự hiểu biết sâu sắc và tinh tế về trải nghiệm của người tham gia, ưu tiên sự phong phú và chiều sâu của diễn giải (Ahmed, 2025).

Ngưỡng bảo hòa chịu ảnh hưởng nhiều yếu tố như mục đích nghiên cứu, thiết kế nghiên cứu, tính đồng nhất của mẫu, chất lượng số liệu muốn thu thập, thời gian và kinh phí cho phép và năng lực của người phân tích. Trong các nghiên cứu nội dung định tính, đôi khi việc đạt được ngưỡng bảo hòa đến rất nhanh. Tức là chỉ sau khi phân tích sâu một vài trường hợp điển hình, nhóm nghiên cứu đã thấy xuất hiện sự lặp lại và không thu thập thêm được thông tin mới nữa. Để tránh việc bỏ sót thông tin và có thể đạt được ngưỡng bảo hòa hiệu quả hơn, nhà phân tích nội dung thường sử dụng phương pháp lấy mẫu lặp đi lặp lại, hay nói cách khác là nhà nghiên cứu liên tục thu thập và phân tích dữ liệu cho đến khi đạt được độ bảo hòa.

Đáng chú ý là trước khi tiến hành chọn mẫu, nhóm nghiên cứu cần xác định loại bảo hòa cần đạt được, thao tác hóa và đưa ra hệ thống các chỉ số cụ thể minh họa cho ngưỡng bảo hòa và lựa chọn chiến lược chọn mẫu phù hợp với mục đích nghiên cứu. Trên cơ sở đó, nhóm nghiên cứu bắt đầu bằng việc xác định một cỡ mẫu cơ sở. Cỡ mẫu ban đầu này là số lượng đơn vị mẫu tối thiểu được lựa chọn chủ đích

và giúp nhóm nghiên cứu có thể từ đó tạo dựng được cơ sở dữ liệu ban đầu sử dụng để đối sánh với các thông tin thu thập thêm sau này (Marynak et al., 2022). Cuối cùng, sau những nỗ lực thu thập và phân tích dữ liệu bổ sung, chỉ khi đạt được ngưỡng bão hoà như mong đợi, nhà nghiên cứu mới ghi nhận được chính xác cỡ mẫu cần thiết cho nghiên cứu của mình. Trong thực tiễn nghiên cứu, tùy thuộc vào loại bão hoà mà nhóm nghiên cứu xác định, một cỡ mẫu vừa đủ có thể dao động trong khoảng từ 9-40 phỏng vấn sâu (Wutich et al., 2024).

Trong các nghiên cứu phân tích định tính nội dung TikTok về định kiến giới, nếu lựa chọn đơn vị phân tích chính là các video thì ngưỡng bão hoà dữ liệu ở đây đồng nghĩa với thời điểm khi các video tiếp theo không cung cấp góc nhìn mới về thực hành định kiến giới mà chỉ xác nhận các kiểu mẫu đã được phát hiện. Theo những tranh luận về ngưỡng bão hoà được trình bày ở trên, nhóm nghiên cứu có thể quyết định và luận giải cho việc lựa chọn một cỡ mẫu cần thiết nào đó tùy thuộc vào: (1) các giả định nghiên cứu của nhà nghiên cứu; (2) mục tiêu phân tích; (3) loại và chất lượng dữ liệu định tính cần phân tích; (4) kỹ năng thu thập thông tin của nhóm nghiên cứu; (5) sự đa dạng của nhóm khách thể và cách thức lấy mẫu; và (6) mức độ phổ biến của các chủ đề phân tích trong tổng thể nghiên cứu (nếu chủ đề phổ biến thì chỉ cần một cỡ mẫu nhỏ và ngược lại (Wutich et al., 2024).

4. Kết luận

Nghiên cứu thông điệp về định kiến giới trên TikTok đặt ra những thách thức phương pháp luận đặc thù xuất phát từ kiến trúc thuật toán của nền tảng, tính phù du của nội dung và sự bất khả thi của lấy mẫu xác suất toàn diện. Bài báo đã luận giải rằng không có chiến lược chọn mẫu đơn lẻ nào tối ưu cho mọi câu hỏi nghiên cứu, và sự lựa chọn chiến lược phải được điều chỉnh theo mục tiêu phân tích, thiết kế phương pháp (định lượng, định tính hay hỗn hợp), và đặc thù của hiện tượng định kiến giới được nghiên cứu. Bên cạnh đó, sự phát triển nhanh chóng của nền tảng TikTok trong thuật toán, trong cơ chế giao diện lập trình ứng dụng cho nghiên cứu và sự xuất hiện của các hình thức nội dung mới đòi hỏi rằng các nhà nghiên cứu phải liên tục cập nhật và điều chỉnh

phương pháp luận của mình. Không có giải pháp phương pháp cổ định nào hoàn toàn thỏa đáng trong một môi trường nghiên cứu đang biến đổi liên tục như truyền thông xã hội thế hệ mới hiện nay./.

TÀI LIỆU THAM KHẢO

1. Ahmed, S. K. (2025). Sample size for saturation in qualitative research: Debates, definitions, and strategies. *Journal of Medicine, Surgery, and Public Health*, 5, 100171. <https://doi.org/10.1016/j.glmedi.2024.100171>
2. Bueno-Fernandes, A.C. and Campos-Pellanda, A. (2022). Gender stereotypes in TikTok and Instagram: A reverse engineering experiment for understanding the mechanisms of social network algorithms. *Universitas-XXI*, 37, 233-255.
3. Bùi Thu Hương (2017). Thiết kế nghiên cứu - Một số vấn đề cơ bản (tr. 77-80). Nxb. Thế Giới.
4. Hennink, M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
5. Krippendorff, K. (2025). Remixing stereotypes: How gendered scripts evolve through TikTok's viral culture. *Exploratio Journal*. <https://exploratiojournal.com/remixing-stereotypes-how-gendered-scripts-evolve-through-tiktoks-viral-culture/> (Truy cập 18/03/2026)
6. Marynak, K.K., Robichaud, O.M., Puryear, T., Kennedy, D.R. and Moran, B.M. (2022). #Nicotineaddiction on TikTok: A quantitative content analysis of top-viewed posts. *Tobacco Induced Diseases*, 20(1), tr. 1-9.
7. Tanner, S., & Gillardin, F. (2025). Toxic communication on TikTok: Sigma masculinities and gendered disinformation. *Social Media+ Society*, 11(1), 1-12.
8. Tirocchi, S., & Taddeo, G. (2024). Unveiling informal learning of gender roles on TikTok: The #Stayathomegirlfriends phenomenon. *Journalism*, 25(9), 1987-2006.
9. Werdhaningtyas, D. T., Utari, P., & Anshori, M. (2025). Non-binary and androgynous gender representation of TikTok influencers in the digital era. *Jurnal Komunikasi: Malaysian Journal of Communication*, 12(2), 212-223.
10. Wutich, A., Beresford, M., & Bernard, H. R. (2024). Sample sizes for 10 types of qualitative data analysis: An integrative review, empirical guidance, and next steps. *International Journal of Qualitative Methods*, 23, 16094069241296206. <https://doi.org/10.1177/16094069241296206>.
11. Yin, Q., & Abdullah, K.B.B. (2024). Analysis of gender discourse bias and gender discrimination in social media: A case study of the TikTok platform. *Intercultural Communication*, 24(2), 93-122.