Digital Codebooks: Facilitating Real-Time Data Formatting in Agricultural Content Analysis

Joshue Lewis, M.S.

joshuele@ttu.edu

Agricultural Education and Communications

Texas Tech University

Krysti Kelley, Ph.D.

krystkel@ttu.edu

Agricultural Education and Communications

Texas Tech University

Digital Codebooks: Facilitating Real-Time Data Formatting in Agricultural Content Analysis

Introduction

Content analysis serves as a systematic research methodology that transforms qualitative communication into quantitative data through the rigorous definition of variables and the numerical assignment of categories (Krippendorff, 2019). This methodology is widely utilized across multiple disciplines, including agricultural fields like agricultural education and communications, where researchers leverage existing datasets to extract meaningful, evidence-based insights. This process relies on a structured codebook where each variable is clearly described, allowing coders to evaluate content and assign it to a specific category based on those definitions. The execution of this research design has evolved across various methodologies for data collection.

Traditionally, many researchers have relied on a manual approach, which involves printing physical copies of the codebook and handwriting the assigned categories onto coding sheets (Shellhouse & Baker, 2024). To improve efficiency, many transitioned to digital spreadsheets, using Microsoft Excel to organize the codebook and key numerical data directly into specific cells (Lewis & Irlbeck, 2025). Most recently, scholars have begun to integrate digital survey platforms like Qualtrics into the content analysis workflow. However, current reports indicate that these platforms are often used in a limited capacity, such as setting up the survey to track simple dichotomous variables where coders merely choose between the presence or absence of a specific attribute (Fanyinkah et al., 2025). While traditional methods often require a secondary, manual step of data entry, adopting an integrated digital platform is more efficient; it not only secures the data collection phase but also significantly accelerates the analysis by providing immediate downloads in the specific formatting required for statistical processing.

How it Works

The first step in this process was to create a well-structured codebook. Before setting everything up in Qualtrics, variables and categories needed to be clearly defined and reviewed. Once that foundation was ready, the actual work on the platform began. The researchers logged into Qualtrics and created a new survey, building it to match the structured codebook exactly. It was vital that the survey followed the same order and logic as the codebook, so it made sense to coders as they moved through the data collection. To make things easier, we included the detailed descriptions for every variable and category directly in the survey questions. This way, coders had the instructions right in front of them and could access them whenever they needed to make a decision.

Once the survey was built, it was ready to be published. Coders simply opened the link and selected the options that aligned with their analysis of the content. This digital setup really

showed its value when it came time to collect and analyze the data. Because the data was already digital, researchers did not have to spend additional time on manual entry. For intercoder reliability, it was exported as a .csv file and rearranged for ReCal. Once it was ready for final analysis, it was exported as a .sav file for use in SPSS. This made the transition from collecting to analyzing data much more efficient because the formatting was already complete.

Results to Date

This coding methodology was implemented within an agricultural communication social media monitoring study. The project utilized content analysis to identify nationwide trending topics related to the cotton industry. Three coders collected data via a Qualtrics survey to analyze 20 posts. The intercoder reliability for each variable (11 variables) ranged between .79 and 1.00, with an average coding time of 104 seconds per post. In agricultural education research, Qualtrics has been used to code data for two separate projects. The first project required coders to watch videos of preservice teachers and to score their teaching performance using an established evaluation instrument. This was completed by three coders across 17 teaching videos, with 24 coded items per video. The initial intercoder reliability ranged from .72 to .85. It was also used to score rubrics in a within-subjects comparative design by two coders, with 32 rubrics and five items each. The initial intercoder reliability ranged from .87 to 1.00.

Future Plans/Advice to Others

Utilizing Qualtrics for coding will continue to be implemented in future research projects. The scope of projects will be expanded to other contexts and collaborations with other disciplines. It is also planned to conduct an evaluation of this method of coding by comparing it with other content analysis coding strategies and assessing their effectiveness and efficiency. We have several recommendations for anyone planning to use Qualtrics for coding. First, we recommend that researchers create a codebook to guide the design of the coder survey before developing it, to aid in organization. It would be advantageous to include all items that need to be coded in a single Qualtrics survey. We also recommend directly incorporating the sources to be analyzed into the survey to improve efficiency. Additionally, including variable descriptions within the survey will allow clarity when selecting codes. Finally, we advise downloading the file from Qualtrics in the format that best aligns with your preferred inter-coder reliability software, such as a CSV file for ReCal or an SAV file for SPSS.

Costs/Resources Needed

The primary resource needed is a Qualtrics subscription, estimated at around $6,500 per year, depending on usage and features. Other digital survey platforms or tools could be used as alternatives, such as SurveyMonkey, Google Forms, or Microsoft Forms. Additionally, each coder will have an internet-enabled device, such as a computer or phone, to complete the survey. Device type should be selected based on the layout of the coding survey.

References

Fanyinkah, K. D., Miller, J. D., Cox, C., & Doss, W. (2025). Assessing the integration of QR codes in extension outreach communication graphics: A visual content analysis. *Journal of Applied Communications, 109*(2). https://doi.org/10.4148/1051-0834.2581

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

Lewis, J., & Irlbeck, E. (2025). Framing agricultural narratives: A content analysis of topics and trends on the U.S. Farm Report (2022-2024). *Journal of Applied Communications, 109*(4). https://doi.org/10.4148/1051-0834.2614

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). SAGE Publications.

Shellhouse, J. A., & Baker, L. (2024). Prominently portrayed? A content analysis exploring queer representation in food and agriculture magazines. *Journal of Applied Communications, 108*(4). https://doi.org/10.4148/1051-0834.2547