

## Validity and the Four “R’s”: (Re-)Establishing Standards for Content Analysis *SCM Special Issue*

### *Guest Editors*

Prof. Dr. Mario Haim (LMU München)

Dr. Valerie Hase (LMU München)

Johanna Schindler, M.A. (LMU München)

Dr. Marko Bachl (Universität Hohenheim)

Prof. Dr. Emese Domahidi (TU Ilmenau)

Content analysis is one of the main methods communication scholars regularly employ (Walter et al., 2018). It is also one of the few methods developed by our discipline (Loosen & Scholl, 2012). However, the method is undergoing fundamental changes: Visual, auditory, and audiovisual data from various online channels, often collected via automated approaches (Jünger et al., 2022), have gained a lot of importance. This has introduced challenges for sampling and analyzing relevant data (Ho, 2020; Jünger et al., 2022; Mahl et al., 2022). At the same time, concepts and methods from computer science (e.g., natural language processing, machine learning) are increasingly being incorporated into media and communication studies (Baden et al., 2022; Hase et al., 2022). This has transformed content analysis by introducing new and extending existing data, methods, and research processes (Bachl & Scharkow, 2017; Günther & Quandt, 2016; Nelson, 2020; Waldherr et al., 2019).

Consequently, the thus far institutionalized method faces several methodological debates concerning quality criteria and standards (Casas & Williams, 2022; Krippendorff, 2018, 2021; Lacy et al., 2015; Wirth et al., 2015). Discussing (the lack of) such criteria and standards, especially but not solely for computational advances (Domahidi et al., 2019; Geise & Waldherr, 2022; Haim, 2022), scholars criticize that, to date, “everyone brings the practices and standards from their original field” (Theocharis & Jungherr, 2021: p. 12). As such, “a lack of currently established standards [...] can jeopardize the scholarship scrutiny which is essential in assuring additive science and replicability” (van Atteveldt et al., 2019: p. 3). Based on its great expertise with the method, communication science can play a central role in establishing these standards for modern content analysis.

This Special Issue focuses on changes and challenges concerning quality criteria and standards for the method of qualitative, quantitative, and computational content analysis. This includes *validity* (e.g., Chan & Sältzer, 2020; Song et al., 2020)—that is, whether results correspond to some external truth and how to approximate such—and the four “R’s” to understand how changes in data or methods influence results: *reliability* to estimate whether repeated measures of the same data yield similar results, *reproducibility* to similarly assess if others yield the same results based on the same data and methods, *robustness* to estimate in how far results change when using different methods but the same data (Denny & Spirling, 2018; Pipal et al., 2022; Wilkerson & Casas, 2017), and *replicability* to evaluate conclusions based on the same methods yet different data.

We invite contributions that (systematically) reflect upon methodological shifts in and for content analysis. We also invite contributions that seek to advance quality criteria and standards in this regard. Ultimately, this Special Issue will showcase contributions that further develop quantitative, qualitative, or computational content analysis. Potential contributions could thus, but are not limited to, speak to one or several of these questions:

- How can mixed-methods approaches combining qualitative, quantitative, or computational perspectives help analyzing content?
- How can mixed-methods approaches help measuring validity?
- How can new tools benefit measuring validity or one of the four “R’s”?
- Which innovations are suitable to overcome the limitations of previous content-analytical methods with regard to validity, reliability, robustness, reproducibility, or replicability?
- How can scholars computationally identify theoretical concepts such as frames or topics?
- What are benchmarks for “valid” results?
- How suitable are computer-science measures like precision/recall for content analysis?
- How to estimate reliability for annotations generated via different modes, e.g., experts, crowd workers, or algorithms?
- What are relevant standards for reporting qualitative, quantitative, and computational content analysis?
- What are relevant standards for reproducing machine-learning approaches?
- How robust are results across (arbitrary) methodological decisions for analyzing content?
- How can scholars identify and illustrate robustness of or uncertainty in results?
- How to test and estimate the replicability of content analysis?

#### *Submission Instructions*

SCM is an Open Access Journal of the German Communication Association (DGPK) and Affiliate Journal of the International Communication Association (ICA). Accepted papers will be published as Open Access without additional costs.

Authors should submit an initial extended abstract (maximum of 500 words including references) and a title page including the list of author(s) to Mario Haim ([haim@ifkw.lmu.de](mailto:haim@ifkw.lmu.de)) as a single PDF by 15.01.2023. The guest editors will then decide about general thematic fit. Notifications as invitation to write a full manuscript will be sent by 31.01.2023. These invitations do not guarantee final publication as full manuscripts will undergo the journal’s usual double-blinded peer review. Full manuscripts can be written in English or German and should target a length of a Full Paper (about 60,000 characters with blanks) or Research-in-Brief (about 30,000 characters with blanks). Further guidelines can be found [here](#). Full manuscripts are due by 31.03.2023. The Special Issue is to be published in December as 4/2023.