

Konzept der AG 5 (Inter-Codierer-Reliabilitätstest)

Giulia D'Amico, Regina Eicher, Almira Medaric, Alexandra Muster

11. Oktober 2020

Inhaltsverzeichnis

1	Thema und Fragestellung	3
2	Stand der Forschung und Relevanz des Inter-Codierer-Reliabilitätstests	3
2.1	Stand der Forschung: Qualitätskontrolle der Inhaltsanalyse (Reliabilität und Validität)	3
2.2	Reliabilitätstest	4
2.3	Relevanz des Inter-Codierer-Reliabilitätstests	4
3	Vorgehensweise, Methode(n) und Lösungsansätze	5
3.1	Inhaltlicher Auftrag: Operationalisierung auf Statement-Ebene	5
3.2	Methodischer Auftrag: Vorbereitung und Koordination des Inter-Codierer-Reliabilitätstest	7
3.2.1	Berechnung des Reliabilitätskoeffizienten	7
3.2.2	Ergebnisdarstellung und Interpretation des Tests	8
3.2.3	Ablauf des Inter-Codierer-Reliabilitätstest in diesem Projekt	9
4	Zielsetzung	10
4.1	Inhaltlicher Auftrag: Operationalisierung auf Statement-Ebene	10
4.2	Methodischer Auftrag: Vorbereitung und Koordination des Inter-Codierer-Reliabilitätstests	10
5	Zeitplan, Schnittstellen, Abhängigkeiten	11
5.1	Zeitplanung und Arbeitspakete	11
5.2	Schnittstellen und Abhängigkeiten	12
5.3	Risiken	12
6	Fragen an Betreuerin	12

Abbildungsverzeichnis

1	Matrix für Codierübereinstimmungen bei einer Variable (Früh, 2017, S. 184) .	8
2	Interpretation der Werte nach Berechnung des Krippendorff's Alpha (nach Krippendorff)	9

Tabellenverzeichnis

1	Arbeitspakete und Meilensteine AG 5	11
---	---	----

1 Thema und Fragestellung

Im Rahmen des Projektkurses 'Lokaljournalismus', der im Herbstsemester 2020 stattfindet, hat die Arbeitsgruppe (AG) 5 zwei Aufträge zugewiesen erhalten:

1. **Inhaltlicher Auftrag:** Operationalisierung der theoretischen Kategorien (Leistungsindikatoren) auf Statement-Ebene
2. **Methodischer Auftrag:** Vorbereitung und Koordination des Inter-Codierer-Reliabilitätstest zu V1 und V2 des Codebuchs

Mit der Ausführung des inhaltlichen Auftrags wird ein Beitrag an das Codebuch geleistet, indem die Kategorien für die Analyseeinheit «Statements» operationalisiert werden. Erst dadurch werden diese messbar und können als Basis für die Beantwortung der Forschungsfrage dienen: **Wie kann die Leistungsfähigkeit von unterschiedlichen Anbietern (Legacy Media und neue Online-Only-Medien) mit Hilfe einer Inhaltsanalyse und dazugehörigen Leistungsindikatoren erfasst werden?**

Die methodische Aufgabe der AG 5 liegt darin, ein Konzept für die Durchführung von Reliabilitätstests zu erstellen und im Rahmen des Projekts auch konkret zwei durchzuführen: Der erste Test wird anhand des Codebuchs Version 1, der zweite mit der angepassten Codebuch Version 2 gemacht. Der Reliabilitätstest hat den Zweck, die Schwachstellen der ersten Version des Codebuchs zu identifizieren, um jene Erkenntnisse anschliessend an die AG 4 für die Überarbeitung des Codebuchs zur zweiten, finalen Version weiterzugeben. Ausserdem können auf diese Weise einzelne Codierer*innen identifiziert werden, die weitere Schulungen zum Codieren benötigen. Somit wird pro Test auf zwei Ebenen geprüft: Es werden einerseits die Kategorien (je Variable) und andererseits die Codierer*innen (paarweise) verglichen. Die im Rahmen der Tests zu untersuchenden Datensätze werden von der AG 2 gesammelt und den Codierenden zur Verfügung gestellt. Für jede Person gilt es pro Test 10 Artikel zu codieren, wobei jeweils zwei Personen denselben Artikel bearbeiten. Dies ergibt gesamthaft 80 unterschiedliche Artikel pro Testdurchlauf.

2 Stand der Forschung und Relevanz des Inter-Codierer-Reliabilitätstests

Nachstehend wird nur auf die Relevanz des methodischen Auftrags – die Vorbereitung und Koordination des Reliabilitätstests – eingegangen. Die Methode und Relevanz der Erstellung des Codebuchs und der Erarbeitung eines Leitfadens für die Codierenden fällt in den Zuständigkeitsbereich der Arbeitsgruppe 4.

2.1 Stand der Forschung: Qualitätskontrolle der Inhaltsanalyse (Reliabilität und Validität)

Nach Rössler (2017, S. 205) hängt die Aussagekraft einer Inhaltsanalyse im Wesentlichen von der Sorgfalt ab, mit der die Forschungsfrage in die Untersuchungsanlage übersetzt und wie das Analyseinstrument angewendet wurde. Ob Instrument und Messung tatsächlich geeignet sind, um Antworten auf die Forschungsfrage zu finden, ist eine Qualitätsfrage und bedarf einer Kontrolle.

Die wichtigsten Gütekriterien für Inhaltsanalysen sind deshalb die **Reliabilität** (auch Zuverlässigkeit) und die **Validität** (auch Gültigkeit) (Brosius, Haas & Koschel, 2016, S. 51 f.):

- **Reliabilität des Messinstruments:** Zuverlässigkeit der Messung und somit auch ihrer Wiederholbarkeit (bei wiederholter Messung sollte das gleiche Ergebnis erzielt werden).
- **Validität der Erhebung:** Gültigkeit der Messung (sie gibt an, ob ein Instrument tatsächlich das misst, was es messen soll).

Gemäss Rössler (2017, S. 206) kann die Reliabilität durch den Vergleich von unterschiedlichen Messungen mit dem gleichen Instrument bestimmt werden und das Ausmass der Abweichung lässt sich dann (ziemlich) genau mit dem Reliabilitätskoeffizienten angeben. Rössler betont, dass hingegen die Validität einer Inhaltsanalyse einer unabhängigen Messung bedarf und darüber hinaus oft auch einer inhaltlichen Diskussion.

Im Rahmen dieses Projekts wird nur die Reliabilität des Messinstruments mit Hilfe des Inter-Codierer-Reliabilitätstests (ICR-Test) überprüft.

2.2 Reliabilitätstest

Der Reliabilitätstest gehört im Ablauf der hier angewendeten Forschungsmethode der Inhaltsanalyse zur Phase der «Systematischen Verdichtung»: Nach der Operationalisierung der zentralen Konzepte und dem Erstellen des Codebuchs, muss dieses auf seine Zuverlässigkeit und somit auch auf die Wiederholbarkeit der Messung, unabhängig von der codierenden Person, überprüft werden (Dahinden & Dalmus, 2020, S. 19).

Nach Rössler (2017, S. 207) gibt es drei verschiedene Typen der Reliabilitätsmessung, die hier – um ein gemeinsames Verständnis der Begriffe zu erhalten – kurz definiert werden:

- **Intra-Codier-Reliabilität** (bei Rössler «Intracoder»): Wenn dieselbe codierende Person am Ende der Analyse nochmals das gleiche Material codiert, ohne ihre ersten Codierungen «zu kennen». Diese wird bei einer längeren Projektdauer mit der Inter-Codierer-Reliabilität kombiniert (Früh, 2017, S. 179).
- **Inter-Codierer-Reliabilität:** Wenn zwei Personen das gleiche Material getrennt voneinander codieren und dann verglichen wird, ob die jeweiligen Objekte den gleichen Kategorien zugeordnet wurden.
- **Forscher-Codierer-Reliabilität:** Wenn die/der Forscher'in selber codiert und die Resultate mit jenen der unabhängigen Codierer'innen verglichen werden.

2.3 Relevanz des Inter-Codierer-Reliabilitätstests

Wenn nach der Durchführung des ICR-Tests die Ergebnisse nicht innerhalb eines bestimmten, festgelegten Grades übereinstimmen, muss untersucht werden, wo die Schwachstellen des Codebuchs und der Codieranweisungen liegen und wie diese zu bereinigen sind. Problematische Kategorien können so identifiziert, das Kategoriensystem angepasst und die Anweisungen für die Codierung verbessert bzw. Schulungen wiederholt werden.

Nach Früh (2017, S. 179) ist der ICR-Test unverzichtbar, da die Reliabilität des Messinstruments für die Objektivität der Inhaltsanalyse notwendig ist. Hingegen wird die Intra-Codier-Reliabilität nur bei Bedarf ermittelt.

3 Vorgehensweise, Methode(n) und Lösungsansätze

3.1 Inhaltlicher Auftrag: Operationalisierung auf Statement-Ebene

Die Operationalisierung bedingt vorerst die Definition der Kategorien im Codebuch (Leistungsindikatoren). Eine Empfehlung derer wird von der Projektleitung (PL) vorgelegt. Nichtsdestotrotz wurde vereinbart, dass je Gruppe erste Überlegungen und Anhaltspunkte gesammelt werden und in das Konzept miteinfließen. Für die Kategorien müssen dann Indikatoren (oder auch: Ausprägungen) gefunden werden, anhand derer eine Messung möglich ist. Schliesslich muss für jeden Indikator das geeignete Skalenniveau eruiert werden. Um Missverständnissen und Unklarheiten vorzubeugen, werden diese Einträge um Definitionen und Beispiele ergänzt.

Anbei werden die Überlegungen für die Operationalisierung der Kategorien auf Statement-Ebene mitsamt möglicher Ausprägungen aufgelistet.

- **Verständlichkeit (Sprachgebrauch)** (Handstein o. J.)
 - Bericht enthält nicht-allgemeinverständliche Fachtermini, welche aufgeschlüsselt werden.
 - Bericht enthält nicht-allgemeinverständliche Fachtermini, welche nicht aufgeschlüsselt werden.
 - Bericht enthält keine nicht-allgemeinverständlichen Fachtermini.
- **Sachlichkeit (Argumentation)** (Stifterverein Medienqualität, 2020, S. 82)
 - Im Beitrag wird vor allem rational argumentiert (kognitiv-normativ): Ereignisse werden differenziert dargestellt und problematisiert, Pro- und Kontra-Positionen werden abgewogen, Schlussfolgerungen werden abgeleitet, Kritik sowie Forderungen werden begründet.
 - Im Beitrag wird vor allem gefühlsbetont, appellierend oder polemisch argumentiert (moralisch-emotional): Stimmungslagen werden wiedergegeben, subjektive Äusserungen von Individuen stehen im Zentrum und Emotionen werden ausgelöst, ohne Argumente abzuwägen oder zu vermitteln.
 - Es werden keine Argumente genannt (z. B. bei reinen Nachrichten der Fall).
- **Quellentransparenz: Text** (Stifterverein Medienqualität, 2020, S. 82)
 - Verwendete Quellen, Vorlagen bzw. Gesprächspartner*innen werden erwähnt und offengelegt.
 - Verwendete Quellen, Vorlagen werden erwähnt aber nicht offengelegt bzw. anonymisiert (z. B. bei Informant*innen im investigativen Journalismus der Fall).
 - Verwendete Quellen, Vorlagen bzw. Gesprächspartner*innen werden nicht erwähnt.
 - Es handelt sich um einen (persönlichen) Beitrag, der sich nicht auf andere Quellen oder Vorlagen stützt.

- **Quellentransparenz: Autorschaft** (Stifterverein Medienqualität, 2020, S. 82)
 - Die Autorschaft wird offengelegt (entweder voller Name, Kürzel, Redaktions- oder Agenturverweis).
 - Der Beitrag ist nicht gezeichnet.
- **Quellentransparenz: Bild (auch Grafik, Tabelle etc.)** (Dahinden & Hänggli, 2008, S. 10)
 - Quelle ist angegeben (bei mind. einem Bild)
 - Quelle ist nicht ersichtlich
 - kein Bild, Grafik, Tabelle, etc. im Bericht enthalten.
- **Trennung von Nachricht und Meinung (optisch/akustisch)** (Dahinden & Hänggli, 2008, S. 10)
 - nur Nachricht
 - nur Meinung, entsprechend gekennzeichnet (z. B. “Kommentar”)
 - nur Meinung ohne entsprechende Deklaration
 - Nachricht und Meinung gemischt und klar getrennt
 - Nachricht und Meinung ohne ersichtliche Trennung
- **Art der Aktualität (bezieht sich auf Hauptthema der Berichterstattung)** (Dahinden & Hänggli, 2008, S. 9)
 - zeitlich aktuell (Das Ereignis, der Anlass der Berichterstattung hat in der letzten Woche stattgefunden bzw. wird in der nächsten Woche stattfinden. Falls dies nicht der Fall ist, dann ‘nicht zuordenbar’ codieren.)
 - problemaktuell (Latent aktuelle Themen sind z. B. soziale Probleme, die eine Gesellschaft zwar durchgängig beschäftigen, für die aber immer erst neue Aktualität geschaffen werden muss.).
 - nicht zuordenbar
- **Relevanzebene (bezieht sich auf Hauptthema der Berichterstattung)** (Dahinden & Hänggli, 2008, S. 9)
 - gesamtgesellschaftlich (national, international)
 - gesellschaftliches Subsystem (kantonale/städtische Problematiken oder Sichtweisen von Bevölkerungsgruppen oder auch bestimmte Branchen)
 - institutionell (einzelne Organisationen, Behörden, Firmen, etc.)
 - individuell/privat
 - nicht feststellbar
- **Perspektivenvielfalt (bezieht sich auf Hauptthema der Berichterstattung)** (Dahinden & Hänggli, 2008, S. 10)
 - eine (Haupt-)Perspektive (nur Pro oder nur Kontra)
 - mehrere Perspektiven (Pro und Kontra)
 - nicht ersichtlich

3.2 Methodischer Auftrag: Vorbereitung und Koordination des Inter-Codierer-Reliabilitätstest

Grundsätzlich wird beim ICR-Test kontrolliert, ob zwei oder mehrere Codierer*innen in ihren Urteilen übereinstimmen und wenn ja, wie sehr dies der Fall ist. Rössler (2017, S. 208) stellt als Basis für die konkrete Durchführung des ICR-Tests die folgenden drei Fragen:

1. Was gilt als Übereinstimmung und was als Abweichung der Werte?
2. Wie berechnet sich der Grad der Übereinstimmung?
3. Wie viel Material muss für den Test bearbeitet werden?

3.2.1 Berechnung des Reliabilitätskoeffizienten

Angeichts der Vielzahl von Reliabilitätskoeffizienten ist es gar nicht so einfach herauszufinden, welcher davon für welches Projekt geeignet ist. Folgend wird zur Übersicht die Berechnung von drei verschiedenen Reliabilitätskoeffizienten beschrieben: der Übereinstimmungskoeffizient ($r_{\bar{U}}$), der Reliabilitätskoeffizient nach Holsti (r_H) und Krippendorff's Alpha (r_α).

Der **Übereinstimmungskoeffizient** wird aus dem Verhältnis von der Zahl aller übereinstimmenden Codierungen zu der Gesamtzahl aller Codierungen berechnet (Raupp & Vogelgesang, 2009). Zum Beispiel ergibt sich aus sechs Codierungen und vier Übereinstimmungen ein Reliabilitätswert ($r_{\bar{U}}$) von $4/6 = 0.67$, was 67% entspricht. Die Berechnung dieses Koeffizienten sagt lediglich aus, wie gut die Übereinstimmung der Codierung über die ganze Erhebung ist, sozusagen nach dem Prinzip des schwächsten Glieds der Kette (Raupp & Vogelgesang, 2009), macht aber keine Aussage über die einzelnen Codierer*innen oder die einzelnen Kategorien.

Rössler (2017, S. 212) empfiehlt deshalb den etwas anspruchsvolleren – und trotz seiner Schwächen – weit verbreiteten **Übereinstimmungsmass nach Holsti**. Mit diesem Mass wird die Zahl der übereinstimmenden Codierungen zur Gesamtzahl aller Codierungen von zwei Personen in Beziehung gesetzt (Paarvergleichs). Die Berechnung des Holsti-Koeffizienten erfolgt nach (Raupp & Vogelgesang, 2009, S. 16) in zwei Stufen: Zuerst werden alle möglichen Paarvergleiche berechnet und danach der Durchschnitt aller möglichen Paarvergleiche bestimmt. Die resultierenden Werte liegen dann als Koeffizient zwischen 0 und 1 und können somit als Übereinstimmung zwischen 0 und 100 Prozent ausgedrückt werden. Diese Formel muss für jede Kategorie einzeln und für jedes Paar von Codierer*innen angewendet werden. Obwohl die Verwendung des Holsti-Koeffizienten durch Rössler (2017, S. 212) empfohlen wird, wird die Verwendung aufgrund folgender Argumente auch stark kritisiert: Die Zahl der Aussprägungen beeinflusst die Höhe des Holsti-Koeffizienten und die Berechnung ist nur für nominalskalierte Variablen zulässig. Auch wird seine Präzision kritisiert, sowie die fehlende Korrektur bei zufälligen Übereinstimmungen des Codierens (Raupp & Vogelgesang, 2009, S. 16).

Aus dieser Kritik heraus wurde **Krippendorff's Alpha** entwickelt. Der Alpha-Koeffizient von Krippendorff, auch **K-Alpha** genannt, basiert auf der Annahme, dass die erwartete Übereinstimmung abhängig ist von der Häufigkeit der einzelnen Kategorien im gesamten Datensatz (Laps, 2017, S. 19). Im Gegensatz zu Holsti (oder auch Cohens Kappa etc., auf die hier nicht weiter eingegangen wird), der die paarweise Übereinstimmung berechnet, fokussiert der K-Alpha auf der paarweisen Abweichung. Ausserdem kommt K-Alpha mit fehlenden Werten sehr gut zurecht, was z. B. bei dem Holsti-Koeffizienten nicht der Fall ist, und kann bei kleinen Datensätzen und unter Verwendung von binären, nominalen und ordinalen Variablen eingesetzt werden.

3.2.2 Ergebnisdarstellung und Interpretation des Tests

Nachdem für je zwei Codierer*innen deren Übereinstimmung bezüglich jeder Kategorie mit ihren Variablen ermittelt wurde, lässt sich das Ergebnis des ICR-Tests anschliessend auf drei Ebenen analysieren:

1. **Nach Kategorien:** Mittelwert aller Koeffizienten pro Kategorie (über alle Codiervergleiche).
2. **Nach Codierer*innen:** Mittelwert aller Koeffizienten für eine*n Codierer*in (über alle Kategorien).
3. **Als Matrix:** Übereinstimmungen nach Kategorien und Codierer*innen.

	Codierer 1	Codierer 2	Codierer 3	Codierer 4	Mittelwert
Codierer 1		.64	.90	.86	.80
Codierer 2			.62	.58	.60
Codierer 3				.85	
Codierer 4			.85		
Mittelwert			.79	.76	.74

Abbildung 1: Matrix für Codierübereinstimmungen bei einer Variable (Früh, 2017, S. 184)

Einerseits entspricht eine perfekte Übereinstimmung 100% ($r=1$). Andererseits kommt Neuen-dorf nach Sichtung verschiedener Methodenliteratur zum Schluss, dass eine Übereinstimmung von über 90% als «sehr gut» und eine Übereinstimmung von über 80% als «gut» bezeichnet werden kann (zit. in. Raupp und Vogelgesang, 2009). Wo die Grenze unterhalb den 80% liegen soll, ist jedoch in der Methodenliteratur heftig diskutiert.

Bei der Interpretation von Reliabilitätswerten sollen die empfohlenen Schwellenwerte gemäss Raupp und Vogelgesang (2009, S. 14) stets in Beziehung zum Schwierigkeitsgrad einer Variable gesetzt werden. Ausserdem sollen nach Rössler (2017, S. 210) auch die formalen und die wichtigen inhaltlichen Kategorien getrennt aufgeführt werden. Weil bei formalen Kategorien eine hundertprozentige Übereinstimmung eigentlich schon fast Pflicht ist, aber bei den schwierigen inhaltlichen Kategorien auch geringere Werte zufriedenstellend sind.

Mögliche Richtwerte von Krippendorff's Alpha und deren Interpretation sind nachstehender Abbildung 2 zu entnehmen:

Wert	Interpretation
$\alpha > 0.8$	Starke Übereinstimmung
$\alpha = 0.67-0.8$	Schwache Übereinstimmung
$\alpha < 0.67$	Sehr schwache Übereinstimmung

Abbildung 2: Interpretation der Werte nach Berechnung des Krippendorff's Alpha (nach Krippendorff)

3.2.3 Ablauf des Inter-Codierer-Reliabilitätstest in diesem Projekt

Die Vorbereitung und Durchführung der beiden Tests werden – in Absprache mit den übrigen Arbeitsgruppen – gemäss folgendem Ablauf durchgeführt:

1. **AG 5** definiert was als Abweichung bzw. als Übereinstimmung gelten soll (Interpretation nach Krippendorff).
2. AG 4 erstellt ein Excel-Template für die Codierung: X-Achse für Artikel-ID, Y-Achse für Kategorien. *Resultat: 1 Template.*
3. AG 2 wählt 160 Artikel aus (80 Artikel für die Test-Codierung mit Codebuch V1 und erneut 80 für den Durchgang mit Codebuch V2).
4. AG 2 verteilt die Artikel per Zufallsprinzip auf die jeweiligen Codierer*innen, so dass jede*r Codierer*in Artikel von unterschiedlichen Plattformen codiert. *Resultat: 20 Excel-Sheets mit den Testcodierungen.*
5. Jede*r Codierer*in erhält ein Excel-Sheet und beschafft sich die Artikel (PDF, online oder anderes vorliegendes Format), die zugewiesen wurden anhand der ID aus der Datenablage und codiert diese im „eigenen“ Excel-Sheet.
6. **AG 5** überträgt die Daten aus den 20 Excel-Sheets in die Excel-Sheets für den ICR-Test. *Resultat: Pro Kategorie gibt es ein Excel-Sheet, in dem die Matrix der Artikel und die Auswahl der Codierer*innen abgebildet ist. (ev. auch pro Codierer*in? siehe Fragen am Ende Dokuments)*
7. **AG 5** führt den ICR-Test nach K-Alpha durch.
8. **AG 5** überreicht die Resultate aus dem ICR-Test an AG 4 und erklärt, wie die Werte zu interpretieren sind.
9. AG 4 meldet an alle Gruppen, welche Teilcodebücher erstellt haben, wo Änderungen notwendig sind und optimiert die Codieranweisung. Gegebenenfalls initiiert sie ausserdem eine zweite Codierschulung für ausgewählte Personen.
10. Alle Gruppen passen ihre Teilcodebücher gemäss Rückmeldung an und senden die optimierte Version an AG 4.
11. AG 4 erstellt Codebuch V2.
12. Für alle Gruppen findet eine zweite Runde Test-Codierungen statt.

13. **AG 5** wendet zum zweiten Mal den ICR-Test nach K-Alpha an.
14. **AG 5** interpretiert die Resultate des Tests nur noch im Rahmen des Abschlussberichts (Rückmeldung an AG 4 erfolgt nur nach Bedarf bzw. Wunsch ihrerseits hin).

4 Zielsetzung

4.1 Inhaltlicher Auftrag: Operationalisierung auf Statement-Ebene

In Anlehnung an die von Rössler (2017, S. 105) formulierten Anforderungen an Kategorien sollte unsere Auswahl für die Analyseeinheit «Statements» inkl. deren Ausprägungen trennscharf und vollständig sein. Dadurch soll verhindert werden, dass mit der Verwendung dieses Instruments fehlerhafte Messungen generiert werden. Des Weiteren sollen treffende Definitionen sowie gut gewählte Beispiele dazu führen, dass zwischen den Codierer*innen eine hohe Übereinstimmung und somit eine objektive Genauigkeit des Codebuchs erreicht wird. Wir gehen davon aus: Je präziser und umfassender die Kriterien im Codebuch V1 vorliegen, desto besser sind die Ergebnisse des Reliabilitätstests zu erwarten – und damit einhergehend; desto kleiner fällt der Aufwand für die Überarbeitung des Buchs für die AG 4 aus.

4.2 Methodischer Auftrag: Vorbereitung und Koordination des Inter-Codierer-Reliabilitätstests

Grundsätzlich hoffen wir, dass die gewählte Methode für die Prüfung der Datensätze in Verbindung mit dem Codebuch V1 erfolgreich angewendet werden kann. Der durch uns zu entwickelnde und durchzuführende Test soll anhand des begutachteten Materials aufzeigen können, welche Kategorien und Variablen im Codebuch V1 keine genügend hohe Übereinstimmung aufweisen. Dadurch wird einerseits ermöglicht, das Potential des erarbeiteten Messinstruments zu bewerten und andererseits jene Massnahmen zu erkennen, die zu einer optimierten Version des Codebuchs führen. Schliesslich wäre es wünschenswert, dass das hier erprobte Vorgehen – sofern es sich bewährt – auch im SNF-Projekt Anwendung finden kann.

5 Zeitplan, Schnittstellen, Abhängigkeiten

5.1 Zeitplanung und Arbeitspakete

Der Zeitplan wird in vorliegendem Projekt durch die Projektleitung erstellt und muss zwingend von allen AGs als verbindlich angesehen werden. Er hängt einerseits stark davon ab, wie lange die AG 2 und 3 für die Datensammlung und das Datenmanagement benötigen und andererseits wie viel Zeit der AG 4 gegeben wird, um die finale Version 2 des Codebuchs zu erstellen. Wir plädieren stark dafür, genug Zeit für die Durchführung des Tests sowie die anschließende Bereinigung und Optimierung des Codebuchs in seine finale Version einzukalkulieren. An dieser Stelle wird lediglich auf die Zeitplanung der PL, die mittels einem Gantt Chart dargestellt ist und sich über das gesamte Projekt spannt, verwiesen um Redundanzen zu vermeiden. Die Tabelle 1 gibt Auskunft über die Arbeitspakete und Meilensteine (MS) der AG 5.

Arbeitspakete	Deadlines
Konzept Version 1	21.09.2020
Recherche Inter-Codierer-Reliabilitätstest	12.10.2020
Recherche Operationalisierung	12.10.2020
MS: Konzept Version 2	12.10.2020
Präsentation Konzept Version 2	14.10.2020
MS: Operationalisierung Analyseinheit Statement	20.10.2020
Durchführung Codierschulung für Teilcodebuch Statement	03.11.2020
Codieren mit Codebuch Version 1	10.11.2020
MS: Datenbereinigung und Berechnung ICR-Test Version 1	15.11.2020
Verbesserung des eigenen Codebuch-Teils (Statement)	05.12.2020
Codieren mit Codebuch Version 2	15.12.2020
Datenbereinigung und Berechnung ICR-Test Version 2	20.12.2020
MS: Projektbericht schreiben	05.01.2021
Abschlusspräsentation	15.01.2021

Tabelle 1: Arbeitspakete und Meilensteine AG 5

5.2 Schnittstellen und Abhängigkeiten

Nachstehend wird lediglich auf diejenigen Abhängigkeiten und Schnittstellen, die als am relevantesten erachtet werden, kurz eingegangen. Das von der Projektleitung erstellte Gantt stellt weitere Schnittstellen übersichtlich dar.

- **AG 1-5 (Alle)** Da das Codebuch ein gemeinsames Produkt ist, bildet es als ganzes eine konzeptuelle Schnittstelle. So gilt es, die Vorschläge der Leistungsindikatoren aufeinander abzustimmen, wobei die definitive Auswahl durch die PL erfolgt.
- **AG 2 (Rohdatensammlung):** Das Sammeln der Beiträge für die Inhaltsanalyse ist Basis für das Durchführen der Codierarbeit.
- **AG 4 (Codebuchentwicklung):** Um den ICR machen zu können, muss die erste Version des Codebuchs vorliegen.
- **AG 3-5 (die Operationalisierenden):** Jedes Mitglied der vier Codier-Arbeitsgruppen muss das Codieren der jeweils 10 Artikel fristgerecht erledigen, damit der Test durchgeführt werden kann.

Weitere Schnittstellen und Abhängigkeiten werden sich nach der Vorstellung der Konzepte der einzelnen Arbeitsgruppen zeigen.

5.3 Risiken

Das Projekt lebt von der Mitarbeit und dem Engagement von knapp 20 Personen, was personelle Risikofaktoren mit sich bringt. Wie die grob skizzierten Abhängigkeiten zeigen, würde ein nicht eingehaltener Zeitplan einer anderen Arbeitsgruppe die Erledigung der eigenen Aufgaben stark beeinträchtigen. Verstärkt wird das für unsere Arbeitsgruppe insbesondere dadurch, weil sich jene Arbeitsschritte, die wir koordinieren, eher am Ende des Projekts befinden. Zudem besteht das Risiko, dass Lieferobjekte, die nicht den Anforderungen entsprechen, die Qualität der weiteren Arbeit damit unumgänglich beeinflussen. Auch zu erwähnen ist die Schnittstelle 'Codebuch', die das Mitwirken aller Projektmitglieder voraussetzt und viel Koordination und Kontrolle verlangt.

6 Fragen an Betreuerin

- Welche Werte wollen wir für die Interpretation festlegen? Entscheiden wir das als ganze Projektgruppe oder önnen wir einfach diese auf Basis der Tabelle von Krippendorf nehmen?
- Müssen wir die Ergebnisse des Inter-Codierer-Reliabilitätstests anschliessend auf allen drei Ebenen analysieren (nach Kategorien, nach Codierer'innen und als Matrix)?

Literatur

- Brosius, H.-B., Haas, A. & Koschel, F. (2016). Inhaltsanalyse I: Grundlagen. In H.-B. Brosius, A. Haas & F. Koschel (Hrsg.), *Methoden der empirischen Kommunikationsforschung: Eine Einführung* (S. 137–152). Studienbücher zur Kommunikations- und Medienwissenschaft. doi:10.1007/978-3-531-19996-2_8
- Dahinden, U. & Dalmus, C. (2020, September). Projektkurs Lokaljournalismus - Folien Unterricht. FH Graubünden.
- Dahinden, U. & Hänggli, R. (2008, 15. Oktober). Codebuch zur Vorlage: Einbürgerungsinitiative (Abstimmung vom 01.06.2008).
- Früh, W. (2017). *Inhaltsanalyse* (9. Aufl.). Konstanz, München: UVK.
- Laps, D. (2017). *Interrater-Reliabilitätskoeffizienten: Analysen und Simulationen* (Diss., Universität Düsseldorf).
- Raupp, J. & Vogelgesang, J. (2009). *Medienresonanzanalyse. Eine Einführung in Theorie und Praxis* (1. Auflage).
- Rössler, P. (2017). *Inhaltsanalyse* (3., völlig überarbeitete Auflage). UTB basics. Konstanz: UVK.
- Medienqualitätsranking 2020. (2020). Zugriff am unter <https://www.mqr-schweiz.ch/files/mqr/pdf/MQR-20.pdf>