

Die Qualität der Radio- und Fernsehforschung in der Schweiz

Management Summary zum Jahresbericht 2018

der Medienwissenschaftlichen Kommission
zuhanden des Verwaltungsrates der Mediapulse AG

Berlin/Leipzig/Zürich, im März 2019

Vorgelegt von
Roger Berger, Universität Leipzig
Benjamin Fretwurst, Universität Zürich
Joachim Trebbe, Freie Universität Berlin

Inhalt

1	Zusammenfassung / Management Summary	3
2	Establishment Survey (ES)	4
3	Fernsehreichweitenforschung.....	7
4	Radioreichweitenforschung	12
5	Schlussfolgerungen und Empfehlungen.....	15

1 Zusammenfassung / Management Summary

Im Jahr 2018 fand in der Medienwissenschaftlichen Kommission (MWK) der Mediapulse ein Generationswechsel statt. Bereits im Sommer 2017 wurde Joachim Trebbe, Professor für Forschungsmethoden und Medienanalyse an der Freien Universität Berlin, vom Verwaltungsrat der Mediapulse als zusätzliches Mitglied in die Medienwissenschaftliche Kommission der Mediapulse AG berufen und mit einem Zwischenaudit der neu konzeptionierten Radioforschung beauftragt. Er übernahm dann im Sommer 2018 den Sitz von Michael Schenk. Ebenfalls im Jahr 2018 wurden dann Roger Berger, Professor für Soziologie und Methodenlehre an der Universität Leipzig und Benjamin Fretwurst, Wissenschaftlicher Mitarbeiter am Institut für Kommunikationswissenschaft und Medienforschung der Universität Zürich als neue Mitglieder in die Medienwissenschaftliche Kommission der Mediapulse berufen. Sie traten damit die Nachfolge von Heinz Bonfadelli und Walter Hättenschwiler an.

Die neuen Mitglieder der MWK sind sich der grossen Fussstapfen, in die sie mit ihrem ersten Bericht über die Aktivitäten der Mediapulse im Jahr 2018 treten, bewusst. Die vorliegenden Berichte der Kommission waren stets gründlich, objektiv, transparent und kritisch. Sie haben die Entwicklung der angewandten Radio- und Fernsehforschung in der Schweiz über lange Jahre konstruktiv und akribisch begleitet. Für die neuen Mitglieder der MWK ist diese Tradition Auftrag und Herausforderung. Sie hat sich vorgenommen, die hohen methodischen Standards der Nutzungs- und Reichweiterehebungen, die von der Mediapulse AG verantwortet werden, auch in der Zukunft wissenschaftlich zu begleiten, objektiv zu evaluieren und transparent zu kommunizieren.

Den Auftakt dazu liefert dieser Bericht. Er beschäftigt sich – wie gehabt – mit den drei grundlegenden Studien, die für die Erhebung von Mediennutzungsdaten der elektronischen Medien in der Schweiz durchgeführt werden: 1. Dem Establishment Survey (ES), einer Rekrutierungs- und Befragungsstudie für die Bereitstellung des Fernsehpanels und grosser Teile des Radiopanel, 2. der Fernsehforschung für die Erhebung der TV-Reichweiten und Marktanteile sowie 3. der Radioforschung, der Längsschnittstudie für die tagesgenaue und tagesaktuelle Messung der Hörfunknutzung.

Die Aufgabe der MWK ist es, den Nutzern der Daten, die von der Mediapulse AG bereitgestellt werden, die Gewähr zu bieten, dass diese Daten einwandfrei, d.h. im Hinblick auf ihre methodische Qualität erhoben, aufbereitet und publiziert werden. Da die drei Studien im Auftrag der Mediapulse von verschiedenen Instituten durchgeführt werden, bezieht sich ein grosser Teil der Aussagen in diesem Bericht auf Dokumentationen und Kennziffern, die als Leistungsindikatoren zur Qualitätskontrolle zwischen diesen vertraglich vereinbart wurden. Zusätzlich wurden der MWK durch die Ansprechpartner bei Mediapulse und in den Instituten weitere, konzeptionelle und die jeweils aktuelle Erhebung betreffende Daten und Dokumente zur Verfügung gestellt, die einen genaueren Blick auf die Durchführung der Studien ermöglichten.

Der Bericht in seiner Langfassung, der im Folgenden in seinen wesentlichen Schlussfolgerungen zusammengefasst werden soll, enthält beschreibende und bewertende Aussagen.

Die Kommission hat sich darum bemüht, methodisch-konzeptionelle Sachverhalte so darzustellen, dass Sie für die Nutzer der Studien und ihrer Daten nachvollziehbar und für eine Qualitätsbeurteilung hilfreich sind. In diesem Zusammenhang sind die neuen Mitglieder der MWK mit vielen Details und Operationalisierungsentscheidungen konfrontiert worden, die zum Teil das Ergebnis längerer Entwicklungs- und Optimierungsprozesse zwischen der Geschäftsleitung der Mediapulse, der Forschungskommission und den operativen Verantwortlichen in den Instituten sind. Der Bericht erfüllt aus dieser Perspektive nicht den Anspruch, alle diese Entscheidungen auf den Prüfstand zu stellen und somit vollständig im Sinne einer kompletten Durchleuchtung aller methodischen Details zu sein.

Das Ziel der Kommission war es, in diesem ersten Bericht nach dem Generationswechsel die wesentlichen methodologischen und methodischen Grundlagen und Kennziffern zu verstehen, zu beschreiben und im Hinblick auf ihre Bedeutung für die wissenschaftliche Qualität der erhobenen Reichweitendaten zu beurteilen. Die Ergebnisse dieser Prüfung werden auf den folgenden Seiten zusammengefasst.

2 Establishment Survey (ES)

Mit dem ES werden zwei Ziele verfolgt: Einerseits sollen damit die Haushalte in der Schweiz und ihre technische Ausstattung inklusive Empfangswege beschrieben werden. Andererseits dienen die Respondenten des ES als Pool von potentiellen Teilnehmern für den Aufbau und vor allem den Betrieb des Fernsehpanels. Für beide Ziele sind der Übergang vom zugrundeliegenden Universum zur gezogenen Stichprobe des ES und die damit verbundenen methodischen Elemente von zentraler Bedeutung.

2.1.1 Grundgesamtheit und Stichprobe

Die Grundgesamtheit des ES sind die Privathaushalte der Schweiz in denen mindestens eine Person eine Landessprache spricht.

Die Stichprobenziehung geschieht zweistufig über die Auswahl von Adressen und daran anschliessend über die zufällige Auswahl von Haushalten an den Adressen, falls an einer Adresse mehr als ein Haushalt vorliegt. Alle Haushalte in denen sich mindestens eine Person über 14 Jahren in einer der Amtssprachen genügend verständigen kann um an der Fragebogenerhebung teilzunehmen, bilden dann die bereinigte Bruttostichprobe. Die Nettostichprobe des ES wird dann durch diejenigen Haushalte gebildet, die an der Befragung teilnehmen. Erhoben werden dabei die Haushalts- und Empfangsstruktur, sowie Angaben zu allen Haushaltsmitgliedern. Der Quotient zwischen Netto- und bereinigter Bruttostichprobe ergibt dann die Ausschöpfung.

Das Konzept, die Probanden der Radio- Fernsehpanel nicht direkt zu rekrutieren, sondern dies über den vorgängig angelegten Pool des ES zu tun, ist methodisch sinnvoll. Bei einer

direkten Rekrutierung nach Massgabe der jeweiligen Panelmortalität wäre die Gefahr, dass es zu systematischen Verzerrungen auf Grund von nicht korrekt möglichen und/oder durchgeführten Nachrekrutierungen kommen könnte, wesentlich höher.

Das Stichprobendesign des ES ist methodisch sinnvoll. Die Ziehung der Stichprobe über das aktualisierte Adressverzeichnis der Post und im zweiten Schritt der zufälligen Auswahl von Haushalten stellt die beste Möglichkeit dar, um die ständige Wohnbevölkerung der Schweiz über 14 abzubilden. Ein anderer Zugang, z.B. über Telefonverzeichnisse, wäre weit weniger erfolgsversprechend und von Anfang an von systematischen Verzerrungen begleitet. Dies zeigt sich z.B. bereits daran, dass 2018 nur gut 77 Prozent der Haushalte in der ES-Stichprobe telefonisch kontaktiert werden konnten.

Von den anfänglich vorhandenen 30'818 Adressen in der Ausgangsbruttostichprobe blieben nach dieser Bereinigung 27'460 übrig. Die bereinigte Bruttostichprobe entspricht damit 89,1 Prozent der Ausgangsadressbruttostichprobe. Dieser Anteil fällt für alle drei Sprachregionen fast identisch aus.

Insgesamt wurden 2018 12'395 Interviews zum ES realisiert. Davon waren 8297 telefonische und 3496 postalische Interviews (inkl. 602 Onlineinterviews). Dies ergibt für die telefonische Befragung eine Ausschöpfungsquote von 49,8 Prozent und für den postalischen Survey eine solche von 74,0 Prozent. Rechnet man die oben als stichprobeneutrale Ausfälle bezeichneten Fälle raus, erhöhten sich diese Werte auf 55,4 Prozent (CATI) und 76,6 Prozent (PAPI inkl. Online) für die bereinigte Ausschöpfungsquote.

2.1.2 Datenerhebung

Der Fragebogen des ES liegt deutsch, französisch und italienisch vor. Der Fragebogen ist kurz. Abgefragt werden die Struktur des Radio- und Fernsehempfangs (Geräte, Empfangswege) und die Sozialstatistik der Haushalte. Erwartungsgemäss ergaben sich Probleme bei der sozial heiklen Frage nach dem Haushaltseinkommen. Obschon nur um Antworten in groben Kategorien gebeten wurde, ergaben sich hier ca. 22 Prozent fehlende Antworten. In der postalischen Befragung ergaben sich Probleme bei komplexeren, technischen Fragen, die es zum Teil erfordern würden, bei Geräten, in Verträgen, Betriebsanleitungen, etc. nachzuschauen. Bei Fragen danach, ob etwa ein Fernseher internetfähig oder eine Spielkonsole vorhanden ist (und Ähnlichem), ergaben sich Ausfallraten im Bereich von 10 bis 20 Prozent.

Unter methodischen Gesichtspunkt sind bei der Befragung im ES zwei Perspektiven von Interesse: (1) Die Gültigkeit der Antworten jedes Befragten und (2) die Gültigkeit der Antworten je Frage. Im ES wird für die erste Perspektive aus 11 Indikatoren ein Summenindex für die Validität der Antworten jedes Befragten gebildet. Der Wert 0 (keine Auffälligkeit) trifft für über 70 Prozent der Fälle zu. Über 97 Prozent aller Fälle haben maximal zwei Auffälligkeiten. Der empirische Maximalwert beträgt jeweils 7. Aus der zweiten Perspektive problematisch sind die Fragen zur Empfangsstruktur bei der postalischen Befragung. Hier ist der Ausfall nicht vernachlässigbar, zumal er systematisch höher ist bei älteren, bei weniger gebildeten Befragten und bei Frauen. Offenbar ist der Ausfall auch bei Personen höher, die

wenig technikaffin sind. Dies bedeutet, dass die Empfangsstruktur der Population vermutlich durch den ES verzerrt dargestellt wird. Dadurch, dass nur etwa ein Drittel der Befragten postalisch befragt wird, hält sich diese Verzerrung jedoch in Grenzen. Weiterhin wird bei den – aus dem ES rekrutierten – Panel-Teilnehmern die Empfangsstruktur noch mal präzise gemessen, was es ermöglicht, eventuelle Verzerrungen abzuschätzen.

2.1.3 Datengewichtung

Gewichtung von Daten kann aus zwei Gründen geschehen. Zum einen kann ein bereits im Design festgelegtes Ungleichgewicht bei der Erhebung korrigiert werden, um für deskriptive Zwecke die Ausgangsproportionen herzustellen («Designgewichte»). Zum anderen kann die Stichprobe im Nachhinein nach Strukturmerkmalen der Grundgesamtheit («Poststratifizierung») gewichtet werden, um die Verteilung einer bestimmten Gewichtungsdimension der entsprechenden empirischen Relation in der Population anzugleichen.

Da im ES durch die disproportionale Schichtung die Haushalte in deutschsprachigen Regionen eine unterproportionale, diejenigen in französisch- und italienischsprachigen Gebieten hingegen eine überproportionale Wahrscheinlichkeit haben, in die Stichprobe zu kommen, müssen diese Disproportionalitäten korrigiert werden. Diese Gewichtung erfolgt nicht entlang der Sprachregionen, sondern nach den fünf WEMF-Grossregionen Suisse Romande, Alpen und Voralpen, Westmittelland, Ostmittelland, Svizzera Italiana. Insofern handelt es sich hier genau genommen um mehr als ein Designgewicht, da gleichzeitig eine Poststratifizierung der Stichprobe nach bekannten statistischen Parametern vorgenommen wird.

Zwei weitere Strukturgewichte werden zur Poststratifizierung angewendet: die Agglomerationsgrössenklassen «mehr als 100'000», «50'000 bis 100'000», «10'000 bis 49'999 Einwohner», «Unter 10'000» und die Haushaltsgrösse mit 1-, bis 4-, sowie 5- und Mehrpersonenhaushalten. Als Populationsparameter werden die Angaben des BfS/STATPOP verwendet.

Im ES wird ein Gewicht pro Erhebungswelle und eines pro Jahr berechnet. Die Verteilungen der beiden Gewichtungsvariablen unterscheiden sich kaum. Auch die Verteilungen in den beiden Wellen sind sich sehr ähnlich. Das maximale Gewicht variiert zwischen 2,4 und 2,7. Das minimale Gewicht liegt immer bei etwa 0,18. Die Standardabweichung beträgt jeweils etwa 0,5. Mit dem Ziel der Hochrechnung von TV- und Radioreichweiten ist eine Gewichtung unumgänglich. Da der ES die Grundlage für die Panels bildet, muss auch hier gewichtet werden. Auch wenn die Gewichtung also zu höheren Fehlerstreuungen der Kennwert führt, wird sie von der MWK nicht in Frage gestellt. Die Berechnungsalgorithmen sind darauf ausgerichtet, die Gewichte möglichst klein zu halten. Das gelingt auch aufgrund des grossen Stichprobenaufwands.

2.1.4 Dokumentation

M.I.S. Trend veröffentlicht pro Erhebungswelle Methodenberichte zum ES. Die MWK regt an, die Ausfälle auf allen Ebenen der Stichprobenziehung zukünftig standardisiert zu dokumentieren und alle geografischen Parameter im Format des Geoinformationssystems (GIS) aufzubereiten, um einen schnelleren Zugriff und ein leichteres Verständnis der methodischen Details zu erlauben.

2.1.5 Fazit

Insgesamt wird der Establishment Survey als eine methodisch korrekte und zuverlässige Basis für die Beschreibung der Haushaltsstruktur, der technischen Ausstattung und der Empfangssituation, sowie für die Rekrutierung der Paneluntersuchungen der Medienforschung in der Schweiz angesehen. Die Stichprobenziehung entspricht dem Stand der empirischen Sozialforschung und sorgt für eine repräsentative Abbildung der angestrebten Grundgesamtheiten in den statistischen Fehlermargen. Die Befragungsqualität ist hoch, sowohl in der telefonischen als auch in der postalischen Befragung. Insbesondere der erhebliche Aufwand im Rahmen der Kontaktversuche wird von der MWK anerkannt.

Der ES ist für die Qualität der Medienforschung der Mediapulse von entscheidender Bedeutung und als vorgeschaltete, repräsentative Stichprobenerhebung eine wichtige Voraussetzung für die Qualität der Radio- und Fernsehpanel.

3 Fernsehreichweitenforschung

Das Ziel der Reichweitenmessung ist die Abbildung der TV-Nutzung in der Schweiz im Jahr 2018. Diese Nutzung wird anhand des Mediapulse-Fernsehpanels gemessen und auf die Gesamtheit der Nutzungsvorgänge hochgerechnet.

3.1.1 Grundgesamtheit und Stichprobe

Die private TV-Nutzung wird zu jedem sekundengenauen Zeitpunkt für die Grundgesamtheit der Personen geschätzt, die 1. zur ständigen Wohnbevölkerung in der Schweiz zählen, 2. in einem Privathaushalt leben in dem mindestens einem TV-fähigen Gerät vorhanden ist und 3. mindestens drei Jahre alt sind.

Die soziodemografischen Schätzungen stammen vom Bundesamt für Statistik (BFS) und werden jeweils aus den vergangenen Jahren für das Berichtsjahr extrapoliert. Die TV-Dichte wird seit 2018 aufgrund des Establishment Surveys (ES mit ca. 30'000 Haushalten der letzten fünf Wellen) geschätzt. Für das Jahr 2018 wurde die Grundgesamtheit der Wohnbevölkerung ab drei Jahren auf 8'291'000 geschätzt. Davon leben 92 Prozent in Fernsehhaushalten. Aufgrund dieser Berechnungen wurde die Grundgesamtheit korrekt auf 7'365'000 Personen

festgelegt (Deutschschweiz 5'186'000, Suisse Romande 1'840'00, Svizzera Italiana 339'000). Aus der Grundgesamtheit wurde eine über die drei Sprachregionen disproportionale Quotenstichprobe gezogen. Die Stichproben von 2057 Haushalten (1100 in der Deutschschweiz, 660 Haushalte in der Romandie und 297 in der italienischen Schweiz) sind im Umfang hinreichend für die angestrebte Genauigkeit.

Als Quoten sind Merkmale aus den Daten des STATPOP (Haushaltsgrösse und Kinder im Haushalt) und des ES (TV-Dichte und Empfangsweg) herangezogen worden, die mit den Fernsehreichweiten zusammenhängen. So wird zum Beispiel sichergestellt, dass nicht zufällig deutlich zu viele oder deutlich zu wenige Haushalte mit Kindern in das Panel gelangen. Da Kinder ein besonderes Fernsehnutzungsverhalten aufweisen, könnte eine durch extremen Zufall verzerrte Stichprobe hier zu verzerrten Anteilsschätzungen führen. Andererseits garantiert die Quotierung keine unverzerrte Erhebung der Merkmale, die nicht stark mit den Quotenmerkmalen zusammenhängen. Für die Rekrutierung des Panels werden keine Haushalte systematisch ausgeschlossen oder bevorzugt. Daher bewertet die MWK die Stichprobenziehung als hinreichend äquivalent zu einer stratifizierten Zufallsstichprobe.

Für die Repräsentativität der Reichweitenschätzung ist die Bedeutung von Ausfällen durch Teilnahmeverweigerungen schwer abschätzbar. Für die Panelauffrischung 2018 konnte von 2205 Haushalten aus dem Establishment Survey eine Teilnahmeentscheidung gewonnen werden. Davon haben 84 Prozent (2063) der Haushalte eine Teilnahme abgelehnt. 16 Prozent der Haushalte (in beiden Semestern etwa gleich) konnten schliesslich für die Teilnahme am Fernsehpanel rekrutiert werden. Die Teilnahmeverweigerung lag 2018 also bei 84 Prozent – ein erklärungsbedürftig hoher Wert. Immerhin hatten alle 2205 Personen eingewilligt, am ES teilzunehmen. Die Ablehnung betrifft also nicht Befragungen generell, sondern die Teilnahme am Fernsehpanel. Die MWK empfiehlt hier statistische Analysen, bei denen Personeneigenschaften und nutzungsrelevante Variablen des ES herangezogen werden, um die Teilnahmeverweigerungen zu erklären. Unabhängig vom Ausmass systematischer Teilnahmeverweigerungen gibt es keine bekannten Gründe von einem bedeutenden Zusammenhang mit der Nutzung bestimmter Fernsehsender und -inhalte. Bis auf Weiteres kann deshalb angenommen werden, dass die Gleichbehandlung der Anbieter von den Teilnahmeverweigerungen nicht verletzt wird.

Die Quotenmerkmale TV-Dichte und der Empfangsweg werden anhand des ES geschätzt. Die Stichprobe wird nach Sprachregionen und 13 Konzessionsgebieten gezogen. Daraus ergeben sich 25 Elementarzonen (EZ). Die Haushalte wurden proportional zu den Bevölkerungsanteilen in diesen EZ gezogen. In kleineren EZ sind daher die Fallzahlen sehr klein und in grösseren EZ unnötig gross. Eine disproportionale Stichprobe wäre vermutlich mit grösserem Aufwand verbunden, könnte aber im ES mit entsprechend höheren Fallzahlen vorbereitet werden. Die Vorgabe an Mediapulse, die Anteile der feingliedrigen EZ proportional zu ziehen, hat zur Folge, dass in mehreren EZ die Soll-Fallzahlen unter 30 Haushalten liegen und damit unter der für Inferenzstatistiken notwendigen Fallzahl. Die Personenzahlen liegen in den meisten EZ über 30, aber das TV-Nutzungsverhalten der Personen eines Haushalts stellt keine unabhängigen Fälle dar. Darum ist die Zahl der Haushalte entscheidend. Dieses Prob-

lem ist also nur durch eine Designentscheidung lösbar: Entweder wird (a) das proportionale Panel auf das Dreifache vergrößert, (b) die Proportionalität für die kleinen EZ aufgegeben und Boosts auf mindestens 30 Haushalte vorgegeben, oder (c) die kleinen EZ werden mit angrenzenden EZ zusammengelegt. Gewichtung ist hier keine Lösung, da sie die Voraussetzungen nicht herstellen kann.

3.1.2 Panelmessung und Datenerhebung

Ausfälle im TV-Panel werden kontinuierlich durch Nachrekrutierungen aufgefrischt. Ab dem Jahr 2018 wurden dafür engere Toleranzgrenzen vereinbart, d.h. es wird schneller nachrekrutiert. Diese Massnahme erhöht die Qualität und die Transparenz der Panelpflege.

Haushalte mit sehr unplausiblen Nutzungsverhalten werden aus dem Panel ausgeschlossen. Diese Massnahme führt zu einer Normalisierung des Panels, da solch extreme Nutzungsmuster als singuläre Ausreisser betrachtet werden. Da das Verhalten der Panelhaushalte auf die Gesamtbevölkerung hochgerechnet wird, würde durch die zufällig ins Panel gelangten Extremhaushalte dieses Verhalten in der Gesamtbevölkerung deutlich überschätzt werden. Die Regeln zum nutzungsbedingten Ausschluss von Haushalten aus dem Panel sind im Sinne der Repräsentativität zu begrüssen.

Einkaufsgutscheine als Incentives für die Teilnahme und die Aufklärung von Fehlermeldungen sind aus Sicht der MWK eine gute Massnahme, die nach wissenschaftlichem Erkenntnisstand keinen verzerrenden Effekt auf das Nutzungsverhalten hat. Wenn möglich sollte die Bindung der Gutscheine an einzelne Detailhändler langfristig gelöst werden.

Technisch wird die Fernsehnutzung seit 2013 mit dem PeopleMeter 5000 gemessen. Die Messung ist also nicht von Interviewern abhängig und daher von hoher Objektivität. An jedes TV-Geräte und seine Peripherie ist ein PeopleMeter angeschlossen. Da zur Fernsehnutzung auch Computer genutzt werden können, wird die Fernsehnutzung über ein auf den Rechnern installiertes Virtual Meter erfasst. In der Deutschschweiz sind 71 Prozent, in der Romandie 63 Prozent und in der italienischen Schweiz 70 Prozent mit mindestens einem Virtual Meter ausgestattet.

Die Erhebung des Fernsehverhaltens erfolgt durch den (IP-basierten) Abgleich der ausgestrahlten TV-Audiosignale und dem Tonausgang am Fernsehgerät (Audiomatching). So kann sehr genau identifiziert werden, welches Programm zu welchem Zeitpunkt im Haushalt genutzt wird. Technische Ausfälle im Referenzierungssystem werden jeweils dokumentiert, fallen aber für das Jahr 2018 kaum ins Gewicht.

Bisher wird die TV-Nutzung auf Smartphones, Spielekonsolen und Tablets nicht abgedeckt. In Zukunft wird diese Leerstelle durch die Installation eines Router Meter adressiert. Diese können dann am Internetrouter die Fernsehnutzungsdaten automatisch erheben.

Uncovered Viewing ist Fernsehnutzung ohne Anmeldung mindestens eines Haushaltsmitglieds. 2018 betrug diese fehlende Zuordnung 4,2 Prozent. Die höchsten Wochendurchschnitte von 5,1 Prozent wurden im April und Juni gemessen. In Haushalten, mit hohem nicht zuordenbarem Viewing, wurde ein BeepMeter installiert, der die Panelteilnehmer beim

Start des TV an die Anmeldung mit der Fernbedienung erinnert. Diese Massnahme hält die MWK für sinnvoll.

Zur Überprüfung der Anmeldedisziplin werden einmal im Jahr alle Haushalte angerufen und nach ihrer augenblicklichen Nutzungssituation im Haushalt befragt. Die Auskünfte aus diesen Koinzidenzinterviews werden dann mit den technischen Messdaten verglichen. Insgesamt kann festgehalten werden, dass in mehr als 10 Prozent der Koinzidenzinterviews die Nutzungszahlen vom System unterschätzt wurden. Aus Sicht der MWK wäre es wünschenswert, wenn dies für die Fälle gesondert ausgewiesen wird, die eine angemeldete Person hatten, aber weitere Personen nicht korrekt angemeldet waren. Denn diese Fälle werden im laufenden Betrieb nicht durch die Uncovered Viewings erfasst.

3.1.3 Gewichtung und Datenaufbereitung

Für die Beurteilung der Gewichtung muss zwischen Hochrechnungsfaktor, Design- und Strukturgewichtung unterschieden werden. Der Hochrechnungsfaktor ist als Projektionsfaktor (wie Prozentuierung) eine reine Darstellungsfrage, bei der gedacht auf die absoluten Bevölkerungszahlen projiziert wird. Die Hochrechnung projizierte 2018 mit der Konstante von 1649 auf 7'365'000 Personen. Die Designgewichte korrigieren das disproportionale Stichprobendesign, sind also für alle Fälle innerhalb einer Sprachregion gleich und daher unter Qualitätsgesichtspunkten eher unproblematisch.

Strukturgewichte spielen für die Qualität der Daten hingegen eine grosse Rolle. Sie geben Personen der Stichprobe bei der Zählung die Bedeutung von mehreren natürlichen Personen (z.B. bei einem Faktor 3, das Dreifache) und anderen Fällen nur einen Bruchteil der Bedeutung einer natürlichen Person (z.B. bei einem Faktor von 0,25 ein Viertel). Strukturgewichte korrigieren so die Verteilungsunterschiede zwischen Stichprobe und Grundgesamtheit. Das gilt beim Fernsehpanel für die bekannten Verteilungen von Geschlecht, Alter, Bildung oder Haushaltsgrössen, die den Registererhebungen des STATPOP (BfS) entnommen werden und für die TV-Dichte und die Empfangswege aus dem ES. Strukturgewichte verbessern die Messung von Fernsehreichweiten nur dann, wenn die Fernsehnutzung stark mit den Gewichtsmerkmalen zusammenhängt. Die Korrelationen zwischen den Quotierungsmerkmalen und den zu untersuchenden Merkmalen sind allerdings eher klein. Die Verschlechterung der Kennwertschätzung (abgebildet in Instar durch grössere Standardfehler der Facts) ist der Preis für die Validität der Hochrechnung.

Die Grössen der Designgewichte in den drei Sprachregionen der Schweiz sind unproblematisch. Das kleinste Strukturgewicht sämtlicher Tage des Jahres 2018 kommt in der Deutschschweiz vor und beträgt 0,28. Diese Fallgruppe wurde also auf 28 Prozent verkleinert. Bis zu 5 Prozent der Stichprobe (5% Perz Min) haben ein Gewicht zwischen 0,28 und 0,6. Der Anteil der Stichprobe mit sehr kleinen Gewichten ist also sehr tief. Die Hochgewichtungen (>1) sind ebenfalls im Rahmen des Zulässigen (Strukturgewichte bis 4 werden als unproblematisch angesehen). Das grösste Gewicht liegt bei 3,19 (DS). Auch hier sind es über die Tage der Jahre gesehen nur maximal 5 Prozent der Zellen, die über 1,98 (95% Perz Max)

liegen. In der Deutschschweiz ist das Designgewicht 1,34; in der Romandie 0,77 und in der italienischen Schweiz ist es 0,31.

Für die Gesamtbeurteilung der Gewichtung im Fernsehpanel sollten nach Ansicht der MWK zukünftig (1) Strukturgewichte, Designgewicht und Hochrechnungsfaktor gesondert ausgewiesen werden, (2) die Betrachtung der Strukturgewichte in den Mittelpunkt der Dokumentation gestellt werden und (3) die Korrelation zwischen Quoten- und Nutzungsmerkmalen dokumentiert werden.

3.1.4 Datenkontrolle und Dokumentation

Die Evening Checks sichern an jedem Abend die automatische Datenlieferung am folgenden Morgen. Diese Datenprüfungen wurden von der MWK geprüft und für eher technisch als methodisch, aber in jedem Fall als nachvollziehbar und sinnvoll erachtet. Die Plausibilitätschecks steigern die Qualität des Panels und sind für die Reliabilität und Validität der gelieferten Daten produktiv.

Während der morgendlichen Routinen der Morning Checks und der Datenlieferung ergänzt Mediapulse die Nutzungsdaten mit Wetterdaten und Sendeprotokollen, die aus methodischer Sicht intervenierende Variablen sind, die Variationen in den Nutzungsdaten erklären können. Das erhöht die korrekte Interpretation der Daten und mithin die Validität der Schlüsse, die aus den Daten gezogen werden. Mit der Datenlieferung werden auch Ausfälle und andere Probleme für die Vertragspartner transparent gemacht.

Die MWK regt auch für die Erhebung der Fernsehreichweiten die Erstellung einer standardisierten, mindestens semesterbezogenen Methodendokumentation mit den jeweils aktuellen Kennwerten zur Stichprobe, zur Datenerhebung und zur Gewichtung sowie Problemberichte (z.B. Toleranzüberschreitungen, Referenzierungsausfälle, Dropouts) an, um langfristig zu einem nachhaltigen und transparenten Berichtswesen der laufenden Erhebung zu kommen. Eine solche Form der Transparenz ist zugleich Dokumentation und Kriterium von Qualität, die die Wertschätzung der Schweizer Fernsehreichweitenforschung nachhaltig steigern und stabilisieren kann.

3.1.5 Fazit

Die MWK bescheinigt der Fernsehforschung der Mediapulse AG und der Fa. Kantar Media eine hohe methodische Qualität. Die im Bericht beschriebenen Verbesserungsvorschläge beziehen sich vor allem auf eine zusätzliche Ausweisung oder Interpretation vorhandener Daten über das Panel und die Gewichtung. Zusätzliche Informationen und eine standardisierte Dokumentation der methodischen Eckdaten können die Transparenz und Nachhaltigkeit der Studie zukünftig erhöhen.

4 Radioreichweitenforschung

Die Radioreichweitenmessung in der Schweiz wurde zum 1. Januar 2018 auf ein neues Erhebungs- und Auswertungssystem umgestellt. Unter anderem wird die Nutzungsmessung mit Hilfe einer neuen Uhr, der Mediawatch 4, vorgenommen, die längere Tragedauern und eine genauere Messung der Audiosignale erlaubt. Nach einer Entwicklungs- und Testphase wurden für das erste Semester 2018 erstmals Reichweitendaten aus dem neuen System freigegeben. Die Konzeption, Durchführung und Datenqualität des neuen Erhebungsmodells wurde in einem Zwischenaudit der MWK evaluiert und positiv bewertet.¹

4.1.1 Grundgesamtheit und Stichprobe

Die Grundgesamtheit der Radioforschung Schweiz ist definiert als die ständige Wohnbevölkerung in Privathaushalten ab 15 Jahren. Der Pool, aus dem die Stichprobe für die tägliche Reichweitenermittlung geschöpft wird, hatte für das Jahr 2018 eine vereinbarte Sollgrösse von n=13'673 Personen. Bis zum dritten Quartal des Jahres 2018 hat sich die tatsächliche Grösse des Hörfunk-Panels kontinuierlich um etwa 500 Personen vermindert. Die Zahl der Austritte je Monat schwankte dabei von Juni bis September zwischen n=96 (im September) und n=110 (im Juni). Die eingebrachten Neurekrutierungen führten dann im September 2018 zu einer hundertprozentigen Deckung (n=13'677), die in den Folgemonaten bis zum Jahresende dann wieder leicht – auf 99 Prozent der Sollstärke – abgesunken ist.

Für die insgesamt 290 Quotenzellen des Panels (29 Gebiete x 5 Altersgruppen, jeweils für Männer und Frauen) zeigen sich einige systematische Unterschreitungen, die weiter beobachtet werden müssen. So wurde zum 1.12.2018 in mehr als 35 Zellen die Sollzahl der Panelisten um mehr als 25 Prozent unterschritten, in fünf Zellen um mehr als 40 Prozent. Vor allem in den jüngeren Altersgruppen, besonders bei den 15-24jährigen zeigen sich überdurchschnittliche Abweichungen von den berechneten Sollzahlen. Vereinzelt deutliche Unterschreitungen gab es auch in der Altersgruppe der über 60jährigen Frauen. Die Sollzahlunterschreitungen haben keinen direkten Einfluss auf die Güte der Reichweitendaten in der jeweils aktuellen Stichprobe für das Radio in der Schweiz. Da das Panel hier aber als Pool für die aktiven Probanden der Tagesstichproben fungiert, sinken mit der (systematischen) Abweichung in den Zellen die freien Wahlmöglichkeiten für die Bestimmung der Probanden. Die Entwicklung der Zellen im Jahresverlauf zeigt jedoch, dass die notwendigen Rekrutierungsanstrengungen unternommen wurden.

Aus dem oben beschriebenen Pool wird eine tägliche Bruttostichprobe von 2400 Probanden vorgehalten, die der Verteilung des Panels hinsichtlich der Strukturvariablen (Rekrutierungszone, Alter, Geschlecht) entspricht. Durch Ausfälle («Missings») reduziert sich diese

¹ Vgl. Trebbe, Joachim (2018): Die Qualität der neuen Radiowährung der Schweiz. Untersuchungskonzeption, Erhebungsmodell und aktuelle Qualitätsindikatoren. Berlin/Bern.

Bruttostichprobe weiter und sinkt – im besten Fall nicht unter die angestrebte Nettostichprobe mit 1840 Probanden. Im Durchschnitt aller Messtage lag der Anteil der Missings an der effektiven Bruttostichprobe bei 19 Prozent mit einem Minimum von 13 Prozent und einem Maximum von 35 Prozent. Systematische Einflüsse auf die Ausfallgründe zeigen sich vor allem zu Ferienbeginn und an Feiertagen. Diese Peaks traten im Jahr 2018 an Ostern (27 Prozent), Auffahrt (25 Prozent), Pfingsten (32 Prozent), vor dem Nationalfeiertag (35 Prozent, das Maximum) und vor Weihnachten (26 Prozent) auf.

Zwischen Juni und Dezember wurde die Sollstärke der Nettostichprobe an 48 von insgesamt 214 Tagen unterschritten. Die Mehrzahl der Unterschreitungen sind quantitativ betrachtet nicht gravierend und liegen unterhalb von 5 Prozent (n=92 Personen) der Sollstärke. An 18 Tagen war diese Abweichung grösser als 5 Prozent, an 3 Tagen wurde die angestrebte Sollstärke um mehr als 10 Prozent (n=184 Personen) unterschritten. Analog zum Pool zeigt sich bei genauerer Betrachtung auch in der Stichprobe eine Häufung der Zellunterschreitungen in der jüngsten Altersgruppe der 15-24jährigen und der über 60jährigen Probanden.

4.1.2 Tragedauern

Die neue Media Watch 4 erlaubt längere Tragedauern. Im Grundsatz wird dabei zwischen Langzeit- und Kurzzeitträgern unterschieden. Kurzzeitträger sind solche Probanden, die die Uhr während 4 Wochen tragen, Langzeitträger teilen sich in Probanden mit 3-monatiger und 6-monatiger Teilnahme an der Radiomessung auf. Für die tägliche Repräsentation der drei Gruppen in der Nettostichprobe wird ein Verhältnis von 60 Prozent (4 Wochen) zu 10 Prozent (3 Monate) zu 30 Prozent (6 Monate) angestrebt. Im Jahresmittel der berichtsrelevanten Monate März bis Dezember 2018 lag der Durchschnitt dieser Anteile in der effektiven Nettostichprobe bei 63 Prozent zu 8 Prozent zu 29 Prozent. Die Schwankungen zwischen den Trägergruppen liegen – insbesondere mit Blick auf die verhältnismässig kleinen Fallzahlen bei den Drei-Monats-Trägern im erwartbaren und vertretbaren Bereich.

Um die Varianzen in der Stichprobe durch die längeren Tragedauern nicht zu stark einzuschränken, wurden Karenzzeiten vereinbart, die für 4-Wochen- und 6-Monatsträger jeweils 24 Wochen und für die 3-Monatsträger 36 Wochen betragen. Die definierten Karenzzeiten waren jeweils für mindestens 95 Prozent der Trägergruppe einzuhalten. Im Jahr 2018 wurden die Karenzzeiten in zwei Trägergruppen unterschritten. Vor allem bei den 1-Monatsträgern ist die Zahl der Unterschreitungen erheblich. Dort wurden 5'263 Teilnahmen (von insgesamt 19'098) in der aktiven Stichprobe realisiert, die unterhalb der Frist von 24 Wochen lagen. D.h. für 72 Prozent anstatt der vereinbarten 95 Prozent der Teilnahmen wurden die Karenzzeiten im Kalenderjahr 2018 eingehalten.

Die durchschnittlichen Tragedauern pro Tag zeigen starke Schwankungen zwischen Wochen- und Wochenendtagen. Wie schon im ersten Semester 2018 liegt der Durchschnitt über alle Tage des Betrachtungszeitraums und alle Probanden über der 12 Stunden-Marke. An den Sonntagen sinkt dieser Wert jedoch um bis zu einer Stunde. Der Anteil der Messtage an denen die Uhr in der Stichprobe mehr als 12 Stunden getragen wird, liegt etwa bei einem

Drittel. Unter methodischen Qualitätsgesichtspunkten interessanter ist dabei weniger der absolute Betrag dieses Werts (der positiv zu bewerten ist) als die vergleichsweise hohe Stabilität der Tragedisziplin im Panel. Die Standardabweichung der Tragestunden pro Tag liegt über alle Messtage betrachtet von Montag bis Sonntag bei etwa 45 Minuten – das spricht für eine vergleichsweise homogene Verteilung der mittleren Tragedauern auf die Messtage.

Im Vergleich zum Zwischenaudit kann man von einer Konsolidierung der Tragedauern sprechen. Die linearen Trends der Mittelwerte für die relevanten Zeitabschnitte «6–12 Stunden» und «12–20 Stunden» verhalten sich nach den Anstiegen im ersten Semester im zweiten Halbjahr konstant horizontal und steigen nicht mehr an. Das spricht für einen Deckeneffekt hinsichtlich der maximalen, durchschnittlichen Tragedauern im Panel.

4.1.3 Datengewichtung und Datenaufbereitung

Durch die Gewichtung der Personendaten werden allfällige Abweichungen zwischen Stichprobenmerkmalen und bekannten Merkmalen der Grundgesamtheit rechnerisch ausgeglichen. Dies geschieht für die Daten der Radioreichweiten auf Tages-, Wochen-, Quartals- und Monatsbasis. Kritisch sind in diesem Zusammenhang sehr kleine Gewichte und vor allem sehr grosse Gewichte, um die angestrebten Merkmalsverteilungen zu erreichen und ein hoher Genauigkeitsverlust durch mangelnde Effektivität der Gewichtung.

Im Zeitraum März bis Dezember 2018 gab es bei der Berechnung der Tagesgewichtung insgesamt fünf geringfügig kritische Fälle. Dabei handelte es sich um drei Fälle im Mai mit einer relevanten Unterschreitung der Gesamtsumme der Personengewichte und um zwei weitere Fälle im Juli bzw. August. Darüber hinaus zeigen sich für das Jahr 2018 an insgesamt 14 Tagen Gewichtungsläufe, in denen die Zahl der Rechendurchgänge überproportional hoch war und 8 weitere Tage, in denen die Summe der Personengewichte geringfügig unterschritten wurde.

Man sieht hier den Zusammenhang zwischen der Güte der Nettostichprobe und dem späteren Gewichtungsaufwand: Unterschreiten die Fallzahlen die Zielgrössen im Sample zu stark, muss bei der Gewichtung auf Extremwerte zurückgegriffen werden. Erfreulicherweise sind diese Fälle im Betrachtungszeitraum des Jahres 2018 jedoch (1) sehr selten und (2) in sehr schwacher Ausprägung aufgetreten, so dass man hier nicht von einer Einschränkung der Datenqualität sprechen kann. Dies zeigt auch die Effektivität der Gewichtung, die auf Tagesbasis im Durchschnitt bei 75 Prozent liegt, bei einem Minimum von 66 Prozent und einem Maximum von 79 Prozent.

Auf der Basis der vorliegenden Daten für die Gewichtungen der Wochen-, Monats- und Quartalsdaten zeigen sich in der Summe keine kritischen Werte für die Personengewichte, die Zahl der Iterationen oder die Effektivität der Gewichtung.

4.1.4 Dokumentation

Die Medienwissenschaftliche Kommission regt an dieser Stelle an, zu einem standardisierten Berichtswesen der methodischen Eckdaten zu kommen. Dafür sollten in regelmässigen Abständen (mindestens für jedes Semester) die wichtigsten Indikatoren zusammengefasst und tabellarisch aufbereitet werden, so dass auch die Kontrolle und Prüfung dieser Daten nicht mehr nur auf der (mehr oder weniger individuellen Durchsicht) durch die Kommissionsmitglieder beruht, sondern auf vergleichbaren, auch für die Vertragspartner aktuellen, stichprobenbezogenen Methodendokumentationen basiert. Aus der Sicht der MWK würde ein solches methodisches Berichtswesen nicht nur die Prüftätigkeit der Kommission erleichtern sondern auch die Akzeptanz und die Wertschätzung der Daten im Markt noch einmal erheblich steigern.

4.1.5 Fazit

Im Hinblick auf die Beurteilung der methodischen Eckdaten des Rekrutierungspools, der Netstichprobe sowie der Datenerhebung, -gewichtung und -auslieferung lassen sich durch die Medienwissenschaftliche Kommission nicht nur keine qualitativen Einschränkungen der Hörfunkdaten feststellen, sondern auch eine weitere Konsolidierung des laufenden Betriebes nach dem offiziellen Wechsel des Erhebungsmodells zum 1. Januar 2018 konstatieren. Nach der Analyse der Kennzeichen und Indikatoren kann der Umstieg auf die neue Uhr mit ihren neuen technischen Möglichkeiten als gelungen betrachtet werden.

Bei Stichprobenerhebungen im Paneldesign betreffen qualitative Einschränkungen fast immer die Aussteuerung und Pflege des Samples. Dies hat jedoch auch etwas mit der Effektivität des Ressourceneinsatzes und der Optimierung methodisch-operationaler Prozesse zu tun. Hier ist die Mediapulse AG mit der Definition und Kontrolle bindender Sollwerte auf dem richtigen Weg. Wenn es gelingt, neben die Einhaltung der hohen Qualitätsstandards auch noch eine standardisierte Methodenberichterstattung zu stellen, wird man die Nachhaltigkeit der Studie im Radio- und Werbemarkt Schweiz langfristig sichern können.

5 Schlussfolgerungen und Empfehlungen

Die Medienwissenschaftliche Kommission hat die Untersuchungskonzeption, die operationelle Umsetzung und die forschungsmethodischen Kennwerte des Establishment-Surveys sowie der Fernseh- und der Radioforschung der Mediapulse AG für das Jahr 2018 geprüft. Den drei empirischen Erhebungen werden im Ergebnis dieser Prüfungen hohe Gültigkeit und Zuverlässigkeit der gewonnenen Daten attestiert.

In dem hier vorgelegten Bericht der MWK für das Jahr 2018 werden konkrete Kritikpunkte und Verbesserungsvorschläge im Detail formuliert, die vor allem auf eine grössere Transparenz und Dokumentation der methodischen Entscheidungen und jeweils aktuellen Kennwer-

te abzielen. Dies hängt sicher auch damit zusammen, dass die Kommission in ihrer jetzigen Besetzung erstmals mit einem Dokumentations- und Zahlenwerk konfrontiert war, das einer langen Tradition und technisch-methodischen Entwicklung entspringt, die erst einmal durchdrungen werden musste. Wir bedanken uns für die offene und jederzeit konstruktive Zusammenarbeit bei den Mitarbeiterinnen und Mitarbeitern der Mediapulse AG und der beauftragten Institute GfK, Kantar Media und M.I.S. Trend.

Drei grundlegende Empfehlungen möchte die Kommission als Ergebnis ihrer Daten- und Dokumenteninspektion für die Weiterentwicklung der Radio- und Fernsehforschung in der Schweiz festhalten.

(1) Wir schlagen für das TV-Panel vor, eine disproportionale Anhebung (Boosts) der Fallzahlen auf mindestens 30 Haushalte in den kleinsten Elementarzonenn oder ihre Zusammenlegung mit angrenzenden EZ, um die statistischen Voraussetzungen für Schlüsse auf die Grundgesamtheit auch für diese Gebiete abzusichern.

(2) Wir plädieren für alle drei hier betrachteten Erhebungen in den Gewichtungsdokumentationen die Stichprobenkorrekturen getrennt nach Design- und Strukturgewichten sowie unabhängig von Hochrechnungsfaktoren auszuweisen, um die Nachvollziehbarkeit der Datenaufbereitung zu erhöhen.

(3) Wir empfehlen – ebenfalls für alle drei Erhebungen – standardisierte, semesterbezogene Methodenberichte mit den wichtigsten Kennzahlen der Stichprobenbildung und Panelpflege, sowie der Datenerhebung und Gewichtung bereitzustellen, um die Qualität der Erhebungen bei den Nutzern transparenter zu machen und mit dieser Transparenz wiederum die Qualität der Forschung zu erhöhen.

Darüber hinaus regt die Kommission an, die gebietsbezogenen Rekrutierungs-, Datenerhebungs- und Gewichtungsstatistiken jeweils auch im GIS-Format zu dokumentieren.

Unabhängig von diesen Diskussionsvorschlägen würdigt die MWK die hohen methodischen Standards der Stichprobenbildung, der Datenerhebung und der Datenaufbereitung und empfiehlt die Publikation der Daten für das Jahr 2018 ohne Einschränkungen.