

Das TV-Duell und die Landtagswahl in Schleswig-Holstein: Das Debat-O-Meter als neues Instrument der politischen Kommunikationsforschung

*Uwe Wagschal, Thomas Waldvogel, Thomas Metz, Bernd Becker, Linus Feiten,
Samuel Weishaupt und Kamaljeet Singh*

1. Ziele der Echtzeitmessung politischer Debatten

TV-Duelle zwischen Spitzenkandidaten politischer Parteien um Spitzenämter sind das mediale Großereignis eines jeden Wahlkampfes, so auch bei Landtagswahlen. Dabei standen die drei Landtagswahlen des ersten Halbjahres 2017 besonders unter dem Vorzeichen der Bundestagswahl am 24. September dieses Jahres, obwohl die agierenden Spitzenpolitiker – wie etwa *Hannelore Kraft* oder *Annegret Kramp-Karrenbauer* – der eigenen Aussage nach dies gar nicht wollten, sondern versuchten, die Landespolitik in den Mittelpunkt ihrer Wahlkämpfe zu stellen. Für verschiedene Akteure der Landespolitik haben TV-Debatten aus ganz unterschiedlichen Gründen eine herausragende Bedeutung: Für Politiker und Wahlkämpfer liegt der besondere Reiz in der großen Reichweite. Vor Hundertausenden Zuschauern können die Spitzenpolitiker auf unmittelbare Weise Positionen beziehen und zentrale Argumente unter weitgehender Meidung medialer Selektionskriterien vortragen. Für die Medien sind TV-Duelle deshalb besonders attraktiv, weil sie leicht kommunizierbar sind, einen hohen Unterhaltungscharakter aufweisen sowie dem Trend von Personalisierung und Mediatisierung von Wahlberichterstattung entsprechen.¹ Für Wählerinnen und Wähler bieten TV-Duelle eine sehr gute Gelegenheit, das politische Spitzenpersonal der für die Besetzung des Amtes des Regierungschefs relevanten Parteien direkt miteinander zu vergleichen und deren Programme gegenüberzustellen. TV-Duelle eröffnen der Wählerschaft damit ein attraktives Angebot der Wahlinformation und politischen Bildung zur häufig weniger beachteten Landespolitik.

Und auch für die Kommunikations- und Politikwissenschaft sind sie ein reizvolles Forschungsfeld: Spätestens seit ihrer Etablierung als zentrales Ereignis des Medienwahlkampfes zur Bundestagswahl 2002 hat sich die Wissenschaft intensiv mit den Inhalten, Wahrnehmungsprozessen und Wirkungsmechanismen von TV-Duellen auf deren Rezipienten befasst. In der Landespolitik gehören diese Duelle bereits seit der Hamburger Bürgerschaftswahl 1997 zum medialen Repertoire; seitdem sind sie fester Bestandteil der Landtagswahlkämpfe in der Bundesrepublik.

Der vorliegende Beitrag präsentiert die erste auf einer großen Fallzahl basierende Echtzeitanalyse (RTR-Messung = Real Time Response) einer politischen Debatte in Deutschland, die nicht unter Laborbedingungen durchgeführt wurde.² Die Daten hierfür wurden

- 1 Vgl. *Harald Schoen*, Wahlkampfforschung, in: *Jürgen W. Falter / Harald Schoen* (Hrsg.), Handbuch Wahlforschung, Wiesbaden 2014, S. 661 – 728.
- 2 Im Rahmen der German Longitudinal Election Study (GLES) wurde anlässlich der Bundestagswahl 2013 eine insgesamt 269 Teilnehmer umfassende Laborstudie an unterschiedlichen Standorten durchgeführt. Dabei wurden Drehregler und Desktop-Computer verwendet. Vgl. *Thorsten Faas / Jürgen Maier / Michaela Maier* (Hrsg.), *Merkel gegen Steinbrück. Analysen zum TV-Duell vor der Bundestagswahl 2013*, Wiesbaden 2017.

anlässlich der Landtagswahl in Schleswig-Holstein mit dem Debat-O-Meter erhoben, einer Web-Applikation, die auf unterschiedlichen internetfähigen Endgeräten wie Smartphones, Tablets oder PCs verwendet werden kann. In dem TV-Duell standen sich am 25. April 2017 der amtierende Ministerpräsident *Torsten Albig* (SPD) und sein Herausforderer *Daniel Günther* (CDU) gegenüber.

In dem Beitrag sollen für die Landtagswahl folgende vier Fragen beantwortet werden: (1) Wer hat die Debatte gewonnen und warum? (2) Welche Themen der Diskutanten überzeugten die Zuschauer besonders und welche nicht? (3) Wie groß war der Anteil der Unentschlossenen, und wie haben sich diese nach der Debatte entschieden? (4) Hat sich der Debattenausgang auf das tatsächliche Wahlergebnis ausgewirkt?

2. Die Bedeutung von TV-Duellen und ihre Vermessung

Sowohl in der nationalen als auch in der internationalen Forschung haben sich Studien zu TV-Duellen vor allem dann als besonders erkenntnisreich erwiesen, wenn sie methodisch die Real-Time-Response-Messung ins Zentrum ihres Designs rücken.³ Diese Methode, im angelsächsischen Raum auch als CRM für Continuous Response Measurement bezeichnet, erlaubt die rezeptionsbegleitende Erfassung individueller Reaktionen auf audio-visuelle Stimuli in Echtzeit. TV-Duelle zwischen Politikern sind dabei besonders häufige und interessante Anwendungen. Die Zuschauer (oder Probanden in Experimenten) können hierbei mit Hilfe technischer Eingabegeräte wie Smartphones (oder in der Vergangenheit Drehreg-

- 3 Zum Beispiel *Dan Schill / Rita Kirk*, Courting the Swing Voter: 'Real Time' Insights Into the 2008 and 2012 U.S. Presidential Debates, in: *American Behavioral Scientist*, 58. Jg. (2014), H. 4, S. 536 – 555; *Marko Bachl / Frank Bretschneider / Simon Ottler*, Die TV-Duell-Studie Baden-Württemberg 2011, in: *dies.* (Hrsg.), Das TV-Duell in Baden-Württemberg 2011, Wiesbaden 2013, S. 7 – 27; *Amber E. Boydston / Rebecca A. Glazier / Matthew T. Pietryka / Philip Resnik*, Real-Time Reactions to a 2012 Presidential Debate. A Method for Understanding Which Messages Matter, in: *Public Opinion Quarterly*, 78. Jg. (2014), H. S1, S. 330 – 343; *Thomas Metz / Uwe Wagschal / Thomas Waldvogel / Marko Bachl / Linus Feiten / Bernd Becker*, Das Debat-O-Meter: ein neues Instrument zur Analyse von TV-Duellen, in: *Zeitschrift für Staats- und Europawissenschaften*, 14. Jg. (2016), H. 1, S. 124 – 149; *Jürgen Maier / Thorsten Faas / Michaela Maier*, Aufgeholt, aber nicht aufgeschlossen: Ausgewählte Befunde zur Wahrnehmung und Wirkung des TV-Duells 2013 zwischen Angela Merkel und Peer Steinbrück, in: *ZParl*, 45. Jg. (2014), H. 1, S. 38 – 54; *Marcus Maurer / Carsten Reinemann / Jürgen Maier / Michaela Maier* (Hrsg.), Schröder gegen Merkel. Wahrnehmung und Wirkung des TV-Duells 2005 im Ost-West-Vergleich, Wiesbaden 2007; *Marcus Maurer / Carsten Reinemann*, Schröder gegen Stoiber: Nutzung, Wahrnehmung und Wirkung der TV-Duelle, Wiesbaden 2003; *Friederike Nagel*, Die Wirkung verbaler und nonverbaler Kommunikation in TV-Duellen. Eine Untersuchung am Beispiel von Gerhard Schröder und Angela Merkel, Wiesbaden 2012; *Georgios Papastefanou*, Reliability and Validity of RTR Measurement Device. Working Paper, http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/gesis_arbeitsberichte/WorkingPapers_2013-27.pdf (Abruf am 4. August 2017); *Barbara Wolf*, Beurteilung politischer Kandidaten in TV-Duellen: Effekte rezeptionsbegleitender Fremdbeurteilungen auf Zuschauerurteile, Baden-Baden 2010; *Dan Schill / Rita Kirk / Amy E. Jasper*, Political Communication in Real Time: Theoretical and Applied Research Approaches, New York / London 2017; *Jürgen Maier / Michaela Maier / Marcus Maurer / Carsten Reinemann* (Hrsg.), Real-Time Response Measurement in the Social Sciences. Methodological Perspectives and Applications, Frankfurt am Main 2009; *Thorsten Faas / Jürgen Maier / Michaela Maier* (Hrsg.), a.a.O. (Fn. 2).

ler) kontinuierlich ihren Eindruck wiedergeben. Die in Echtzeit erfassten Datenpunkte werden zentral auf einem Speichermedium gesammelt und stehen dann für die weitere statistische und grafische Aufbereitung zur Verfügung.

RTR-Messungen werden dabei zumeist mit Inhaltsanalysen verbunden und von (Panel-) Befragungen gerahmt. Somit erlauben diese Studiendesigns eine detaillierte Untersuchung von Wahrnehmungsprozessen und Wirkungsmechanismen politischer TV-Debatten auf seine Rezipienten unter Berücksichtigung individueller, soziodemographischer und politischer Prädispositionen. In der US-amerikanischen Forschung hielt die rezeptionsbegleitende Messmethodik bereits 1945 mit dem „program analyzer“ von *Paul Lazarsfeld*, seinerzeit Direktor des Bureau of Applied Social Research an der Columbia University in New York, und *Frank Stanton*, Medienforscher und später Präsident des kommerziellen Fernsehkonzerns CBS, Einzug. Hier entwickelte sie sich in der Politikwissenschaft vor allem ab den 1960ern mit dem Aufkommen von „presidential debates“ und deren Etablierung in den 1970er Jahren weiter. Besondere Dynamik entfaltete diese Entwicklung in den Folgejahren auch durch die Öffnung der Wissenschaft zu kognitivistischen Ansätzen, die die Erforschung individueller Informationsverarbeitungsprozesse in den Blick nahmen und mit dem Durchbruch von Mikrocomputer-Systemen in den 1980ern zusätzlich befördert wurden. In Deutschland ist der Durchbruch von Real-Time-Response hingegen sehr eng mit dem Aufkommen von TV-Duellen zwischen Spitzenpolitikern vor Bundestagswahlen um die Jahrtausendwende verknüpft.⁴

Seit ihrer Etablierung haben sich unterschiedliche technische Spezifikationen des RTR-Messinstrumentariums entwickelt. Sie reichen von physischen Drehreglern, Druckknopf-Systemen (Push-button) und Schieberegglern (Slider) über Joysticks bis hin zu internetbasierten Verfahren. In ihren Konfigurationen (Datenniveau, Eingabemodus und der Anzahl der Bewertungselemente) zuweilen sehr unterschiedlich, versprechen sie eine zuverlässige und gültige Erfassung der individuellen Informationsverarbeitungsprozesse von Zuschauern politischer TV-Duelle.⁵ Wenngleich die Reliabilität und Validität des Messverfahrens auch kritisch hinterfragt werden⁶, konnte in der methodischen Forschung gezeigt werden, dass

4 Vgl. *Thomas Waldvogel / Thomas Metz*, Real-Time-Response-Messungen, in: *Sebastian Jäckle* (Hrsg.), Neue Trends in den Sozialwissenschaften, Wiesbaden 2017, S. 307 – 331; *Marcus Maurer*, Real-Time Response Messung: Kontinuierliche Befragung in Echtzeit, in: *Wiebke Möhring / Daniela Schlütz* (Hrsg.), Handbuch standardisierte Erhebungsverfahren in der Kommunikationswissenschaft, Wiesbaden 2013, S. 219 – 234; *Simon Ottler*, RTR-Messung: Möglichkeiten und Grenzen einer sozialwissenschaftlichen Methode, in: *Marko Bachl / Frank Bretschneider / Simon Ottler* (Hrsg.), a.a.O. (Fn. 3), S. 113 – 134; *Frank Biocca / Prabu David / Mark West*, Continuous Response Measurement (CRM): A Computerized Tool for Research on the Cognitive Processing of Communication Messages, in: *Annie Lang* (Hrsg.), Measuring Psychological Responses to Media Messages, Hillsdale 1994, S. 15 – 64.

5 Vgl. *Carsten Reinemann / Jürgen Maier / Thorsten Faas / Marcus Maurer*, Reliabilität und Validität von RTR-Messungen, in: *Publizistik*, 50. Jg. (2005), H. 1, S. 56 – 73; *Jürgen Maier / Marcus Maurer / Carsten Reinemann / Thorsten Faas*, Reliability and Validity of Real-Time Response Measurement: a Comparison of Two Studies of a Televised Debate in Germany, in: *International Journal of Public Opinion Research*, 19. Jg. (2007), H. 1, S. 53 – 73; *Jürgen Maier / Marcus Maurer / Carsten Reinemann / Thorsten Faas*, a.a.O. (Fn. 3).

6 Vgl. *Marko Bachl*, Analyse rezeptionsbegleitend gemessener Kandidatenbewertungen in TV-Duellen. Erweiterung etablierter Verfahren und Vorschlag einer Mehrebenenmodellierung, Hohenheim 2014; *Georgios Papastefanou*, a.a.O. (Fn. 3).

Druckknopf- und Drehregler-Systemen eine akzeptable Reliabilität aufweisen, die im Vergleich von physischen Drehreglern und digitalen Slidern noch übertrroffen wird.⁷

Innovation erfährt die rezeptionsbegleitende Forschung zu TV-Duellen seit wenigen Jahren durch die Virtualisierung seines Messinstrumentariums. Virtualisierte Formen der RTR-Messtechnik machen dabei mobile Endgeräte wie Smartphones, Tablets oder Notebooks als Messgeräte nutzbar, indem sie die RTR-Technologie als softwarebasierte Anwendung implementieren. Entweder als native App⁸ oder als Web-Anwendung⁹ programmiert, gleichen virtualisierte Formen des RTR-Instrumentariums bisher stark bestehenden Messsystemen physischen Typus: Die grafischen Benutzeroberflächen (graphical user interface; GUI) werden dabei äquivalent zum klassischen Instrumentarium als Slider¹⁰ beziehungsweise Dialer¹¹ im latched mode (das heißt es wird kontinuierlich die momentane Position aufgezeichnet) oder als Push-Button-Version im reset mode (das heißt die Geräte springen nach der Eingabe automatisch wieder in die Ausgangsstellung zurück) umgesetzt.

Die Virtualisierung des Messinstrumentariums bringt eine Vielzahl an Vorteilen mit sich: Da Studienteilnehmer eigene Endgeräte zur Eingabe nutzen, entfällt ein Großteil der erheblichen Kosten für die Hardware. Das virtualisierte Messinstrumentarium zeichnet sich zudem durch eine hohe Flexibilität aus. Unterschiedliche Konfigurationen der GUI (Slider, Dialer, Push-Button) können kosteneffizient softwareseitig umgesetzt werden. Außerdem ist in diesem Bereich von einem großen Innovationspotenzial auszugehen, denn durch informationstechnische Implementierung können bestehende Oberflächengestaltungen weiterentwickelt und gänzlich neue Funktionen und Darstellungsformen umgesetzt werden, die es Forschern ermöglichen, bisher unzugängliche Konstrukte und Prozesse der individuellen Informationsverarbeitung zu vermessen.¹² Weiterhin lassen sich die Skalierung und damit das Niveau der erhobenen Daten und der Eingabemodus (reset mode vs. latched mode) flexibel konfigurieren. Die Anzahl der Bewertungsitems kann zudem an den vorliegenden Medienstimulus angepasst werden, was vor allem im Hinblick auf die Diversität politischer TV-Diskussionen, die von einer Duell-Situation bis zu einem Vielpersonen-Podium reichen können, von großer Bedeutung ist. So ist es etwa mit den bisherigen Drehreglern kaum möglich, Debatten zwischen mehreren Teilnehmern zu erfassen. Weitrei-

7 Vgl. Thomas Metz / Uwe Wagschal / Thomas Waldvogel / Marko Bachl / Linus Feiten / Bernd Becker, a.a.O. (Fn. 3); Carsten Reinemann / Jürgen Maier / Thorsten Faas / Marcus Maurer, a.a.O. (Fn. 5).

8 Vgl. Jürgen Maier / J. Felix Hampe / Nico Jahn, Breaking Out of the Lab Measuring Real-Time Responses to Televised Political Content in Real-World Settings, in: *Public Opinion Quarterly*, 80. Jg. (2016), H. 2, S. 542 – 553.

9 Vgl. Thomas Metz / Uwe Wagschal / Thomas Waldvogel / Marko Bachl / Linus Feiten / Bernd Becker, a.a.O. (Fn. 3); Amber E. Boydston / Rebecca A. Glazier / Matthew T. Pietryka / Philip Resnik, a.a.O. (Fn. 3).

10 Vgl. Thomas Metz / Uwe Wagschal / Thomas Waldvogel / Marko Bachl / Linus Feiten / Bernd Becker, a.a.O. (Fn. 3); Jürgen Maier / J. Felix Hampe / Nico Jahn, a.a.O. (Fn. 8); Jan Kercher / Marko Bachl / Frank Vohle, The “MediaLiveTracker”. A New Online Tool for Real-Time-Response-Measurement, https://www.uni-hohenheim.de/fileadmin/einrichtungen/komm/PDFs/Komm/Publikationen/Vortrag_MediaLiveTracker_GOR_2012.pdf (Abruf am 4. August 2017).

11 Vgl. ebenda.

12 Vgl. Philip Resnik / Amber E. Boydston / Rebecca A. Glazier / Matthew T. Pietryka, Scalable Multi-dimensional Response Measurement using a Mobile Platform, in: *Dan Schill / Rita Kirk / Amy E. Jasper (Hrsg.), Political Communication in Real Time. Theoretical and Applied Research Approaches*, New York 2017.

chend für den gesamten Forschungsbereich dürften zudem die methodologischen Implikationen sein, die eine Virtualisierung des Messinstrumentariums mit sich bringt: Messungen sind nun leicht auch außerhalb des Labors und damit in natürlichen Rezeptionssituationen realisierbar, was die berechtigte Kritik der mangelnden externen Validität der bisherigen rezeptionsbegleitenden Forschung entkräftet.

Es ist zudem zu erwarten, dass Effekte der Selbstselektion der Teilnehmerschaft abgeschwächt werden: Bisher stellte die Teilnahme an einer RTR-Studie hohe Anforderungen (Motivation, Anfahrt, Anreizsetzungen usw.) an die Versuchspersonen. Dies führte dazu, dass vor allem stark interessierte Personen für eine Teilnahme zu gewinnen sind, die sich aber regelmäßig in ihrer politischen Informiertheit¹³, der Klarheit und Stabilität ihrer Orientierungen¹⁴ und in der Art der politischen Informationsverarbeitung¹⁵ von der breiten Bevölkerung unterscheiden. Diese Selbstselektion birgt die Gefahr, dass Effekte von TV-Debatten auf die Rezipienten systematisch unterschätzt werden. Als applikative Heimanwendung senken virtualisierte Formen des RTR-Messinstrumentariums wie das Debat-O-Meter diese Hürden beträchtlich und dürften bei gewissen Teilen der Nutzer auch das Motiv der Teilnahme verändern: Nicht die Teilnahme an einer wissenschaftlichen Studie oder ein besonderer Anreiz bilden die Motivation, sondern vielmehr das Angebot eines „Second Screen“-Erlebnisses, also die Möglichkeit, über einen „zweiten Bildschirm“ parallel zur Fernsehsendung seine Meinung in Echtzeit äußern und damit an der Diskussion und der folgenden Berichterstattung partizipieren zu können. Auch das Angebot eines „digitalen Wahlhelfers“ (so genannte Voting Advice Application) als innovatives Instrument der politischen Wahlinformation und Bildung dürften Grund für eine Nutzung der Applikation sein, denn der Nutzer hat durch die Offenlegung des individuellen Abstimmungsverhaltens eine direkte Rückmeldung, mit welchem Kandidat er in unterschiedlichen Themenfeldern jeweils übereinstimmt.

Neben der höheren Kosteneffizienz, der gesteigerten Flexibilität und Mobilität des RTR-Instrumentariums verbunden mit einem hohen Innovationspotenzial und der Verbesserung der externen Validität des Messverfahrens wirkt sich die Virtualisierung auch auf die Größe und Qualität der Stichprobe aus. Dies erhöht zum einen die Robustheit und Generalisierbarkeit der abgeleiteten Befunde und erlaubt zudem die Entfaltung eines feingranulierten Quotenplans, in dem die einzelnen Felder ausreichend gefüllt werden können.¹⁶ Hinzu kommt, dass mit der online-basierten Datenerhebung die räumliche Repräsentation der Befragten noch deutlich verbessert werden kann, da die einzige Bedingung eine Netzverbindung ist. Da diese Abdeckung durch Personalcomputer (inklusive Notebooks und Tablets rund 89 Prozent der Haushalte) und Smartphones (rund 95 Prozent der Haushalte) mittlerweile auch besser ist als etwa die in Telefonumfragen verwendeten Festnetznummern, dürfte die potenzielle Reichweite des Instruments sogar mittlerweile höher liegen als die der klassischen Wahlumfragen mittels Festnetztelefon (rund 91 Prozent der Haushalte).

13 Vgl. *Philip E. Converse*, Assessing the Capacity of Mass Electorates, in: *Annual Review of Political Science*, 3. Jg. (2000), H. 1, S. 331 – 353.

14 Vgl. *ders.*, The Nature of Belief Systems in Mass Publics (1964), in: *Critical Review*, 18. Jg. (2006), H. 1, S. 1 – 74.

15 Vgl. *Angus Campbell / Philip E. Converse / Warren Miller / Donald Stokes*, *The American Voter*, New York 1960.

16 Vgl. *Thomas Metz / Uwe Wagschal / Thomas Waldvogel / Marko Bachl / Linus Feiten / Bernd Becker*, a.a.O. (Fn. 3); *Philip Resnik / Amber E. Boydston / Rebecca A. Glazier / Matthew T. Pietryka*, a.a.O. (Fn. 12).

3. Das Debat-O-Meter als neue Erhebungsmethode: Methodik und Design

Während die Duelle zwischen den Kanzlerkandidaten vor Bundestagswahlen detaillierten politik- und kommunikationswissenschaftlichen Analysen unterzogen wurden¹⁷, ist dies für TV-Debatten zu Landtagswahlen nur bedingt der Fall. Eine den wissenschaftlichen Untersuchungen zur Bundestagswahl in Intensität und Detailreichtum äquivalente Studie auf Landesebene haben bisher einzige *Marko Bachl* zum TV-Duell *Stefan Mappus* gegen *Nils Schmid* zur Landtagswahl 2011 in Baden-Württemberg vorgelegt.¹⁸

Diese Situation ist unzureichend, weil derzeit vollkommen ungeklärt scheint, ob Erkenntnisse zu TV-Duellen, die insbesondere im Rahmen von US-amerikanischen „presidential debates“ und bundesdeutschen Kanzlerduellen gewonnen wurden, auch Gültigkeit im Kontext von Landtagswahlen besitzen. Generell sind die wissenschaftlichen Analysen zu den TV-Debatten in Deutschland überschaubar und haben bisher niemals mehr als 300 Befragte als Grundlage gehabt; es wurden sogar Studien publiziert, die auf nur rund 30 Befragten basierten.¹⁹

Ursachen für dieses Defizit dürften darin liegen, dass nur wenige Forschergruppen in Deutschland über das notwendige Instrumentarium verfügen und hohe Kosten für die technische Infrastruktur zur Analyse von TV-Duellen eine Ausweitung des Forschungsbereichs auf die als wenig „attraktiv“ geltende Ebene der Landespolitik verhindern. Mit dem Debat-O-Meter, einem von Informatikern und Politikwissenschaftlern der Universität Freiburg entwickelten onlinebasierten Messinstrument zur Untersuchung von Inhalten, Wahrnehmungen und Wirkungen von TV-Debatten, liegt nun ein Analyseinstrument vor, das eine vergleichsweise ressourcenschonende Untersuchung von TV-Duellen in Echtzeit auch auf Landesebene ermöglicht und zahlreiche Restriktionen des bisherigen Messinstrumentariums überwindet (siehe oben).

Als virtualisiertes Messinstrument zeichnet sich das Debat-O-Meter durch seine modulare Struktur aus. Es reicht in seinem Funktionsumfang damit weit über ein klassisches RTR-Eingabegerät hinaus. Vielmehr ist das Debat-O-Meter eine virtuelle Plattform, mit der sich „normale“ RTR-Studiendesigns online durchführen lassen und das deren Phasenstruktur in einem „virtuellen Versuchslabor“ umsetzt. Unter den bisherigen Anwendungen des Debat-O-Meters waren etwa die dritte US-Präsidentenstsdebatte zwischen *Hilary Clinton* und *Donald Trump* (unter anderem mit Teilnehmern aus den USA), die Brexit-Debatte, das TV-Duell der baden-württembergischen Spitzenkandidaten am 14. Januar 2016 (*Winfried Kretschmann* versus *Guido Wolf*) sowie die Elefantenrunde am 16. März 2017 anlässlich der Landtagswahl im Saarland und die erste Runde der Präsidentschaftswahl in Frankreich am 20. März 2017 (in Französisch). Hierbei wurde das Erhebungsinstrument erprobt und weiterentwickelt, insbesondere wurde die Datensicherheit verbessert. Zudem wurden verschiedenen Eingabemasken getestet. Allerdings betrug die maximale Teilnehmerzahl in dieser Phase noch nicht mehr als 250 Personen.

17 Siehe zum Beispiel *Marcus Maurer / Carsten Reinemann / Jürgen Maier / Michaela Maier* (Hrsg.), a.a.O. (Fn. 3); *Thorsten Faas / Jürgen Maier / Michaela Maier* (Hrsg.), a.a.O. (Fn. 2); *dies.*, a.a.O. (Fn. 3).

18 Vgl. *Marko Bachl / Frank Brettschneider / Simon Ottler* (Hrsg.), a.a.O. (Fn. 3).

19 Vgl. *Jürgen Maier / Marcus Maurer / Carsten Reinemann / Thorsten Faas*, a.a.O. (Fn. 5); *Jürgen Maier / J. Felix Hampe / Nico Jahn*, a.a.O. (Fn. 8).

Auf Seiten des Benutzers (Client) ist das Debat-O-Meter als einfache Website programmiert, deren Aussehen einer gewöhnlichen Smartphone-Applikation nachempfunden ist. Die Modulstruktur des Debat-O-Meter folgt dabei dem klassischen RTR-Studienaufbau: Der anonymisierten Registrierung folgt ein Tutorial, in dem die Benutzeroberfläche erklärt und die Messanweisung erläutert wird. Daran schließt sich eine Vorbefragung an, in der soziodemographische Variablen, politische Einstellungen und Verhaltensweisen der Teilnehmer sowie Einschätzungen und Erwartungen an die Debatte und die Diskutanten abgefragt werden. Im Kern der Web-Anwendung steht das RTR-Modul, das es den Nutzern ermöglicht, ihre (der Messanweisung entsprechende) Bewertung in Echtzeit abzugeben, die sekundengenau vom Debat-O-Meter erfasst und mit einem Zeitstempel versehen gemeinsam mit der pseudonymisierten Nutzerkennung in der Datenbank auf einem Server gespeichert wird. Am Ende der Debatte werden die Nutzer unmittelbar zur Nachbefragung weitergeleitet. Daran anschließend erhalten sie als Anreiz und Informationsangebot in einem Analysemodul Einblick in erste Ergebnisse der RTR-Messung.

Die Gesamtergebnisse werden den Benutzern und allen Interessierten zudem auf einer interaktiven Webseite zur Verfügung gestellt, auf der auch eigene deskriptive Auswertungen durchgeführt werden können.²⁰

Ein weiteres, für den Anwender nicht sichtbares Instrument („Annotierungstool“) ermöglicht die Erhebung kommunikativer Strategien. Hier wird beispielsweise erfasst, ob der Diskutant aggressiv, emotional oder empathisch ist. Ferner werden Angriffs- oder Verteidigungsstrategien erhoben, außerdem die Themen, die Dauer der einzelnen Beiträge sowie wichtige thematische Äußerungen.

Bisher liegen für die Rekrutierung von Studienteilnehmern für RTR-Messungen mit virtualisierten Eingabegeräten, die das Potenzial dieser Innovation nutzbar machen, kaum Erfahrungen vor. Neben klassischen Rekrutierungsstrategien von Studierenden des eigenen Campus²¹ hat sich das Colleague Crowdsourcing²² als erfolgreiche Methode erwiesen, die Studien mit großen Teilnehmerzahlen möglich macht. Beide Strategien wurden mit dem Debat-O-Meter etwa bei der Landtagswahl 2016 in Baden-Württemberg (1. Weg) und bei der dritten TV-Debatte zwischen *Trump* und *Clinton* (Oktober 2016) sowie der TV-Debatte in Frankreich zur Wahl in der ersten Präsidentschaftsrunde (März 2017) angewendet (2. Weg). Bei letzteren wurden die Erhebungen zudem in Englisch und Französisch durchgeführt.

Die hier vorliegende Studie wählte einen anderen Ansatz: Mit Hilfe von Medienpartnerschaften wurde für das Debat-O-Meter als „Second Screen“-Angebot, als Beteiligungsprojekt und als Angebot der niederschwülligen Wahlinformation breit geworben. Für die Vorberichterstattung erhielten die Medienpartner verschiedenartige Inhalte zur Verfügung gestellt (z.B. Experteninterviews), mit denen potenzielle Teilnehmer auf das Angebot aufmerksam gemacht wurden. Ungefähr eine halbe Stunde nach dem Ende der TV-Debatte erhielten drei

20 Die URL lautet für die Debatte in Schleswig-Holstein etwa <https://debatometer.shinyapps.io/ltwsh17/>.

21 Vgl. *Jürgen Maier / J. Felix Hampe / Nico Jahn*, a.a.O. (Fn. 8).

22 Vgl. *Amber E. Boydston / Jessica T. Fiezell / Rebecca A. Glazier / Timothy P. Jurka / Matthew T. Pietryka*, Colleague Crowdsourcing: A Method for Incentivizing National Student Engagement and Large-N Data Collection, in: PS: Political Science & Politics, 47. Jg. (2014), H. 4, S. 829 – 834.

Zeitungssredaktionen eine verschriftlichte Blitzauswertung mit ersten Ergebnissen der Daten für die unmittelbare Berichterstattung online und gedruckt. Am nächsten Tag folgte eine Detailanalyse für ausführliche Nachberichte. Dabei konnten in Schleswig-Holstein drei reichweitenstarke Zeitungen (Kieler Nachrichten, Lübecker Nachrichten und Segeberger Zeitung) gewonnen werden, die über eine intensive Vorberichterstattung – sowohl gedruckt als auch online – Werbung für das Projekt machten. Ergänzt wurde dieses Vorgehen durch Social Media-Werbung auf Facebook und Twitter sowie in Online-Anzeigenblättern.

Für Erhebungen zu den Landtagswahlen in Schleswig-Holstein erfolgte die Eingabe in der Push-Button-Implementierung im reset mode mit dem eigenen mobilen Endgerät. Über eine graduelle Abstufung von „+“ für eine sehr gute bis hin zu „-“ für eine sehr schlechte Bewertung sollten die Versuchsteilnehmer ihren momentanen Eindruck der Spitzenkandidaten wiedergeben. Das Nicht-Drücken eines Knopfes wurde dabei als neutraler Eindruck interpretiert. Für die Auswertung werden im Debat-O-Meter diese Daten auf eine Skala von -2 bis +2 umkodiert.

Tabelle 1: Soziodemographische Verteilung der Teilnehmer in Schleswig-Holstein

	Debat-O-Meter		landesweit
	N	%	%
Geschlecht			
weiblich	228 (274)	32,8 (32,0)	51,1
männlich	467 (582)	67,2 (68,0)	48,9
Alter			
unter 24 Jahre	165 (210)	23,7 (24,5)	24,0
25 bis 29 Jahre	83 (106)	11,9 (12,4)	5,7
30 bis 39 Jahre	124 (152)	17,8 (17,8)	11,1
40 bis 49 Jahre	123 (155)	17,7 (18,1)	14,5
50 bis 59 Jahre	97 (114)	14,0 (13,3)	15,9
60 bis 69 Jahre	68 (75)	9,8 (8,8)	11,9
über 70 Jahre	35 (44)	5,0 (5,1)	16,9
Bildung			
Schüler	30 (37)	4,3 (4,3)	13,3
Haupt-/Volksschulabschluss	14 (18)	2,0 (2,1)	26,1
Mittlere Reife	91 (125)	13,1 (14,6)	32,0
Fachhochschul- oder Hochschulreife	235 (291)	33,9 (34,1)	31,5
Hochschulabschluss	324 (383)	46,7 (44,8)	14,9
Anmerkung: Angaben aller Teilnehmer, die eine Bewertung abgegeben haben. Da später teilnehmende Personen keine Vorbefragung mehr ausgefüllt haben, liegen nicht für alle Personen demographische Daten vor. In Klammern: Alle in der Vorbefragung registrierten Angaben. Differenzen ergeben sich auch daraus, dass einzelne nach der Vorbefragung die Teilnahme abgebrochen haben. Die fast identische Zusammensetzung beider Gruppen lässt vermuten, dass die Abbrüche unsystematisch sind.			
Quelle: Eigene Daten; landesweit: Daten für Geschlecht: Statistisches Amt für Hamburg und Schleswig-Holstein, Die Bevölkerung in Schleswig-Holstein nach Alter und Geschlecht 2015. Endgültige Ergebnisse Fortschreibung auf Basis des Zensus 2011 (A I 3 - j 15 SH), Hamburg 2016. Daten für Bildung: dass., Statistisches Jahrbuch Schleswig-Holstein 2016/2017, Hamburg 2017, S. 45, https://www.statistik-nord.de/fileadmin/Dokumente/Jahrb%C3%BCcher/Schleswig-Holstein/JB16SH_02_fertig.pdf (Abruf am 4. August 2017).			

Der eingeschlagene Weg zur Rekrutierung von Studienteilnehmern erwies sich quantitativ als sehr erfolgreich. In Schleswig-Holstein beteiligten sich insgesamt 856 Personen an der RTR-Bewertung. Es zeigt sich generell, dass die Erhebung in Schleswig-Holstein näher an der landesweiten Demographie lag als unsere eine Woche später in Nordrhein-Westfalen durchgeführte. Bei der Bildungsvariable sind die Unterschiede im Vergleich zum Landesdurchschnitt jedoch beachtlich. Dennoch ist dieses Verfahren nicht schlechter als bisherige Verfahren, die etwa mit Quotenplänen und Vorabrekrutierungen operieren. So liegt die durchschnittliche absolute Abweichung bei der Altersvariablen unter der Experimentalstudie für die Bundestagswahl 2009 und auf ähnlichem Niveau wie bei der Studie von 2013.²³ Daher wurde für die empirische Untersuchung eine Gewichtung mit den soziodemographischen Faktoren Geschlecht und Alter vorgenommen.

Es haben sich mehr Männer als Frauen am Debat-O-Meter beteiligt (vgl. Tabelle 1), und auch der formale Bildungsabschluss der Teilnehmer ist deutlich höher als im Landesdurchschnitt. Überdies haben vor allem politisch Interessierte am Debat-O-Meter teilgenommen: 78,8 Prozent der Teilnehmer waren stark beziehungsweise sehr stark politisch interessiert.²⁴ Dieser Befund zeigt sich im Übrigen auch bei den bisherigen Laborstudien zur Bundestagswahl 2009 und 2013.

4. Die TV-Debatte in Schleswig-Holstein

4.1. Ergebnisse der Vorbefragung

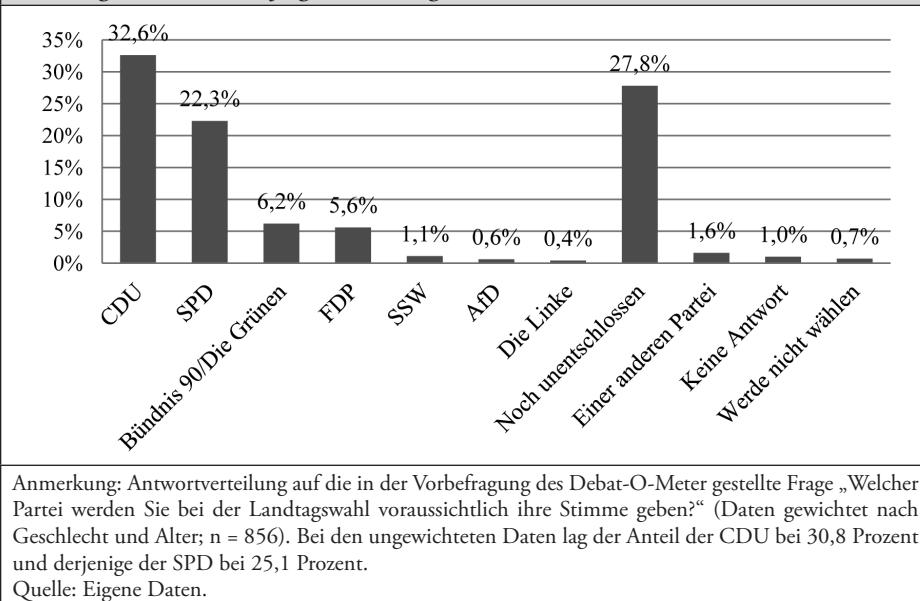
Die Ausgangssituation war zwei Monate vor der Landtagswahl – gemessen an den Wahlumfragen – und der politischen Stimmungslage („Schulz-Effekt“) eher positiv für den Ministerpräsidenten *Albig* (SPD) und die Fortführung der so genannten Küstenkoalition aus SPD, Grünen und Südschleswigschem Wählerverband (SSW). Hinzu kamen interne Parteiquerelen (Rücktritt des ursprünglichen Spitzenkandidaten) bei der CDU, so dass erst am 19. November 2016 mit *Daniel Günther* ein neuer Spitzenkandidat gefunden werden konnte, der zudem nur einen geringen Bekanntheitsgrad aufwies.

Das TV-Duell zwischen ihnen fand am 25. April 2017 statt, zwölf Tage vor dem Wahltermin. Dabei folgte die Konzeption der 60-minütigen Live-Sendung des NDR weniger einem klassischen Duell, sondern lehnte sich nach US-amerikanischem Vorbild vielmehr an die Idee eines Town Hall Meetings an. So wurden die Fragen nicht vom Moderator und Chefredakteur des NDR Fernsehens *Andreas Cichowicz*, sondern von rund 140 wahlberechtigten Zuschauern gestellt. Eine wichtige Neuerung zur Landtagswahl in Schleswig-Holstein betraf die Senkung des aktiven Wahlalters auf 16 Jahre, weshalb die Live-Sendung auch von einer jugendlichen Fragestellerin eröffnet wurde. Allein in Schleswig-Holstein verfolgten rund 270.000 Menschen das Duell vor den Bildschirmen. Dies entsprach einem Marktanteil von 4,6 Prozent und war eine deutliche Steigerung gegenüber dem TV-Duell

23 Vgl. GLES 2013, TV-Duell-Analyse. Befragung, Inhaltsanalyse, Real-Time-Response Messung (Dial und Push Button) ZA5709.

24 Auf Grundlage von 768 Befragten ermittelte das Politbarometer 2016 61,1 Prozent sehr stark beziehungsweise stark Politikinteressierte für Schleswig-Holstein (Datensatz abrufbar unter <https://dbk.gesis.org/dbksearch/sdesc2.asp?no=6888&search=Politbarometer&search2=&field=all&field2=&DB=d&tab=0¬abs=&nf=1&af=&ll=10> (Abruf am 31. August 2017)).

Abbildung 1: Wahlabsichtsfrage zur Landtagswahl



Anmerkung: Antwortverteilung auf die in der Vorbefragung des Debat-O-Meter gestellte Frage „Welcher Partei werden Sie bei der Landtagswahl voraussichtlich ihre Stimme geben?“ (Daten gewichtet nach Geschlecht und Alter; n = 856). Bei den ungewichteten Daten lag der Anteil der CDU bei 30,8 Prozent und derjenige der SPD bei 25,1 Prozent.

Quelle: Eigene Daten.

zur Landtagswahl 2012. Seinerzeit hatten lediglich 150.000 dem TV-Duell zugesehen.²⁵ Mit dem Debat-O-Meter selbst konnten rund 860 Teilnehmer gewonnen werden, so viele wie noch nie bei einer RTR-Analyse eines TV-Duells in Deutschland.

Eine wichtige Frage der Debattenforschung lautet: Wer profitiert stärker von dem TV-Duell? In der Regel ist dies der Herausforderer beziehungsweise die Kandidaten mit einem geringen Bekanntheitsgrad.²⁶ Der Amtsinhaber kann, weil er einen Amtsbonus hat, fast nur verlieren – es sei denn, er liegt stark zurück. Man sieht dies etwa auch bei der Bundestagswahl 2017, bei der Kanzlerin *Angela Merkel* nur ein TV-Duell gegen *Martin Schulz* wollte, der wiederum zwei Debatten mit *Merkel* forderte. Bei der Landtagswahl in Schleswig-Holstein zeigte sich das gleiche Muster: Amtsinhaber und Ministerpräsident *Albig* wollte auch nur eine Debatte, während sein Herausforderer *Günther* zweimal mit dem Ministerpräsidenten diskutieren wollte.²⁷ Empirisch spricht ebenfalls einiges für die These: Die Telefonumfragen zu den TV-Duellen 2009 und 2013 sahen überwiegend die Heraus-

25 Vgl. NDR, Kein klarer Sieger: TV-Duell auf Augenhöhe, 26. April 2017, http://www.ndr.de/nachrichten/schleswig-holstein/landtagswahl_2017/Kein-klarer-Sieger-TV-Duell-auf-Augenhoehe,wahlarena236.html (Abruf am 4. August 2017).

26 Vgl. *Mitchell S. McKinney / Benjamin R. Warner*, Do Presidential Debates Matter?, in: *Argumentation and Advocacy*, 49. Jg. (2013), H. 4, S. 238 – 258; *Marko Bachl*, Die Wirkung des TV-Duells auf die Bewertung der Kandidaten und die Wahlabsicht, in: *ders. / Frank Bretschneider / Simon Ottler* (Hrsg.), a.a.O. (Fn. 3), S. 171 – 198.

27 Vgl. „Albig lehnte Zeitungs-Duell ab“, in: Kieler Nachrichten vom 25. April 2017, S. 11. Im Artikel wird als Grund dafür sogar ausdrücklich der oben beschriebene Effekt zitiert, auch wenn offiziell andere Ursachen angegeben wurden: „Im Landeshaus hört man allerdings auch ein anderes Argument für *Albigs* Entscheidung: Ein solches Duell nutze tendenziell eher dem Herausforderer als dem amtierenden Regierungschef, heißt es aus SPD-Kreisen.“

forderer *Frank-Walter Steinmeier* und *Peer Steinbrück* als Debattensieger. Auch die auf dem Laborexperiment basierenden Studien der GLES weisen in diese Richtung.²⁸

In der Befragung vor Beginn der Debatte hatten sich 43,3 Prozent der Teilnehmer bei der Frage nach der Direktwahl für *Albig* als Ministerpräsident ausgesprochen, 41,6 Prozent für *Günther*; 15,1 Prozent waren noch unentschieden. Dieses Kopf-an-Kopf-Rennen zeigte sich auch an einer ausgeglichenen Bewertung der beiden Kontrahenten, wobei *Günther* (auf einer Skala von -2 bis +2) mit 0,17 im Durchschnitt leicht besser bewertet wurde als *Albig* (0,09).

Betrachtet man bei den Ergebnissen aus der Vorumfrage die Antworten nach der Sonntagsfrage, das heißt die Wahlabsicht, dann ergibt sich ein etwas anderes Bild der Zuschauer: Die CDU lag vor der SPD. Insgesamt waren vor der Debatte jedoch noch fast 28 Prozent der Teilnehmer unentschlossen (vgl. Abbildung 1). Auffallend ist zudem, dass für die Anhänger kleiner Parteien das TV-Duell weniger attraktiv ist. Gemessen an den Vorwahlumfragen liegen die Prozentwerte jeweils deutlich niedriger im Vergleich zu den Wahlumfragen, was vor allem an dem hohen Anteil der Unentschlossenen liegt.

Bei der Frage nach der Parteineigung lagen in der Vorbefragung die Anhänger der CDU (34,7 Prozent) vor jenen der SPD (23,6 Prozent); rund 22 Prozent der Befragten wiesen keine längerfristige Parteiidentifikation auf.

4.2. Unterschiede zwischen der Vor- und der Nachbefragung

Betrachtet man nur die (gewichteten) Teilnehmer, die sowohl die Vor- als auch die Nachbefragung ausgefüllt haben, zeigen sich folgende Unterschiede. In der Vorbefragung wurde *Günther* mit 0,32 (-2 bis +2) leicht positiv bewertet, der Eindruck von *Albig* war mit -0,02 neutral bis durchwachsen. Nach der Debatte hatte sich *Günther* auf 0,58 verbessert, während *Albig* auf -0,18 abrutschte.

Was das Durchsetzungsvermögen in der Diskussion anging, hatte *Günther* zu Beginn einen leichten Favoritenstatus (35,9 Prozent glaubten an seinen Sieg), während 30,3 Prozent von einem Sieg *Albigs* ausgingen und 33,7 Prozent an ein „Unentschieden“ glaubten. Nach der Debatte gaben 60,8 Prozent an, dass *Günther* die Debatte gewonnen habe. *Albig* sahen 32,3 Prozent der Befragten vorn, während sieben Prozent für Unentschieden votierten. *Günther* konnte dabei nicht nur den größeren Teil der Unentschiedenen an sich ziehen, sondern vielfach auch Personen, die zuvor *Albig* als Sieger erwartet hatten.

Ebenfalls deutlich wird der Sieg *Günthers* bei der Veränderung der Wahlabsichten. Hier waren vor der Diskussion die größten Gruppen: CDU (32,6 Prozent), SPD (22,3 Prozent) und Unentschieden (27,8 Prozent).²⁹ Das Unentschieden-Lager schmolz durch die Diskussion deutlich auf 7,5 Prozent ab. Die NDR-Wahlarena half also zahlreichen Zuschauern, sich eine Meinung zu bilden. So kletterte die SPD zwar von 22,3 auf 27,1 Prozent, wurde aber von der stärkeren Wahlentscheidung zugunsten der CDU (von 32,0 auf 49,4 Prozent) überflügelt. Der Rest verteilt sich wieder auf die anderen Parteien und Nichtwähler. Die Zahl der Antwortenden ging jedoch von 856 auf 470 zurück, ein Rückgang von 45 Prozent. Gründe hierfür können nur vermutet werden. So kann kein Interesse an einer Nachbefragung vorgelegen haben, oder ganz praktische Überlegungen können hinter dem Abbruch

28 Vgl. *Thorsten Faas / Jürgen Maier / Michaela Maier* (Hrsg.), a.a.O. (Fn. 2).

29 Die Unterschiede zur Vorbefragung ergeben sich, weil nur Personen ausgewertet wurden, die sowohl an der Vor- als auch an der Nachbefragung teilgenommen haben.

stehen. Was auffällig war, ist jedoch die Rückkehr zahlreicher Teilnehmer nach einem gewissen Zeitraum. Das Zeitfenster für das Offenlassen der Nachbefragung war rund eine Stunde.

Die zuvor Unentschlossenen wechselten in der Nachbefragung etwas häufiger zur CDU: 48 Personen (20,2 Prozent der Unentschlossenen) bekundeten nun eine Wahlabsicht für die CDU und 41 Personen (17 Prozent der Unentschlossenen) für die SPD. Allerdings veränderten nur drei Personen ihre Wahlentscheidung von CDU auf SPD, während 19 nach der Diskussion nicht mehr die SPD, sondern die CDU wählen wollten. Die Wechsler gingen also klar an *Günther*, ebenso die Unentschlossenen.

4.3. Die RTR-Bewertungen

Im Debat-O-Meter wurden 193.038 Bewertungen bei den Echtzeitmessungen während der Debatte registriert, 856 Personen haben dabei mindestens einmal eine Bewertung abgegeben. Im Mittel sind das 225 Bewertungen pro Teilnehmer, im Durchschnitt also knapp vier Eingaben pro Minute und pro Person. Dies ist jedoch nach oben verzerrt, erstens, weil einzelne Teilnehmer außerordentlich aktiv waren und zweitens, weil nicht alle Voten gezählt werden. Wenn mehr als einmal pro Sekunde eine Wertung einging, zählte für beide Bewertungsrichtungen jeweils die stärkere; anschließend wurden für die fragliche Sekunde beide Wertungsrichtungen saldiert. Insofern reduzieren sich die gültigen Werte.

Für *Albig* wurden auf Grund dieser Begrenzung insgesamt 48.224 einzelne Voten abgegeben, für *Günther* waren es 54.124. Wenn man über alle Bewertungen mittelt („++“ = 2, „+“ = 1, „-“ = -1, „--“ = -2), erreicht *Albig* einen Durchschnittswert von -0,76, während *Günther* bei 0,98 steht. Dies stützt die Befunde der Vor- und Nachbefragungen, dass es einen positiven Eindruck vom CDU-Kandidaten bei den Wählern gab. Zudem waren die CDU-Anhänger auch beim Debat-O-Meter etwas aktiver.

Allerdings nutzten die Anwender des Debat-O-Meter dieses Instrument durchaus differenziert: Knapp 61 Prozent haben sowohl *Albig* als auch *Günther* positiv bewertet. Rund 26,5 Prozent haben nur positive Voten für *Günther* und keine einzige positive für *Albig* abgegeben, während 12,5 Prozent der Teilnehmer ausschließlich *Albig* positiv bewerteten. Daher kann man – vor dem Hintergrund der Vorbefragungsergebnisse – von einem gewissen Bias zugunsten der CDU bei den RTR-Voten sprechen. Allerdings lag *Günther* auch innerhalb der Gruppe der „differenziert Abstimmenden“ vorne.

Die Aktivität im Debat-O-Meter stieg im Debattenverlauf immer weiter an, was einerseits daher kommt, dass immer noch eine Reihe „Nachzügler“ in die Diskussion einstiegen, andererseits aber wohl auch damit zusammenhängt, dass die Konfrontation in der zweiten Hälfte der Debatte etwas höher war. Die höchste Aktivität gab es bei den Koalitionsaussagen, vor allem zu *Albigs* Ausweichen auf die Frage nach einem Bündnis mit der Linkspartei („kommen nicht in den Landtag“). Ferner war beim Infrastruktur-Abschnitt (Ausbau der Autobahn A20) und bei der Inneren Sicherheit (bei *Günthers* Bekenntnis zu einem Ausbau der Polizeikapazitäten) eine besonders hohe Aktivität zu verzeichnen.

Wertet man nach dem gleichen Muster, indem positive und negative Bewertungen saldiert werden, die Voten für die SPD- und CDU-Anhänger sowie die Unentschlossenen aus, ergeben sich folgende Zahlen: Die SPD-Wähler bewerteten ihren Kandidaten *Albig* im Durchschnitt mit 0,36, während *Günther* auf -0,09 kommt. Die CDU-Wähler benoteten *Günther* mit 1,85, *Albig* mit -1,80. Die CDU-Wähler bewerteten ihren Kandidaten damit

deutlich besser als die SPD-Wähler ihren Spitzenvertreter; insofern konnte im Duell nur *Günther* seine Reihen schließen. Eine weitere Perspektive ergibt sich, wenn man die jeweiligen Anhänger eines Kandidaten (und damit potenzielle „Parteisoldaten“) ausschließt. Für alle Teilnehmer ohne die SPD-Wähler ergeben sich hier für *Albig* Werte von -1,0 und 1,18 für *Günther* (jeweils Durchschnittswerte). Klammert man die CDU-Wähler aus, so liegen die Durchschnittswerte bei -0,25 für *Albig* und bei 0,52 für *Günther*. Die Unentschiedenen bewerteten *Albig* im Mittel mit -0,48 und *Günther* mit 0,77. *Daniel Günther* kommt damit auch hier generell besser an als der Amtsinhaber – sowohl bei Zuschauern jenseits der eigenen Anhängerschaft als auch bei den Unentschlossenen. Neben der starken Ablehnung von *Albig* bei den CDU-Anhängern fällt auf, dass der amtierende Ministerpräsident sogar von seinen eigenen Anhängern nur sehr zögerliche Unterstützung erhielt. Dagegen wurde *Günther* im gegnerischen Lager nicht besonders abgelehnt.

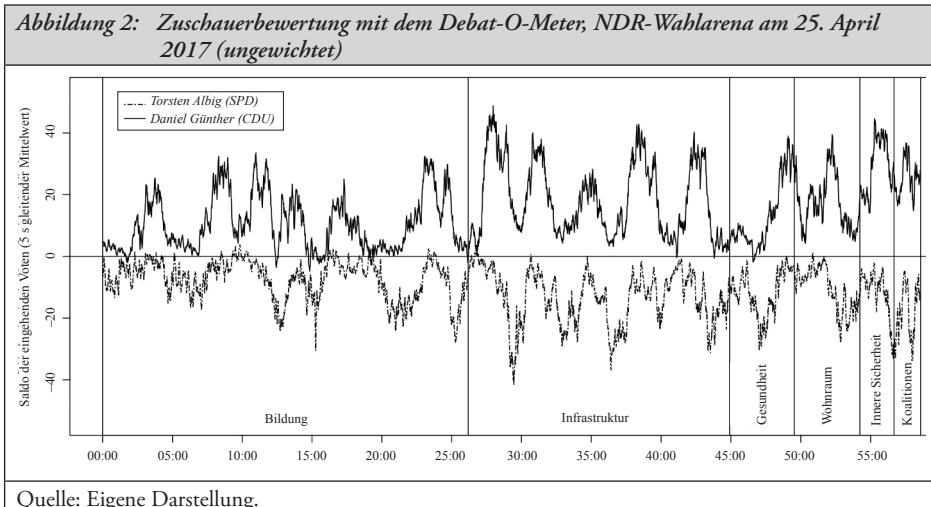
Diese Ergebnisse scheinen im Widerspruch zu einer Vorwahlumfrage der Forschungsgruppe Wahlen zu stehen, die am Tag des TV-Duells veröffentlicht wurde. Danach wurde *Albig* besser als *Günther* bewertet.³⁰ Bemerkenswert war dabei jedoch, dass der Spitzenkandidat der FDP *Wolfgang Kubicki* am besten von allen Politikern in Schleswig-Holstein abschnitt. In der Wahlanalyse der Forschungsgruppe Wahlen, basierend auf einer Umfrage in der Woche vor der Wahl und damit nach der TV-Debatte, kommt *Günther* jedoch auch auf eine etwas bessere Bewertung als *Albig* (1,8 zu 1,7 auf einer Skala von +5 bis -5). Dieser Umschwung in der Kandidatenbewertung bei der Forschungsgruppe Wahlen lässt sich unseres Erachtens vor allem mit dem Sieg *Günthers* in der TV-Debatte erklären.

Betrachtet man die besten und schlechtesten Bewertungen für *Albig*, wurde er während der Debatte bei der Bildungsfrage am besten eingeschätzt, als er an *Günther* folgende Aussage zum Abitur nach acht Jahren richtete: „G8 wurde unter Carstensen eingeführt, das verschweigen Sie.“ Die schlechteste Bewertung erzielte er dagegen, als er sich über die CDU-Vorgängerregierung lustig machte, dass sie keine Fledermäuse zählen könne, und monierte, er habe durch einen Planfeststellungsbeschluss der CDU „Jahre verloren“.

Nach den Echtzeitbewertungen im Debat-O-Meter konnte als bester Moment für *Daniel Günther* die Kritik an der SPD-Landesregierung identifiziert werden: So warf er der Regierung *Albig* mangelnde Kommunikation beim Bau der Autobahn A20 vor („Minister reden nicht miteinander“, „leerer Adlerhorst“). Die schlechteste Note für *Günther* konnte festgestellt werden, als *Albig* ausführte, dass nun 300 mehr Lehrer da seien als vor 2012.

Ein besonderer und unvorhergesehener Punkt der Debatte war jedoch der Vorwurf einer Zuschauerin an *Daniel Günther*, er habe sie als „Ver.di-Schlampe“ bezeichnet. Diese Zuschauerin, die sich im Nachhinein als SPD-Funktionärin aus Flensburg herausstellte, warf dem Herausforderer vor: „Sie haben mich mal Verdi-Schlampe genannt im Landtag“ (im Nachgang konnten in den Landtagsprotokollen hierfür jedoch keine Belege gefunden werden). Auch diese Passage konnte im Debat-O-Meter gemessen werden; sie wird insbesondere dann nachvollziehbar, wenn man die Bewertung vor der Saldierung heranzieht und als eingehende Ströme positiver und negativer Voten darstellt. Nach dem Vorwurf der Zuschauerin blieb der negative Strom eingehender Bewertungen für *Günther* konstant, während der Strom positiver Bewertungen kurzfristig stark nachlässt, wodurch in Summe der Saldo kurzfristig rückläufig wird. Nach wenigen Augenblicken (und *Günthers* Replik) pen-

30 Vgl. Politbarometer Extra 4/17, S. 2. Auf einer Skala von +5 bis -5 erhielt *Albig* +1,5, während *Günther* +1,2 erhielt.



Quelle: Eigene Darstellung.

deln sich beide Ströme wieder auf dem ursprünglichen Niveau ein. Die Teilnehmer haben *Günther* den Vorwurf also offenbar nicht negativ angerechnet, wahrscheinlich weil seine Beteuerung, das nicht gesagt zu haben, auf sie glaubhaft wirkte. Dasselbe zeigt sich, wenn man die SPD-Anhänger allein betrachtet.

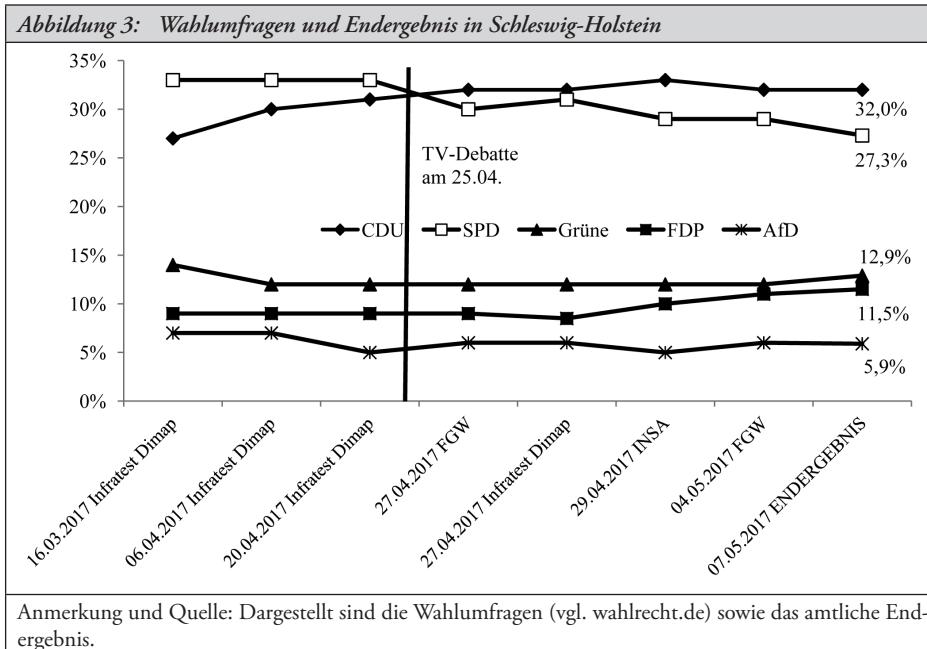
Im Nachgang zur TV-Debatte und in Folge der Diskussion in der Presse hat diese Be- schuldigung wohl eher negativ für *Albig* gewirkt, da die in der Vergangenheit skandalge- prägten Wahlkämpfe in Schleswig-Holstein bis zu diesem Zeitpunkt keine Rolle in der Diskussion gespielt hatten. Insofern konnte hier ein negativer Backlash zu Lasten der SPD identifiziert werden. Verstärkt wurde die Wahrnehmung wohl auch dadurch, dass der NDR im Vorfeld angekündigt hatte, keine politischen Funktionäre in der TV-Debatte als Fragesteller auftreten zu lassen, was auch den NDR in die Kritik geraten ließ.

Die Analyse der Bewertungen nach den in der Diskussion behandelten Politikfeldern (Durchschnittsberechnung aus dem Saldo der positiven und negativen Bewertungen) zeigt, dass *Günther* generell bei allen Themen besser bewertet wurde als *Albig*. Die besten Felder für *Albig* waren Bildung, Gesundheitspolitik und Innere Sicherheit, für *Günther* Innere Si- cherheit, Koalitionsaussagen sowie ebenfalls die Gesundheitspolitik.

Für eine weitere Betrachtungsweise können die RTR-Kurven aggregiert und so der Puls der Debatte nachgezeichnet werden. Die Verlaufskurve der Debatte mit den saldierten Durchschnittsbewertungen wird in Abbildung 2 dargestellt. Dabei fällt auf, dass sich die Schere zwischen *Günther* und *Albig* immer weiter öffnete, der Herausforderer also im Laufe des TV-Duells Zugewinne verbuchen konnte, während der Amtsinhaber an Boden verlor. Insofern verfestigt sich der Sieg *Günthers* auch graphisch.

4.4. Effekte der Debatte

Um die Wirkungen der Debatte zu betrachten, können nicht nur die Angaben aus Vor- und Nachbefragung verglichen werden, denen zufolge *Günther* sowohl bei der Frage nach dem Gewinner als auch bei den Bewertungsfragen vorne lag (siehe oben). Ein weiterer



Plausibilitätscheck ist zudem die Überprüfung durch andere Daten, zum Beispiel die Einschätzung in den Printmedien, in den Sozialen Medien oder auch in den Entwicklungen der Umfragen. Dieses Vorgehen kann man als „nomologische Validierung“ bezeichnen. Es beruht auf der Annahme, dass die externe Variable das gleiche beziehungsweise das erwartete Äquivalent misst wie das zu prüfende Messinstrument³¹: „Nomological validation assesses the performance of indicators in relation to causal hypothesis in order to gain leverage in evaluating measurement validity.“³²

Die „nomologische“ Validierung oder exogene Validierung untersucht die Gültigkeit einer Messung durch den Vergleich mit Referenzgrößen außerhalb des Messkonstruktions selbst. Vergleicht man die in 2017 durch die Umfrageinstitute erhobenen Vorwahlumfragen und das Endergebnis in ihrer zeitlichen Dimension (vgl. Abbildung 3), fallen der Zuwachs der CDU und das Abfallen der SPD nach der Debatte besonders auf. Alle anderen Parteien blieben in der Reihenfolge gleich und veränderten sich auch sonst nur geringfügig. Für die Frage, wen die Bürger direkt zum Ministerpräsidenten wählen würden, konnte man aus den Umfragen der Meinungsforschungsinstitute ebenfalls einen deutlichen Swing ableSEN: So verlor *Albig* als gewünschter Ministerpräsident verglichen mit den Umfragen vor der Debatte zwölf Prozentpunkte, während *Günther* elf Prozentpunkte hinzugewann.

31 Zur Erläuterung des Konzepts der nomologischen Validierung siehe *Richard A. Zeller / Edward G. Carmines* (Hrsg.), *Measurement in the Social Sciences. The Link between Theory and Data*, Cambridge 1980, S. 78 – 84.

32 *Robert Adcock / David Collier*, *Measurement Validity. A Shared Standard of Qualitative and Quantitative Research*, in: *American Political Science Review*, 95. Jg. (2001), H. 3, S. 529 – 546, S. 542.

4.5. Das Interview *Albigs* in der Illustrierten „Bunte“

Neben der TV-Debatte, die auch einen großen Nachhall in der tagespolitischen Presse fand, spielte gleichzeitig noch ein Interview *Albigs* vom 20. April 2017 in der Illustrierten „Bunte“ eine gewisse Rolle. In dem Interview, in dem er auch auf seine neue Lebenspartnerin sowie seine geschiedene Ehefrau einging, machte *Albig* verschiedene Aussagen, die sich im Nachhinein als kontraproduktiv für den Wahlkampf herausstellten beziehungsweise von zahlreichen Medien als Erklärungsfaktor für die Niederlage herangezogen wurden. So äußerte er, dass sich sein Leben schneller als das seiner Ehefrau entwickelt habe. Ein Satz, der später in der Presse besonders diskutiert wurde, lautete: „Wir hatten nur noch ganz wenige Momente, in denen wir uns auf Augenhöhe ausgetauscht haben“. Der Tenor in der Presse nach der Wahl war unisono, dass dieses Interview *Albig* vor allem bei den Wählerinnen geschadet habe. Dies deckt sich auch mit den letzten verfügbaren Daten aus der amtlichen Wahlstatistik aus dem Jahre 2005. Betrug die Differenz zwischen weiblichen und männlichen SPD-Wählern 2005 noch fast sechs Prozentpunkte (41,8 zu 36,2) so ist diese Differenz 2017 laut der Wahltagbefragung von Infratest dimap auf nur einen Prozentpunkt zusammengeschmolzen.

Bei der Debat-O-Meter Auswertung zeigt sich zwar, dass *Albig* im Laufe des TV-Duells relativ stark bei den Frauen verlor, jedoch noch stärker bei den Männern. *Günther* dagegen konnte besonders bei weiblichen Teilnehmern punkten. Dies sagt etwas über die Effekte bei den Teilnehmern der RTR-Messung aus, jedoch nichts über die Nachberichterstattung und andere intervenierende Faktoren wie etwa das Bunte-Interview *Albigs*. Allerdings wurde darüber erst ausführlich etwa nach der Wahl berichtet, so dass die Wirkungen eher begrenzt sein dürften.

5. Quantitative Analysen

Den wichtigsten Bestimmungsfaktoren für den Debattensieg kann auch mit multivariaten Analysemethoden näher gekommen werden. Dazu werden die in den Befragungen gewonnenen Daten mit Geschlecht und Alter der Rezipienten gewichtet. Als statistische Untersuchungsmethode wird die binäre logistische Regression verwendet, die eine dichotome abhängige Variable analysiert. Als abhängige Variable wurde dabei kodiert, wer in der Nachbefragung *Torsten Albig* beziehungsweise *Daniel Günther* als Sieger der Debatte in der Wahlarena gesehen hatte. Die Unentschiedenen – etwa sieben Prozent in der Nachbefragung – werden nicht weiter betrachtet. Die Tabellen 2 und 3 präsentieren die binären logistischen Regressionen für den Erfolg *Albigs* beziehungsweise *Günthers* aus Sicht der Befragten.

Als Einflussgrößen werden erstens Faktoren der Parteidifferenz (Parteidifferenz und Wahlabsicht für SPD beziehungsweise CDU) als unabhängige Variable verwendet. Bei der Parteidifferenz (wahlweise wurde auch noch die Wahlabsicht verwendet) wird ebenfalls ein positiver Zusammenhang mit der Einschätzung, wer die Debatte gewonnen hat, erwartet. Hypothese 1 lautet daher: „Liegt eine Parteidifferenz für die SPD vor, steigt die Wahrscheinlichkeit, *Torsten Albig* als Sieger zu sehen.“ Analog gilt (H2): „Liegt eine Parteidifferenz für die CDU vor, nimmt die Wahrscheinlichkeit zu, *Daniel Günther* als Sieger zu sehen.“

Ein weiterer Erklärungsfaktor ist die Einschätzung der beiden Diskutanten in der Vorbefragung. Dazu wird die Hypothese (H3) aufgestellt, dass ein Kandidat umso eher als Sieger gesehen wird, je besser seine Bewertung vor der Debatte war. Die Bewertung wurde dabei für beide Kandidaten getrennt auf einer 5er-Skala von -2 bis +2 erfasst.

Darüber hinaus wird in der Regressionsmodellierung der Bedeutung von Emotionen Rechnung getragen. In den vergangenen Jahren wurde hier in der Literatur eine immer stärker werdende Diskussion geführt³³, wonach Emotionen die Bewertung von Kandidaten in Debatten stark beeinflussen. In der Vorbefragung wurden die Teilnehmer daher nach ihrem momentanen emotionalen Zustand gefragt und konnten für die Gefühle Wut, Verachtung, Furcht, Ekel, Freude, Traurigkeit, Überraschung und Gleichgültigkeit auf einer Skala von -2 bis +2 angeben, wie sehr diese ihre momentane Verfassung beschreiben. Die Hypothese (H4) lautet dabei, dass der Amtsinhaber eher abgestraft wird, wenn Zuschauer wütend sind: „Je wütender der Befragte, desto schlechter ist die Bewertung für den Amtsinhaber.“ Korrespondierend dazu sollte der Herausforderer eher von wütenden Befragten in der Debatte profitieren. Hypothese 5 lautet daher: „Je wütender der Befragte, desto eher wird der Herausforderer positiv bewertet.“ Positive Emotionen wie etwa Freude oder Glück dürften dagegen eher dem Amtsinhaber nützen (H6), während wir davon ausgehen, dass sie keinen Effekt auf die Bewertung des Herausforderers haben (H7).

Als zentrale unabhängige Variable wurde noch die Echtzeitbewertung für die einzelnen Kandidaten saldiert, das heißt positive gegen negative Bewertungen aufgerechnet und über das Duell gemittelt. Die Hypothese H8 lautet dazu: „Je mehr ein Zuschauer für einen Kandidaten in der Debatte wertet, desto eher sieht er ihn auch als Gewinner der Debatte an.“ Die Vermutung ist dabei, dass die Echtzeitreaktionen der Zuschauer direkt auf die Bewertungen der Kandidaten als Debattensieger wirken.

Die Ergebnisse in Tabelle 2 zeigen für den perzipierten Sieg des Ministerpräsidenten *Torsten Albig*, dass die getesteten Kontrollvariablen Alter und Geschlecht insignifikant sind, das heißt die Bewertung wird von ihnen nicht beeinflusst. Es sind jedoch signifikante Einflüsse der Parteiidentifikation sowie der Vorabeinschätzung und -bewertung von *Albig* beziehungsweise *Günther* festzustellen. In Gleichung 3 zeigt sich allerdings, dass die Parteiidentifikation (dies würde auch für die Sonntagsfrage gelten) insignifikant wird, weil die saldierten Echtzeitbewertungen die weitaus stärkste und prägendste Variable in den logistischen Regressionen sind. Insofern hat es der „Wahrnehmungsfilter“ der Parteiidentifikation also nicht geschafft, den Eindruck aus dem Duell zu drehen. Die Erklärungskraft von Modell 3 ist für logistische Regressionen sehr hoch, Nagelkerkes R² beträgt hier 0,83. Auch die anderen Modelle in Gleichung 1 und 2 sind gemessen an den Pseudo-R² von Cox und Snell sowie Nagelkerke sehr hoch. Diese beiden Kennziffern dienen zur Interpretation der Erklärungsgüte der Modelle, die verglichen mit anderen logistischen Modellen sehr gut ausfällt.

33 Vgl. *Ines C. Vogel / Lukas Otto*, Die Bedeutung von Emotionen für die Rezeption der TV-Debatte 2013, in: *Thorsten Faas / Jürgen Maier / Michaela Maier*, a.a.O. (Fn. 2), S. 87 – 103; *Jaeho Cho / Yerheen Ha*, On the Communicative Underpinnings of Campaign Effects: Presidential Debates, Citizen Communication, and Polarization in Evaluations of Candidates, in: *Political Communication*, 29. Jg. (2012), H. 2, S. 184 – 204; *Thorsten Faas / Judith Schliephake*, Affective Intelligence und Debattenrezeption: Eine experimentelle Studie zum Einfluss von Emotionen auf die Informationsverarbeitung und Entscheidungsfindung am Beispiel eines Fernsehduells, in: *Politische Psychologie*, 4. Jg. (2015), H. 2, S. 252 – 274.

Tabelle 2: Bestimmungsfaktoren für die Beurteilung Albigs als Sieger der TV-Debatte

	Modell 1	Modell 2	Modell 3
Alter (4 Kategorien, aufsteigend)	-0,17 (0,13)	-0,19 (0,14)	-0,19 (0,21)
Geschlecht (0 = männlich, 1 = weiblich)	-0,12 (0,28)	-0,27 (0,29)	-0,16 (0,44)
Durchschnittliche Bewertung <i>Torsten Albig</i> in der Vorbefragung (-2 bis +2)	1,36** (0,15)	1,35* (0,15)	0,87** (0,24)
Parteiidentifikation für die SPD	0,69* (0,31)	0,54 (0,32)	-0,57 (0,47)
Emotion „Wut“ in der Vorbefragung (-2 bis +2)		-0,11 (0,15)	
Emotion „Freude“ in der Vorbefragung (-2 bis +2)		-0,23 (0,16)	
Saldo der Echtzeitbewertungen (RTR-Messungen) für <i>Torsten Albig</i>			2,78** (0,37)
Konstante	-0,95* (0,43)	-0,79 (0,49)	-2,03* (0,77)
-2*Log-Likelihood	337,79	323,43	140,35
Cox & Snell-R ²	0,37	0,35	0,59
Nagelkerke-R ²	0,52	0,50	0,83
Fallzahl	400	378	381

Anmerkungen: * Signifikant auf 95 %-Niveau, ** Signifikant auf 99 %-Niveau. Dargestellt sind in der ersten Zeile die Regressionsparameter „b“. Werte in Klammern geben Standardfehler an. Abhängige Variable: Antwort „*Torsten Albig*“ auf die Frage „Alles in allem, wer hat Ihrer Meinung nach in der Diskussion am besten abgeschnitten?“

Quelle: Eigene Berechnung.

Die Überprüfung auf Multikollinearität ergibt lediglich bei der Sonntagsfrage („Wahlabsicht“) sowie der Parteiidentifikation bedenkliche Werte, was aus theoretischer Sicht nicht verwundert. Entsprechend wurde entschieden, nur eine der beiden Variablen in den Gleichungen 1 bis 3 (je in den Tabellen 2 und 3) zu verwenden. Bei der Vorhersagegenauigkeit zeigt sich, dass mit dem dritten Modell immerhin 93,4 Prozent aller Fälle korrekt vorhergesagt würden – auch dies ist ein bemerkenswerter hoher Wert.

Die Hypothesen für die Emotionen finden in keiner Richtung Unterstützung. Weder der Herausforderer noch der Amtsinhaber profitierten von den vor dem Duell vorherrschenden emotionalen Zuständen der Zuschauer. Zwar stimmen die Vorzeichen für die Variable Wut (negativ für *Albig*, positiv für *Günther*) mit den Erwartungen überein, die Koeffizienten sind jedoch insignifikant.

Die Erklärungsmodelle für die Einschätzung *Günthers* als Debattensieger sind im Wesentlichen deckungsgleich mit den Befunden für *Albig* (vgl. Tabelle 2). Insgesamt kann man feststellen, dass sowohl in den Befunden für *Albig* als auch für *Günther* die RTR-Bewertungen, die Parteiidentifikation sowie die Einschätzung der beiden Diskutanten von ihrer jeweiligen Anhängerschaft die prägenden Variablen sind. Am stärksten tragen dabei die Echtzeitmessungen zur Erklärung bei, die eine Odds Ratio von über 16 (nicht in Tabel-

Tabelle 3: Bestimmungsfaktoren für die Beurteilung Günthers als Sieger der TV-Debatte

	Model 1	Model 2	Model 3
Alter (4 Kategorien, aufsteigend)	-0,01 (0,12)	-0,26 (0,12)	-0,07 (0,19)
Geschlecht (0 = männlich, 1 = weiblich)	0,18 (0,25)	0,31 (0,26)	-0,20 (0,37)
Durchschnittliche Bewertung für <i>Daniel Günther</i> in der Vorbefragung (-2 bis +2)	1,12** (0,13)	1,09** (0,13)	0,60** (0,19)
Parteiidentifikation für die CDU	1,04** (0,32)	1,06** (0,33)	0,66 (0,46)
Emotion „Wut“ in der Vorbefragung (-2 bis +2)		0,10 (0,13)	
Emotion „Freude“ in der Vorbefragung (-2 bis +2)		-0,02 (0,14)	
Saldo der Echtzeitbewertungen (RTR-Messungen) für <i>Daniel Günther</i>			2,19** (0,25)
Konstante	-0,11 (0,39)	-0,54 (0,44)	-1,0 (0,65)
-2*Log-Likelihood	397,43	380,19	205,85
Cox & Snell-R2	0,36	0,34	0,57
Nagelkerke-R2	0,48	0,46	0,77
Fallzahl	399	375	382

Anmerkungen: * Signifikant auf 95 %-Niveau, ** Signifikant auf 99 %-Niveau. Dargestellt sind in der ersten Zeile die Regressionsparameter „b“. Werte in Klammern geben Standardfehler an. Abhängige Variable: Antwort „*Daniel Günther*“ auf die Frage „Alles in allem, wer hat Ihrer Meinung nach in der Diskussion am besten abgeschnitten?“
Quelle: Eigene Berechnung.

le 3 aufgeführt) aufweisen. Odds bedeutet dabei Chancenverhältnis (man kennt den Begriff eher aus Wetten), was in diesem Fall ein Verhältnis von 16 zu 1 für die Bewertungen von *Albig* gegenüber keinen Bewertungen angibt. Insgesamt sind die binär logistischen Regressions in Tabelle 3 wieder sehr erkläzungskräftig.

Als weiterer Test der Anpassungsgüte wurde noch der Hosmer-Lemeshow-Test durchgeführt. Dieser überprüft die Differenz zwischen den beobachteten und erwarteten Differenzen für zufällig aufgeteilte Gruppen. Für alle sechs Modelle wird dabei bestätigt, dass eine gute Modellanpassung vorliegt.

6. Echtzeitmessungen von Debatten ermöglichen neuartige Analysen

Mit dem Debat-O-Meter wurde erstmals in der Bundesrepublik Deutschland eine Echtzeitbewertung eines TV-Duells in großem Maßstab außerhalb eines experimentellen Labor-settings vorgenommen. Mit 856 Teilnehmern konnte über Medienpartner als bisher unerprobten Rekrutierungsweg eine hinreichend große Zahl von Befragten gewonnen werden, die die Echtzeitbewertung vornahmen. Gekoppelt mit einer Vor- und Nachbefragung lag damit umfangreiches Datenmaterial vor, das es möglich machte, die Debatte am 25. April

2017 zwischen dem schleswig-holsteinischen Ministerpräsidenten *Torsten Albig* (SPD) und seinem Herausforderer *Daniel Günther* (CDU) zu analysieren.

Die Debattenforschung in Deutschland, aber auch in den USA hat gezeigt, dass nicht nur die Debatte selbst wichtig ist, sondern auch die Nachberichterstattung die Wahlentscheidung beeinflussen kann.³⁴ Um jedoch diese Effekte herauszuarbeiten, bedürfte es einer detaillierten Medienanalyse, inklusive den Fokus auf Soziale Medien und Häufigkeit der Mediennutzung. Für Schleswig-Holstein kann daher bestenfalls spekuliert werden, da diese für die Landtagswahl nicht vorliegt. Hier könnte auch eine Gefahr des Debat-O-Meter liegen, in dem die Ergebnisanalysen und die Veröffentlichungen in den Medien eine nicht-intendierte Wirkung entfalten.

Im Extremfall wären auch Szenarien denkbar, dass Parteianhänger sich als Wähler einer anderen Partei ausgeben. Durch eine Änderung der Parteipräferenz könnte dann die andere Partei „schlechter“ aussehen. Kontrollieren kann man dies durch ein Gegenspiegeln der Bewertungen mit der Echtzeitmessung. Für die Wahl in Schleswig-Holstein gab es keine Belege für ein solches Verhalten.

Die vier eingangs gestellten Forschungsfragen dieses Beitrags konnten mit Hilfe der gewonnenen Daten eindeutig beantwortet werden. Der Gewinner der Debatte war Herausforderer *Daniel Günther* von der CDU. Er konnte vor allen Dingen bei Themen wie der Infrastruktur, etwa dem Ausbau der Autobahn A20, bei der Koalitionsaussage und der Inneren Sicherheit punkten. Ministerpräsident *Albig* hatte seine besten Aussagen beim Thema Bildung. Betrachtet man die Faktoren, die den Sieger der Debatte festlegen, kristallisierten sich in der empirischen Analyse besonders die Parteiidentifikation, die Häufigkeit der Bewertungen mit dem Debat-O-Meter, aber auch die Vorbewertung von *Albig* beziehungsweise *Günther* als relevante Einflüsse heraus. Es zeigt sich aber auch, dass Teilnehmer, die sich als „wütend“ bezeichneten, *Albig* nicht besonders abstrafen. Bei „freudigen“ Teilnehmern war ein positiv wirkender Effekt aber ebenfalls nicht sichtbar. Und die Unentschiedenen sowie die Parteiwechsler orientierten sich in höherem Maße zu *Daniel Günther* als zu *Torsten Albig*.

Insgesamt hat sich also die Alltagstauglichkeit des Debat-O-Meters als Echtzeitmessinstrument bewiesen. Die Vorteile der Web-Anwendung gegenüber traditionellen RTR-Methoden liegen dabei klar auf der Hand: Es besteht eine potentiell sehr große räumliche Repräsentation, da es keine Gebundenheit an Labore oder Hörsäle für die Datenerhebung gibt. Zudem ist es möglich, über Quotierungen oder Gewichtung einen möglichen Bias im Datensatz auszugleichen. Ein weiterer deutlicher Vorteil des Debat-O-Meters liegt in der Geschwindigkeit der Analyse im Nachgang der Debatten in Kombination mit den für die Rekrutierung großer Fallzahlen notwendigen Medienpartnern.

34 Vgl. *Tobias Claßen / Lisa Guckel / Marvin Kuhn*, Das TV-Duell, in: *Oliver Strijbis / Kai-Uwe Schnapp* (Hrsg.), Aktivierung und Überzeugung im Bundestagswahlkampf 2013, Wiesbaden 2015, S. 159 – 176; *Daron R. Shaw / Brian D. Roberts*, Campaign Events, the Media and the Prospect of Victory: The 1992 and 1996 US Presidential Elections, in: *British Journal of Political Science*, 30. Jg. (2000), H. 2, S. 259 – 289.