

**SHAPING.**

**ONLINE.**

**RESEARCH.**

**25<sup>th</sup> GENERAL ONLINE  
RESEARCH CONFERENCE**  
20 - 22 SEPTEMBER IN KASSEL



**GOR**  
organized by **DGOF**

BELLA STRUMINSKAYA, SIMON KÜHNE, FLORIAN KEUSCH, OTTO HELLWIG,  
STEFAN OGLESBY, ALEXANDRA WACHENFELD-SCHELL (EDS.)

## 25TH GENERAL ONLINE RESEARCH CONFERENCE PROCEEDINGS KÖLN 2023



ALL RIGHTS RESERVED. NO PART OF THIS PUBLICATION MAY BE REPRODUCED,  
STORED IN A RETRIEVAL SYSTEM, OR TRANSMITTED, IN ANY FORM OR BY ANY MEANS,  
WITHOUT THE PRIOR PERMISSION IN WRITING OF THE PUBLISHER:

ISBN 978-3-9822985-3-5

BELLA STRUMINSKAYA, SIMON KÜHNE, FLORIAN KEUSCH, OTTO HELLWIG,  
STEFAN OGLESBY, ALEXANDRA WACHENFELD-SCHELL (EDS.)

GERMAN SOCIETY FOR ONLINE RESEARCH  
DEUTSCHE GESELLSCHAFT FÜR ONLINE-FORSCHUNG (DGOF) E.V.  
[WWW.DGOF.DE](http://WWW.DGOF.DE)

EDITING: Anna Hristova, Anna-Lena Kemper  
LAYOUT AND TYPESET: Nicole Propach, Köln

04

ORGANIZATION



05

INTERNATIONAL  
BOARD



06

GREETINGS  
FROM DGOF



07

ABOUT DGOF



08

PORTRAITS OF  
THE BOARD



10

GREETINGS FROM  
THE LOCAL PARTNER



11

SPONSORS AND  
ORGANIZER



12

PROGRAMME  
OVERVIEW



14

WORKSHOPS



17

KEYNOTES



18

GOR BEST PRACTICE  
AWARD / ABSTRACTS



21

GOR  
POSTER AWARD



22

GOR  
THESIS AWARD



23

DGOF  
BEST PAPER AWARD



34

ABSTRACTS



## DGOF BOARD



**DR. OTTO HELLWIG**  
Chairman DGOF Board  
Bilendi & respondi, Germany



**PROF. DR. FLORIAN KEUSCH**  
University of Mannheim,  
Germany



**PROF. DR. SIMON KÜHNE**  
Bielefeld University, Germany



**PROF. DR. BELLA STRUMINSKAYA**  
Utrecht University, The Netherlands



**DR. STEFAN OGLESBY**  
data IQ AG, Switzerland



**ALEXANDRA WACHENFELD-SCHELL**  
GIM Gesellschaft für  
Innovative Marktforschung,  
Germany

## CONFERENCE CHAIR

**DR. OTTO HELLWIG**

Chairman DGOF Board, Bilendi & respondi, Germany

## PROGRAMME COMMITTEE

**PROF. DR. BELLA STRUMINSKAYA**

(Programme Chair), DGOF Board & Utrecht University

**PROF. DR. SIMON KÜHNE**

(Vice Programme Chair), DGOF Board & Bielefeld University

**DR. OTTO HELLWIG**

(Track A), DGOF Board, Bilendi & respondi, Germany

**DR. CARINA CORNESSE**

(Track A), DIW Berlin

**PROF. DR. FLORIAN KEUSCH**

(Track B), DGOF Board & University of Mannheim

**ANNA RYSINA**

(TRACK C), KANTAR PUBLIC

**DR. BERNHARD CLEMM VON HOHENBERG**

(Track C), GESIS – Leibniz Institute of the Social Sciences

**YANNICK RIEDER**

(GOR Best Practice Award 2023 Competition),  
The Janssen Pharmaceutical Companies of Johnson & Johnson

**ALEXANDRA WACHENFELD-SCHELL**

(GOR Best Practice Award 2023 Competition),  
DGOF Board & GIM Gesellschaft für Innovative Marktforschung

**DR. OLAF WENZEL**

(GOR Thesis Award 2023 Competition), Wenzel Marktforschung

**DR. FREDERIK FUNKE**

(GOR Workshops), datenmethoden.de & LimeSurvey GmbH

## DGOF OFFICE

**ANNA-LENA KEMPER, ANNA HRISTOVA**

## SUPPORT

**ARMIN DOBRATZ**

## LOCAL PARTNER

**University of Kassel, Germany**

## INTERNATIONAL BOARD

- DR. CHRISTOPHER ANTOUN** University of Maryland  
**DR. FRANCES M. BARLAS** Ipsos Public Affairs  
**KAREN BLANKE** Destatis  
**ORIO J. BOSCH** The London School of Economics and Political Science  
**DR. ALEXANDRU CERNAT** University of Manchester  
**DR. BERNHARD CLEMM VON HOHENBERG** GESIS Leibniz  
 Institute for the Social Sciences  
**DR. CARINA CORNESSE** German Institute for Economic Research  
 (DIW Berlin) and Research Institute Social Cohesion (RISC)  
**PROF. MARCEL DAS** Centerdata  
**PROF. EDITH DE LEEUW** Utrecht University  
**LISA DUST** Facts and Stories  
**BEAT FISCHER** intervista  
**DR. FREDERIK FUNKE** Dr. Funke SPSS & R trainings & LimeSurvey GmbH  
**DR. TOBIAS GUMMER** GESIS Leibniz Institute for the Social Sciences  
**DR. GEORG-CHRISTOPH HAAS** Institute for Employment Research  
**DR. CAROLA HÄRING** GIM Gesellschaft für innovative Marktforschung mbH  
**NIKLAS HAUPT** Mios GmbH  
**DR. OTTO HELLWIG** Bilendi & respondi / DGOF  
**FRANK HEUBLEIN** Frank Heublein M3 Services  
**DR. JAN KAREM HÖHNE** DZHW; University of Hannover  
**PROF. ANNETTE HOXTELL** VICTORIA  
 International University of Applied Sciences  
**DR. WOJCIECH JABLONSKI** Netherlands Interdisciplinary  
 Demographic Institute  
**BIRGIT JESKE** infas GmbH  
**FLORIAN KEUSCH** University of Mannheim  
**AIGUL KLIMOVA** National Research University Higher School of Economics  
**DR. FRANK KNAPP** Psyma  
**FABIENNE KRAEMER** GESIS Leibniz Institute for the Social Sciences  
**DR. THOMAS KRÜGER** Umfragezentrum Bonn –  
 Prof. Rudinger GmbH (uzbonn GmbH)  
**PROF. SIMON KÜHNE** Bielefeld University  
**DR. TANJA KUNZ** GESIS Leibniz Institute for the Social Sciences  
**ANDRÉ LANG** Insius  
**PROF. DR. BETTINA LANGFELDT** University of Kassel,  
**ELLEN LAUPPER** Swiss Federal  
 Institute for Vocational Education and Training  
**DANIIL LEBEDEV** HSE University  
**PROF. JI-PING LIN** Academia Sinica  
**DR. OLIVER LIPPS** FORS  
**FRANK LÜTTSCHWAGER** EARSandEYES  
**DR. EMANUEL MAXL** Context-Research  
**PROF. NATALJA MENOLD** TU Dresden  
**PROF. GUSTAVO S. MESCH** University of Haifa  
**PROF. JOCHEN MUSCH** University of Düsseldorf  
**DR. STEFAN NIEBRÜGGE** Exo-Q GmbH  
**MAREIKE OEHL** September Strategie & Forschung GmbH  
**DR. STEFAN OGLESBY** data IQ AG  
**DR. TOBIAS RETTIG** University of Mannheim  
**YANNICK RIEDER** Janssen-Cilag GmbH  
**DR. EIKE MARK RINKE** University of Leeds  
**DR. JOSS ROSSMANN** GESIS Leibniz Institute for the Social Sciences  
**ANNA RYSINA** Kantar Public  
**CAMILLA SALVATORE** University of Amsterdam  
**DR. INES SCHAURER** City of Mannheim  
**DR. KAZUHIKO SHIBUYA** Globiz Professional University  
**DR. HENNING SILBER** GESIS Leibniz Institute for the Social Sciences  
**ROBIN SPICER** Bilendi & respondi  
**DR. MARKUS STEINBRECHER** Bundeswehr Center  
 for Military History and Social Sciences  
**PIRMIN STÖCKLE** Vocatus AG  
**PROF. BELLA STRUMINSKAYA** Utrecht University  
**OLIVER TABINO** Q I Agentur für Forschung GmbH  
**DR. ASLI TELLI-AYDEMIR** University of Witwatersrand  
**DORIAN TSOLAK** Bielefeld University  
**PROF. ANNA-SOPHIE ULFERT-BLANK** Eindhoven University of Technology  
**ALEXANDRA WACHENFELD-SCHELL** GIM  
 Gesellschaft für Innovative Marktforschung  
**ANJA WENKE** anwema / anja wenke marktforschung  
**DR. ALEXANDER WENZ** University of Mannheim  
**DR. OLAF WENZEL** Wenzel Marktforschung  
**ZAZA ZINDEL** Bielefeld University

# DEAR GUESTS OF THE GOR 23!

**25 YEARS OF DGOF AND THE 25<sup>TH</sup> EDITION OF GOR  
WELCOME TO THE 25TH EDITION OF THE GENERAL  
ONLINE RESEARCH CONFERENCE AND THE 25TH  
ANNIVERSARY OF THE DGOF. FOR THIS REASON, WE  
ARE VERY HAPPY TO HOST THIS YEAR'S GOR IN COOP-  
ERATION WITH THE UNIVERSITY OF KASSEL.**

It all started in 1997. In that year, members of the German Internet Research List met in Cologne for the first time for a two-day conference. It was addressed to all those interested in recent developments in the field of Internet research. The focus of the GOR 97 was on data collection and analysis procedures that require the Internet as a medium. Emphasis was to be placed on survey research, conducting interviews and evaluating log file data.

One year later, in 1998, the same pioneers of online research founded the German Society for Online Research (DGOF). Since 1997, with exceptions in 2000 and 2004, the GOR has opened its doors annually to an ever-growing community of online researchers. Due to the two years without GOR, this year in Kassel we can celebrate not only 25 years of DGOF, but also the 25th edition of GOR. What a wonderful reason to meet in Kassel for GOR 23.

As in previous years, we have a great conference programme lined up for you including keynotes, discussions, presentations, awards, posters, workshops and much more. You can choose between four simultaneous conference tracks: Track A covers "Survey Research: Advancements in Online and Mobile Web Surveys". Track B deals with "Data Science: From Big Data to Smart Data". Track C features "Politics, Public Opinion and Communication" and Track D covers "Digital Methods in Applied Research". In addition, we have three award competitions: i) the GOR Best Practice Award 2023 for the best practice study from applied online research, ii) the GOR Thesis Award 2023 for the best thesis in online research, iii) the GOR Poster Award 2023 for the best poster of the conference. The DGOF Best Paper Award 2023 for the best paper in online research will also be awarded at GOR.

Our keynotes this year take different looks at the development of our profession. On Thursday, 21 September, Steffen Schmidt, Director Marketing Science & Agile Insights at LINK Marketing Services AG will present "The Evolution of Marketing Research: From Insights 1.0 to 5.0". And on Friday, 22 September, Prof. Arie Kapteyn, professor of economics at University of Southern California will talk about "Population Representative Surveys, Wearables, and Contextual Data".

On the pre-conference day, Wednesday 20 September, three workshops will take place and on Wednesday evening we will meet for the GOR 23 Get Together with drinks and snacks. Thursday evening is GOR party time. The GOR Get-Together will take place at the "Caricatura Bar" (Hauptbahnhof Kassel, Rainer-Dierichs-Platz 1, 34117 Kassel). Kassel's Restaurant & Lifestyle Bar "Bolero" (Schöne Aussicht 1A, 34117 Kassel) opens its doors and terrace for the GOR Party on Thursday at 8pm.



We are particularly grateful for the enthusiastic support of and collaboration with our partners at University Kassel: Prof. Dr. Bettina Langfeldt and her team. We would also like to thank our sponsors and media partners. And, of course, a big THANKS to you, the conference participants, presenters, and speakers at this event!

**HAVE A GREAT TIME AT THE GENERAL ONLINE  
RESEARCH CONFERENCE 2023!**

**DR. OTTO HELLWIG**

# DEUTSCHE GESELLSCHAFT FÜR ONLINE-FORSCHUNG E.V.

ONLINE RESEARCH IS A DYNAMIC, INNOVATIVE FIELD, WITH CONSTANTLY EMERGING CHALLENGES AS WELL AS OPPORTUNITIES FOR RESEARCH AND PRACTICE.

The German Society for Online Research (Deutsche Gesellschaft für Online Forschung) (DGOF) is a modern, innovative association, which has focused on the interests of the actors in the field of online research since its establishment in 1998.

It is the association's goal to be the leader in this field. DGOF seeks to bridge different research fields (such as sociology, psychology, political science, economics, market and opinion research, data science) using online research methods and facilitates the transfer between academic research and the industry.

DGOF campaigns for the establishment and the development of online research as well as the interests of online researchers in Germany. Online research ranges from online based data collection methods (e.g., web surveys in online panels); to mobile research with smartphones, tablets, and wearables; to the collection and analysis of social media data, administrative data, data from passive measurements, and other big data sources.

DGOF organizes the General Online Research (GOR) conference and the Research Plus event series which support professional and collegial exchanges between researchers and practitioners across academia and the industry. By bringing together scientific findings, commercial needs, and practical applications for best practices, DGOF provides a sustainable input for further developments in online research.

**CHANGE THROUGH INNOVATION  
IS A KEY CHARACTERISTIC OF  
OUR RESEARCH FIELD. DGOF IS A  
FACILITATOR FOR THIS CHANGE:**

## 1 DGOF MEANS DEVELOPMENT:

Online research is more than just web surveys. We constantly expand our portfolio and our expertise with the development, encouragement, and establishment of innovative digital methods, passive measurement, and big data methods. In addition, we focus on the relationship between the Internet and society.

## 2 DGOF CONNECTS:

We are a bridge between different research disciplines and across commercial applications.

## 3 DGOF IS DIVERSE:

We support our members' interests, for the dissemination of knowledge, for exchange, and for discussion, as well as for the establishment and implementation of scientific standards.

## 4 DGOF IS INNOVATIVE:

We are a facilitator of new issues such as big data and data science.

## 5 DGOF IS DISRUPTIVE:

We support change. It is our practice to foster acceptance for new methods in research, and we are always on the lookout for new developments.



DEUTSCHE GESELLSCHAFT FÜR ONLINE-FORSCHUNG – DGOF E. V.  
GERMAN SOCIETY FOR ONLINE RESEARCH  
HUHNSGASSE 34B  
50676 COLOGNE (GERMANY)

PHONE: +49 (0)221-27 23 18 180  
E-MAIL: OFFICE@DGOF.DE

WWW.DGOF.DE • WWW.GOR.DE



## DR. OTTO HELLWIG



**OTTO HELLWIG WAS THE CEO OF RESPONDI AG FROM ITS FOUNDATION IN 2005 UNTIL THE SALE OF THE RESPONDI GROUP TO BILENDI IN 2021. SINCE THEN HE HAS BEEN CORPORATE INTEGRATION DIRECTOR OF BILENDI & RESPONDI.**

He has been active in the field of market and social research since the early 1990s. Dr Hellwig has a degree in social sciences, psychology and media studies. He worked for several years as a research assistant at the Institute for Applied Social Research at the University of Cologne and received his doctorate in 2000. Otto Hellwig has been Chairman of the DGOF board since March 2013.

## PROF. DR. FLORIAN KEUSCH



**FLORIAN KEUSCH IS PROFESSOR OF SOCIAL DATA SCIENCE AND METHODOLOGY IN THE SCHOOL OF SOCIAL SCIENCES AT THE UNIVERSITY OF MANNHEIM AND ADJUNCT RESEARCH PROFESSOR IN THE JOINT PROGRAM IN SURVEY METHODOLOGY (JPSM) AT THE UNIVERSITY OF MARYLAND.**

He received his PhD in Social and Economic Sciences from WU, Vienna University of Economics and Business, Austria. In his research, he develops, implements, and assesses modern methods of collecting digital data for the behavioral and social sciences. He has been a member of the DGOF board since 2017.

## PROF. DR. SIMON KÜHNE



**SIMON KÜHNE IS PROFESSOR OF APPLIED SOCIAL DATA SCIENCE AT THE FACULTY OF SOCIOLOGY AT BIELEFELD UNIVERSITY. HIS METHODOLOGICAL RESEARCH IS IN THE AREAS OF SURVEY METHODOLOGY AND COMPUTATIONAL SOCIAL SCIENCE, ESP. SOCIAL MEDIA.**

He also works on aspects of social inequality in the areas of sexual and gender diversity, racism and discrimination, and health. Simon Kühne has been actively involved in shaping the GOR for several years. After serving as a reviewer and track chair, he has shared responsibility for the program as Vice Program Chair since 2022. He has been a member of the DGOF board since 2023.

## DR. STEFAN OGLESBY, MBA



**STEFAN OGLESBY IS CHAIRMAN OF DATA IQ AG, A CONSULTING AGENCY SPECIALIZING IN DATA STRATEGY AND ANALYTICS SERVICES. .**

He is also active in academia, with publications about digital methods, and in his role as lecturer for consumer research at the University of Lucerne. He has more than 20 years of experience in marketing and social research, including roles as research director and CEO at a leading Swiss market research agency. He is a board member of DGOF since 2019.



## PROF. DR. BELLA STRUMINSKAYA



**BELLA STRUMINSKAYA IS A PROFESSOR IN METHODS AND STATISTICS AT UTRECHT UNIVERSITY IN THE NETHERLANDS AND AN AFFILIATED RESEARCHER AT STATISTICS NETHERLANDS. SHE HOLDS A DOCTORAL DEGREE IN SURVEY METHODOLOGY FROM UTRECHT UNIVERSITY.**

Her research focuses on innovations in data collection methods, data quality in studies that use smartphone apps, sensors, and wearables, and combining survey data with other data sources. She has published on various aspects of data quality in sensor and survey data, nonresponse and measurement error, panel conditioning, and device effects. Bella Struminskaya has been a board member of the German Society for Online Research (DGOF) since 2017 and is the programme chair of the GOR 23 conference.

## ALEXANDRA WACHENFELD-SCHELL



**ALEXANDRA WACHENFELD-SCHELL IS SENIOR RESEARCH DIRECTOR AT GIM GESELLSCHAFT FÜR INNOVATIVE MARKTFORSCHUNG. SHE WAS PREVIOUSLY CUSTOMER EXPERIENCE MANAGER AT SGBDD AND MANAGING DIRECTOR AT FORSA.MAIN, A FULL-SERVICE INSTITUTE SPECIALIZED IN MARKET, MEDIA AND SOCIAL SURVEYS.**

She has broad experience in the area of quantitative market and social research. She began working in market research more than 20 years ago as a project manager at LINK Institut. For over 10 years she was responsible for strategy, marketing and business development in the area of online research in her role as Research Director at LINK. She, moreover, supervised the actively recruited (by representative telephone interviews) LINK online panel with regards to methods. She focuses on online and methodological research and is a regular speaker at (market) research conferences. She is a member of the DGOF board since March 2013.

## GREETINGS FROM THE LOCAL PARTNER

# DEAR PARTICIPANTS OF THE GOR 23,



**by** **PROF. DR. BETTINA LANGFELDT**  
Chair of Methods of Empirical Social Research, University of Kassel

I am delighted to welcome you to the General Online Research Conference (GOR) 2023, which is located at the University of Kassel for the very first time. This year's conference holds a special significance. As one of the first conferences in Europe, the GOR began in the 1990s to deal with digital research in detail. Since 1997 it has been held regularly once a year in different cities in the German-speaking area, only interrupted by two years of meeting online due to the worldwide pandemic. Not only are we commemorating the 25th edition of GOR this year, but also the 25th anniversary of the German Society for Online Research (DGOF) as a remarkable milestone that highlights the enduring commitment of our community.

I have closely followed the evolution of the DGOF, and now, as the local host of this "anniversary edition", I could not be more excited. The GOR conference has always been synonymous with openness, innovation, interdisciplinary collaboration, and a vibrant exchange between academia and industry. These values resonate deeply with the ethos of the University of Kassel, making it a perfect setting for GOR 23.

During the three conference days, we will delve into the latest innovations, trends, and developments in the ever-evolving field of digital research. Main topics are advancements in online and mobile web surveys, data science (from big data to smart data), and politics, public opinion, and communication. Besides, the use of AI in different areas of society, and its influence on teaching, learning, and writing exams at universities is certainly a highly topical issue that we will reflect on. The program is characterized by different formats. It provides a platform for insightful discussions, thought-provoking presentations, and fruitful interactions among researchers and practitioners from diverse backgrounds. Together, we will contribute to the advancement of digital research and shape the future of our field for the next 25 years to come.

Apart from the stimulating exchange of ideas, I am also looking forward to celebrating this double jubilee with all of you in person and on premise. I am confident that GOR 23 will be an unforgettable experience, filled with valuable connections, and inspiring moments. On behalf of the University of Kassel and the GOR organizing committee, I extend a warm invitation to all participants to join us and to create lasting memories.

## ORGANIZER

U N I K A S S E L  
V E R S I T Ä T

## SPONSORS



INTERROGARE



norstat



## MEDIA PARTNER

planung&amp;analyse

marktforschung<sup>de</sup>

## COOPERATION PARTNER

25-26 October 2023, RMCC  
Wiesbaden/Frankfurt, Germany

## WEDNESDAY 20/09/2023

09:00 AM - 10:00 AM	<b>BEGIN CHECK-IN</b>
10:00 AM - 01:00 PM	<b>WORKSHOP 1</b> Targeted Advertising for Survey Recruitment: Leveraging Social Media Platforms to Reach and Engage Survey Participants INSTRUCTOR: Zaza Zindel [Bielefeld University, Germany]
01:00 AM - 01:30 AM	<b>BREAK</b>
01:30 PM - 04:30 PM	<b>WORKSHOP 2</b> Be careful of the careless: Data quality assessment for online surveys INSTRUCTOR: Nivedita Bhaktha, Thomas Knopf [GESIS – Leibniz Institute for the Social Sciences, Germany]
	<b>WORKSHOP 3</b> Trustworthy Analytics with Generative AI: ChatGPT/GPT-4 Hands-on INSTRUCTORS: Georg Wittenburg, Patrick Mertes [Inspirient GmbH, Germany]
04:30 PM - 05:00 PM	<b>BREAK</b>
05:00 PM - 07:00 PM	<b>DGOF MEMBERS GENERAL MEETING</b>
07:00 PM - 8:00 PM	<b>EARLY CAREER SCIENCE SPEED DATING</b>
08:00 PM - 11:00 PM	<b>GOR 23 GET TOGETHER</b>

## THURSDAY 21/09/2023

08:00 AM - 09:00 AM	<b>BEGIN CHECK-IN</b>
09:00 AM - 10:15 AM	<b>GOR 23 OPENING &amp; KEYNOTE 1:</b> THE EVOLUTION OF MARKETING RESEARCH: FROM INSIGHTS 1.0 TO 5.0 Keynote Speaker: Steffen Schmidt [LINK Marketing Services AG, Switzerland]
10:15 AM - 10:45 AM	<b>BREAK</b>
10:45 AM - 11:45 AM	<b>A1:</b> Survey Measurement / sponsored by GESIS <b>B1:</b> Text as Data <b>C1:</b> Media Consumption Habits / sponsored by aproxima <b>D1:</b> GOR Best Practice Award 2023 Competition 1 <b>T1:</b> GOR Thesis Award 2023 Competition 1 / sponsored by TIVIAN
11:45 AM - 12:00 AM	<b>BREAK</b>
12:00 AM - 01:00 AM	<b>A2:</b> Respondent Behavior and Data Quality / sponsored by GESIS <b>B2:</b> ML and AI in Surveys <b>D2:</b> GOR Best Practice Award 2023 Competition 2 <b>T2:</b> GOR Thesis Award 2023 Competition 2 / sponsored by TIVIAN
01:00 AM - 02:15 AM	<b>LUNCHBREAK</b>
02:15 PM - 03:30 PM	<b>POSTER SESSION</b>
03:30 PM - 03:45 PM	<b>BREAK</b>

...

## THURSDAY 21/09/2023

03:45 PM - 04:45 PM	<b>A3.1:</b> Recruitment Processes for Online Surveys / sponsored by GESIS <b>A3.2:</b> Mixed Mode and Mode Transition I / sponsored by GESIS <b>B3:</b> Smartphones, Sensors, and Geodata <b>C3:</b> Extremism, Hate Speech and Political Engagement <b>D3:</b> Innovation in Practice (1): Transforming MR with Data-Driven Tools
04:45 PM - 05:00 PM	<b>BREAK</b>
05:00 PM - 06:00 PM	<b>A4.1:</b> Data Quality in Online Surveys / sponsored by GESIS <b>A4.2:</b> Mixed Mode and Mode Transition II / sponsored by GESIS <b>B4:</b> Combining Self-Reports with Digital Behavioral Data <b>D4.1:</b> Innovation in Practice (2): Opportunities and Challenges <b>D 4.2:</b> Podiumsdiskussion: 25 mal GOR, 25 Jahre DGOF – aus der Vergangenheit in die Zukunft. Eine Diskussion ehemaliger und amtierender Vorstände
08:00 PM - 11:00 PM	<b>GOR 23 PARTY</b>

## FRIDAY 22/09/2023

09:30 AM - 10:00 AM	<b>BEGIN CHECK-IN</b>
10:00 AM - 10:45 AM	<b>KEYNOTE 2:</b> POPULATION REPRESENTATIVE SURVEYS, WEARABLES AND CONTEXTUAL DATA Keynote Speaker: Arie Kapteyn (University of Southern California, United States of America)
10:45 AM - 11:15 AM	<b>GOR AWARD CEREMONY</b>
11:15 AM - 11:45 PM	<b>BREAK</b>
11:45 PM - 12:30 PM	<b>A5:</b> Increasing Survey Participation / sponsored by GESIS <b>B5:</b> QR-Code Technology <b>C5:</b> Political Issue Salience / sponsored by aproxima <b>D5:</b> Podiumsdiskussion AI / LLM
12:30 PM - 01:45 PM	<b>LUNCHBREAK</b>
01:45 PM - 02:45 PM	<b>A6.1:</b> Respondent Interaction / sponsored by GESIS <b>A6.2:</b> Bias Assessment and Correction / sponsored by GESIS <b>B6:</b> Online Search Data <b>C6:</b> Trust in Science / sponsored by aproxima



## TARGETED ADVERTISING FOR SURVEY RECRUITMENT: LEVERAGING SOCIAL MEDIA PLATFORMS TO REACH AND ENGAGE SURVEY PARTICIPANTS

Instructor: **Zaza Zindel** (Bielefeld University)

Workshop language: **English**

### Goals of the workshop:

1. Introduce participants to the concept of targeted advertising on social media platforms for survey recruitment.
2. Highlight the advantages and disadvantages of targeted advertising in reaching and engaging specific survey respondents, particularly in the context of web surveys.
3. Provide participants with practical knowledge and skills necessary to create, configure, and optimize targeted advertising campaigns on Facebook for effective survey recruitment.
4. Enable participants to utilize Facebook's API to monitor the progress of their survey recruitment campaigns, allowing them to track key metrics and assess the effectiveness of their ads in reaching the intended target audience.
5. Equip participants with the ability to review survey results for quality violations, enabling them to identify and address issues such as duplicate responses or suspicious patterns that may impact data integrity.

### About the instructor:

Zaza Zindel is a doctoral researcher and research assistant in sociology at Bielefeld University. Her research focuses on various areas within social science, including survey methodology, the utilization of social media for empirical social research, and the examination of social inequalities. In her doctoral dissertation, she specifically explores the application of social media advertising as a means to recruit rare populations for web surveys. With a strong academic background and expertise in these areas, Zaza brings valuable insights and knowledge to the workshop.

**Description of the content of the workshop:** This workshop aims to provide participants with a comprehensive understanding of how to use targeted advertising on social media platforms, specifically Facebook, to recruit survey participants. Through a step-by-step guide, participants will learn how to create and configure targeted advertising campaigns tailored to reach and engage specific respondents for web surveys. The workshop will emphasize the impact of creative elements within ads, highlighting their influence on audience attention, interest, and data quality. Furthermore, participants will gain practical insights into utilizing Facebook's API to monitor the progress of survey recruitment campaigns.

Key metrics and analytics will be covered, empowering researchers to evaluate the effectiveness of their ads in reaching the desired audience. This data-driven approach enables participants to make informed decisions and adjustments throughout the recruitment process. Lastly, the workshop will explore techniques for reviewing survey results to identify and address data quality violations. Participants will learn strategies to detect issues such as duplicate responses or suspicious patterns, ensuring the integrity and reliability of collected data.

WEDNESDAY  
20 September 2023  
01:30 PM - 04:30 PM

WORKSHOP  
2

## BE CAREFUL OF THE CARELESS: DATA QUALITY ASSESSMENT FOR ONLINE SURVEYS

Instructors: **Dr. Nivedita Bhaktha and Mr. Thomas Knopf**  
(GESIS – Leibniz Institute for the Social Sciences)

Workshop language: **English**

### Goals of the workshop:

1. Learn about data quality taxonomy, issues, and concept.
2. Detect and evaluate low quality data from the sample.
3. Gain hands-on experience of data quality assessment in R using commonly applied data quality indicators in online survey research.

### About the instructor:

Dr. Nivedita Bhaktha is a senior researcher at GESIS, a Leibniz Institute for the Social Sciences in Germany. She completed her Ph.D. in quantitative research, evaluation, and measurement from the Ohio State University. She has had an extensive training in both theoretical and applied statistics. She has had a rich interdisciplinary research and teaching experience having taught statistics at different levels to students from varied disciplines. Her research interests are in psychometrics and data quality assessment. Some of her current projects include impact of low data quality on psychometric properties of scale, research synthesis of data quality indicators used in social sciences. She is also working on creating R packages to aid data quality assessment in survey research.

Mr. Thomas Knopf is a researcher in the team “Scale Development and Documentation” at GESIS - Leibniz Institute for the Social Sciences. He holds a Master's degree in psychology and a Diploma in communication design. His research interests are in psychometrics and data analysis of multi-item rating scales and cognitive assessment tests. In his dissertation, he deals with statistical modelling of “short-text answers as data” and conducts applied research in the realm of personality and cognitive abilities, like creativity and intelligence. As part of his daily work for ZIS (Open Access Repository for Measurement Instruments), he develops guidelines and packages in R and Stata, for e.g., data cleaning, wrangling, and modelling. He has had a broad teaching experience for introductory and intermediate courses in presence and blended learning formats.

**Description of the content of the workshop:** This is a short introductory course on detecting and handling data quality issues in datasets from online survey research. This course will equip the audience with the knowledge and practical skills to perform some basic and advanced data quality procedures. Some of the topics covered are as follows:

- Introduction – taxonomy of response quality indicators (RQIs)
- Check items – usage of e.g., attention checks and similar
- Para data – usage of e.g., time variables
- Item responses – usage of three classes of RQIs that address outlier, consistency, and pattern

We will provide demonstration & hands-on training on data quality assessment using our inhouse developed R packages, covering the topics mentioned above. The three R packages covered are as follows:

- requin – Response Quality Indicators
- restin – Response Style Indicators and
- respin – Response Pattern Indicator



WEDNESDAY  
20 September 2023  
01:30 PM - 04:30 PM

WORKSHOP  
3

## TRUSTWORTHY ANALYTICS WITH GENERATIVE AI: CHATGPT/GPT-4 HANDS-ON

Instructors: **Dr. Georg Wittenburg** and **Mr. Patrick Mertes**  
(Inspirient GmbH)

Workshop language: **English**

### Goals of the workshop:

1. Allow participants to gain first-hand experience in analyzing survey data using a combination of Generative AI systems
2. Review limitations and prerequisites of the technology and build skills in operating it without compromising analytical rigor or the soundness of results
3. Come up with an initial personal “cookbook” of prompts that reliably produce draft deliverables, e.g., a report on key analytical findings from a survey

### About the instructor:

Dr. Georg Wittenburg founded Inspirient after five years of ICT research and three years as a management consultant at the Boston Consulting Group. He wholeheartedly disagrees with software “eating the world”, mostly on the precedents that the wheel, the steam engine, and the calculator didn’t develop a taste for the world either. Instead, he’s looking forward to the new possibilities that Artificial Intelligence will add to humanity’s tool chest, and to the great things we’ll all get to do once this transition is done.

Georg studied Computer Science and Business Administration in Berlin and Vancouver. He received his PhD (“summa cum laude”) from Freie Universität Berlin in 2010, and subsequently conducted research in topics including distributed systems, pattern recognition, and the Internet of Things at INRIA Paris-Saclay. He started working at the Boston Consulting Group (BCG) in 2012 and advised international corporations in the logistics, automotive, and media sectors. From 2013 to 2014, he was selected to join the BCG Henderson Institute, the firm’s think tank to push the envelope on business strategy, and worked on topics including innovation strategy and organizational algorithmics. Georg kicked off Inspirient in the summer of 2014 and is now serving as Inspirient’s CEO.

**Description of the content of the workshop:** ChatGPT doesn’t do math beyond adding three numbers. GPT-4 improves upon this, but the core question remains: Can Generative AI, and in particular Large Language Models (LLMs), help in deriving relevant analytical results from survey data?

This workshop takes a hands-on, end-to-end perspective on answering this question. We start with a real-world survey dataset, briefly review data annotations (as per DIN SPEC 32972), and then utilize Inspirient’s Analytics Engine and OpenAI’s GPT-4 to query the data for key quantitative findings, high-level insights, and draft deliverables. We will study how pre-rendered analyses interact with the LLM’s data ingestion requirements, and cover prompting alternatives to produce multi-modal deliverables (text + charts) such as executive summaries, report outlines, and press release text snippets.

Participants are invited to bring their own laptops to get the most out of this workshop. Required account credentials will be made available free of charge during the session. No personal / confidential data will be shared with third-parties, incl. OpenAI.

## POPULATION REPRESENTATIVE SURVEYS, WEARABLES AND CONTEXTUAL DATA



**by** **ARIE KAPTEYN** Professor (Research) of Economics, Center for Economic and Social Research, University of Southern California

Surveys have become both more challenging and easier to implement. Response rates to probability-based surveys have shown a downward trend for several decades now. At the same time, the advent of the Internet allows literally everyone to conduct surveys at close to zero cost. Furthermore, big data and wearable devices are sources of new “passive” information that can complement or replace survey information. Navigating this landscape requires considerable skill in avoiding potentially fatal pitfalls.

Non-probability panels can do a terrible job in representing a target population, both in terms of estimating population parameters and when used to estimate relations of substantive interest. Even within probability-based surveys, measurement may be difficult due to several well-known biases, including mode effects, social desirability, and simply the difficulty respondents may have articulating answers to complicated or vague questions. Big data often represent selective populations, such as social media users, drivers with smartphones that can be tracked, or owners of a particular device, such as a Fitbit. Wearable devices are typically owned by parts of the population with more money, and an interest in new technologies.

Nevertheless, the many new opportunities for data collection create opportunities for novel and more accurate information collection than has been possible in the past. For a start, the way respondents answer surveys provides information beyond the content of a questionnaire. These paradata have shown to be powerful indicators of the cognitive status of respondents, for instance. Combining self-reports on physical activities with the provision of activity trackers yields superior information on an important dimension of health behavior. Beyond measurement of air quality by ground monitors, providing wearable air quality monitors allows for fine grained measurement of variation in air quality by time and place at the individual level. Activity trackers, like Fitbit, can detect illness, sleep interruption and how that affects cognition on a day-to-day basis.

I will illustrate both the pitfalls of modern data collection and the promise of combining population representative survey data with information from wearable devices and contextual information.

## THE EVOLUTION OF MARKETING RESEARCH: FROM INSIGHTS 1.0 TO 5.0



**by** **STEFFEN SCHMIDT** Director Marketing Science & Agile Insights; LINK Marketing Services AG, Switzerland

The field of marketing research has evolved and changed significantly over the past three decades, passing through five key phases referred to as Insights 1.0 to Insights 5.0. In this context, a comprehensive overview of this evolution is provided, showing how marketing research has evolved from traditional techniques to modern methods powered by artificial intelligence [AI].

In particular, an overview of the changing landscape of marketing research is provided, demonstrating its adaptability and adoption of new technological and scientific advances for a better consumer understanding. However, not only is a journey taken through the dynamic evolution of marketing research, but also a reflection on the possible future directions the discipline could take in light of current and future innovative methodologies.

## GOR BEST PRACTICE AWARD 2023:



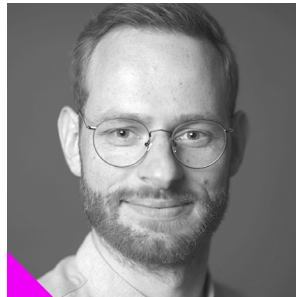
**YANNICK RIEDER**  
Janssen-Cilag GmbH,  
Jury Chair



**HOLGER GEISSLER**  
marktforschung.de



**SABINE MENZEL**  
L'Oréal Deutschland



**SEBASTIAN VERFÜHRTH**  
RTL DISNEY Fernsehen  
GmbH & Co. KG

# GOR BEST PRACTICE AWARD 2023

**THURSDAY, 21 SEPTEMBER 2023:**  
**10:45 AM - 11:45 AM AND 12:00 PM - 1:00 PM**

Nominees for the GOR Best Practice Award 2023 come from GIM Gesellschaft für innovative Marktforschung mbH (Germany), Erason (Germany), Produkt + Markt (Germany), LINK Marketing Services AG (Switzerland), Rivella AG (Switzerland), advise research GmbH (Germany) and Hochland Group (Germany).

The "best practice in commercial online market research" competition was introduced at the 2008 GOR conference and since then has become an important tradition at the General Online Research Conference. Real world case studies with a strong focus on online market research are presented in the competition. The prize is awarded annually to the study which has most effectively answered a key question of online market research with the help of innovative digital methods.

The winners of the award will be selected by the jury and by a random sample of the audience via a mobile survey. The award will be presented to the winners at the GOR Party Thursday evening and the laudatory speech will be held at the GOR Award Ceremony on Friday 10:45 - 11:15.

# ABSTRACTS

## DEEPER TARGET GROUP UNDERSTANDING THROUGH ARTIFICIAL INTELLIGENCE

SILKE MOSER

GIM Gesellschaft für innovative Marktforschung mbH, Germany

NIKLAS MRUTZEK

Erason, Germany

**RELEVANCE & RESEARCH QUESTION:** Revision and readjustment of ARD's typological target group model "MNT" to explain media usage with the aim to better reflect the middle of society as well as younger user types and to map and explain the dynamic media change in digital markets. Additional wish: Obtaining a deeper understanding of the target group and broader information on the individual types in order to optimize the target group approach. Furthermore, obtaining a tool that is able to directly map the ever faster changing usage habits and preferences in order to agilely align the own offer to the dynamically changing target group needs.

**METHODS & DATA:** Baseline survey (GIM):

A combination of CATI and online survey was used to improve the sample quality and to better reach younger and more mobile target groups in particular. In addition, a dual frame sample was selected for the CATI survey. In order to be able to differentiate sufficiently in the segmentation of up to 10 types, a total of n=2,500 interviews (nationally representative of the German-speaking population aged 14 and over) were realized in the ratio 70% CATI / 30% online. For the online sample, primarily the purely offline recruited GIMPulse panel was used.

Artificial Intelligence Allon (Erason): Allon obtains anonymized digital footprints from statistics, news, and social media that yield affinities and sociodemographic and geographic attributes, which are enriched via NLP. The enriched affinities and attributes are captured by Allon in a network that reflects the interrelationships and their strength of all attributes within society. Via Evolutionary cascaded-deep learning, unseen attributes can be accurately predicted. Via the intelligence and structures of the network, Allon can compute up to 70,000 attributes for each input with great algorithmic accuracy.

Connectivity: Defining media user types in Allon to evaluate its broad database for each user type.

**RESULTS:** The application of a mixed-mode study approach and thus the population-representative mapping played a decisive role in achieving the above-mentioned goals. As a result of the readjustment, 9 types could be identified, whose media use can now be better explained, especially in digital terms. In addition, through the connection with AI, it was possible to describe the target groups and their usage behavior even more precisely and in greater detail. On the other hand, trends in usage behavior or changes in preferences over time can be shown through the dynamization of the data situation in the AI.

**ADDED VALUE:** Based on the representatively collected data, changes in the needs and preferences of digital media types can be analyzed at any time. In conjunction with AI, the digital strategic orientation of media companies is thus consistently supported, while at the same time program makers are provided with a valuable tool with which they can evaluate data on their target groups according to their individual requirements and call it up on a daily basis.

## BETTER TOGETHER IN SPEED, AGILITY AND QUALITY

KATJA BIRKE, SCHOMBERG JESSICA

Produkt + Markt, Germany

**RELEVANCE & RESEARCH QUESTION:** Improving the ownership of insights, merging **knowledge** and improving **speed, agility** and the **quality** of insights – our show case will present the mode of working and learnings from running various Insight Lounges for pharmaceutical clients. Beyond speed, in-house developed AI allows our teams to work more agile in qualitative online research.

**METHODS & DATA:** The Insight Lounge is a pre-recruited expert online panel of participants who are available for various projects over a period > 6 months. Stakeholders launch various projects ranging from a single question to IDs or iterative workshops. Surveys from a boost sample quantify results.

Regular meetings allow the team to share results, plan the next session and put forward "hot topics" that emerge from daily work. This is where our AI tool QUALITYTIS enables researchers to produce results in shorter time.

**RESULTS:** Market research designed with an Insight Lounge provided key advantages:

**Speed:** Respondents are recruited once, no time-consuming recruitment. QUALITYTIS transcribes the interviews in real time, creates initial word clouds, merges likes or dislikes, or encodes answers to a pre-defined grid.

**Quality:** The automated evaluation of complete transcripts allows research teams to put attention to the interpretation of results. Quality is also driven by identifying important aspects that nobody has thought of before. More depth: At the same time, knowledge grows. The different stakeholders feel very close to their target group and share information right away.

**ADDED VALUE:** We did better together in various ways: We observed a strong commitment from our clients' teams since stakeholders felt more involved and heard. We built strong relationships with our participants since they work on one topic during various sessions. We shared feedback on the results increasing their sense of meaning for market research. In terms of results, we created a big picture over time by

- linking results from one session to the other
- searching for "nuggets" in answers and word clouds
- building up knowledge in all stakeholders
- writing personas with a comprehensive understanding of the target group

## CALM DOWN AND PUSH THE PREDICTIVE DEMAND SCORING (PDS) BUTTON TO REVITALIZE BRAND GROWTH

STEFFEN SCHMIDT, MARIANNE ALTGELD, MATTHIAS BIEDERMANN

LINK Marketing Services AG, Switzerland

SARA JERMANN, DÉSIRÉE KOLLER, NIKOLAS PETSCHEN

Rivella AG, Switzerland

**RELEVANCE & RESEARCH QUESTION:** FOCUSWATER, a subbrand of Rivella, recognizes the need to adapt towards the prevailing zeitgeist to remain competitive. This includes primarily to address the preferences of both existing and potential consumers, as well as the broader segment of non-consumers. In order to align better with the current zeitgeist towards healthier, sugar-free products, and thus to enhance its competitiveness, FOCUSWATER developed two sugar-free drink concepts, concretely 'push' and 'calm', each with two distinct recipes. The management question to get answered from a decision-driven perspective was: Which of these product variations will effectively boost the brand to the next level of consumer demand and growth.

**METHODS & DATA:** The core of the Predictive Demand Scoring (PDS) approach is the combined application of two advanced methods from neuromarketing and artificial intelligence (AI), namely fast reaction time (FRT) measurement and universal structural modeling (USM) using Bayesian Neural Networks. On top, the visual strength of the product packaging design was predicted by using AI-based eye tracking plus the attribute strength of the product features by using MaxDiff-TURF. Applying an experimental design to determine the product demand, participants were randomly sent a product sample for each product and conducted an online survey while consuming the respective product sample..

**RESULTS:** The applied PDS identified two winner products with a substantially highest market potential (e.g., penetration). Findings also refined product features and packaging design based on insights from MaxDiff-TURF and AI-based eye tracking. In 2022, the launch of these products contributed to a 12% growth for Rivella, primarily fueled by FOCUSWATER's expanded portfolio with the new 'push' and 'calm' products.

**ADDED VALUE:** The PDS provides an evidence-based approach to select top-performing products. It reveals key product features and associations, but also determines packaging elements drawing distinct attention, and optimized product perception ease for better shelf performance. The PDS can be applied to all consumer categories, from insurance to water, with a high explanation and prediction power between 80% to 90% on average. Overall, PDS offers a comprehensive customer understanding for targeted product development and communication, leading to significant business impact.

## APPLICATION OF AI: LLM MODELS IN PRACTICAL MARKET RESEARCH WITH SPECIAL FOCUS ON CHATGPT

ANDREAS WOPPMANN, TABEA WEISSMANN

advise research GmbH, Germany

KATRIN MATTUSCH

Hochland Group, Germany

**RELEVANCE & RESEARCH QUESTION:** The use of so-called large language models has revolutionized the analysis of large volumes of text since the end of 2022. Until then, the analysis of large volumes of text was associated with a lot of AI training or the use of humans. Both options are quite time wasting and expensive.

**METHODS & DATA:** For the first time, it is possible to analyze large volumes of text without prior training and evaluate them in terms of content. As a basis, the collected data is coded using AI via an interface developed by advise, which communicates via api with ChatGPT. This analysis serves as a quantitative analysis of open-ended responses and replaces the currently used manual data coding. Additionally, data is also qualitatively evaluated by the AI. For the analysis a new LLM model (in this case ChatGPT 4.0 established March 2023) has been used.

**RESULTS:** Using concrete examples from Hochland projects, very good results regarding the comparison from manual coding and the coding via LLM could be achieved. Although the coding itself still needs to be instructed and briefly checked for accuracy, a significant time and cost advantage can be achieved in the evaluation of responses. Besides increased efficiency (including translation in English for all languages), the effectiveness can be enhanced through additional qualitative analyses. The analysis attains greater depth with the inclusion of supplementary qualitative evaluations and is perceived as more comprehensible by many participants. Furthermore qualitative analysis of interviews can be supported by the AI model. However, it remains crucial to involve a human project supervisor in the process, utilizing an interface that enables rapid and intuitive action and response.

**ADDED VALUE:** We see this new AI-driven approach as a disruptive innovation that accelerates market research projects, reducing costs and timelines, while also enhancing the quality of results through a combination of quantitative and qualitative approaches. Embracing this symbiotic relationship between highly standardized technology and human expertise, our approach marks a significant advancement in market research practices.

## GOR POSTER AWARD 2023 JURY:



**JAN KAREM HÜHNE**  
University of Duisburg-Essen;  
& Jury Chair



**FRANZISKA KERN**  
infas 360 GmbH



**MAREIKE OEHL**  
September Strategie &  
Forschung GmbH



**CATHLEEN STÜTZER**  
TU Dresden



**VAKA VÉSTEINSDÓTTIR**  
University of Iceland

# GOR POSTER AWARD 2023

**THURSDAY, 21 SEPTEMBER 2023:  
02:15 PM - 03:30 PM**

Posters offer the opportunity to present late breaking research, short research findings or discuss work in progress at the GOR conference. The presented works will be evaluated by a jury. The GOR Poster Award 2023 comes with a prize money of € 500.

Posters are presented in a plenary session on Thursday. Access to the poster presentations will be possible during the whole conference. A poster may cover any topic of online research. All submissions in this category are considered for the GOR Poster Award 2023. The poster award ceremony will take place on Friday, 22 September 2023 at 10.45.

### PAST WINNERS OF THE GOR POSTER AWARD ARE:

- GOR Poster Award 2022: Lisa Marie Natter for her Poster "The Impact of Monetary incentives on Retention Rates in a Panel Study with Mixed Web and Mail Modes"
- GOR Poster Award 2021: Tobias Rettig, Bella Struminskaya and Annelies Blom for their Poster "Memory Effects in Online Panel Surveys: Investigating Respondents' Ability to Recall Responses from a Previous Panel Wave"
- GOR Poster Award 2020: Adrian Hoffmann, Julia Meisters and Jochen Musch (University of Duesseldorf, Germany) for their Poster "Indirect Questioning Techniques: An Effective Means to Increase the Validity of Online Surveys"



## GOR THESIS AWARD 2023 JURY:



**DR. OLAF WENZEL**  
Wenzel Marktforschung &  
Jury Chair



**DR. CARINA CORNESSE**  
DIW Berlin



**DR. FREDERIK FUNKE**  
LimeSurvey &  
datenmethoden.de



**PROF. DR.  
BETTINA LANGFELDT**  
University of Kassel



**ASSISTANT PROFESSOR  
ANNA-SOPHIE  
ULFERT-BLANK**  
Eindhoven University of  
Technology

# GOR THESIS AWARD 2023

**THURSDAY, 21 SEPTEMBER 2023:**  
**10:45 AM - 11:45 AM AND 12:00 PM - 1:00 PM**

The GOR Thesis Award Competition is an integral part of the GOR conference series and takes place annually. It comes with a prize money of 500 € for each of the two parts of the competition.

All submissions relevant to online research are welcome. Presentations in the past years covered a broad range of topics, be it online surveys or research on the Internet or social aspects of the Web. Theses must have been submitted in 2021 or 2022, the Thesis language can be either English or German. The programme committee assigned three members of the Thesis jury to review each submission. The review process was completely anonymised and reviews were randomly distributed in a way that conflicts of interests were avoided.

Three Doctoral Theses, two Master Theses and one Bachelor Thesis are nominated for this year's GOR Thesis Award. The selected authors will present their findings at the GOR conference and the best presentations will be awarded.

### NOMINATED FOR THE GOR THESIS AWARD 2023 ARE:

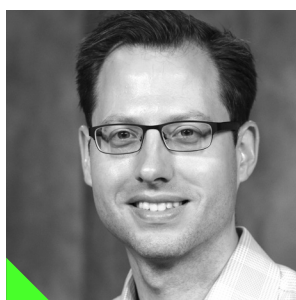
- Sebastian Kocar (University of Tasmania, Australia):  
Approaches to Dealing With Survey Errors in Online Panel Research
- Tobias Rettig (University of Mannheim, Germany):  
Respondent and Response Behavior in Online Panel Surveys
- Tabea Tesch (Royal Canin/Mars Incorporated, Germany):  
Online Product Testing for Generating Online Product Reviews –  
An Analysis of the Effects on Review Rating and Review Quality
- Janna Hämpke (University of Münster, Germany):  
How to Communicate in a Crisis? The Influence of Local Governments'  
Humorous Crisis Response Strategy and Crisis Responsibility on Trust,  
Emotions and Behavioral Intentions
- Raphael Riege (Kassel University, Germany):  
Homophily in Online Friend Selection. Empirical Investigation of Friend  
Selection in the App FriendsUp.
- Inan Eren Bostanci (Zuse Institute Berlin, Germany), Yvonne Gootzen  
(Statistics Netherlands, Netherlands), Peter Lugtig (Utrecht University,  
Netherlands):  
Data Linkage to Validate and Calibrate Traffic Estimations on a  
Nationwide Scale: A Framework for Official Statistics

The GOR Thesis Award 2023 is sponsored by

[ TIVIAN ]



## DGOF BEST PAPER AWARD 2023 JURY:



**ASSISTANT RESEARCH PROF.  
CHRISTOPHER ANTOUN**  
University of Maryland



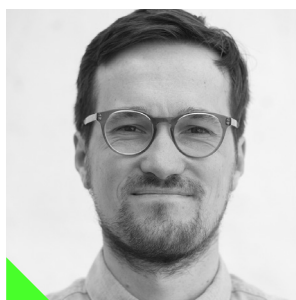
**PROF. DR. NICOLA DOERING**  
Ilmenau University of  
Technology



**PROF. DR. FLORIAN KEUSCH**  
University of Mannheim &  
Jury Chair



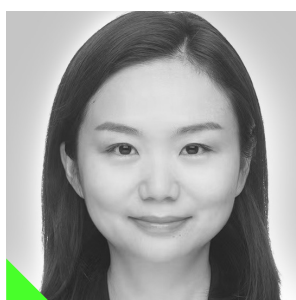
**DR. HENNING SILBER**  
GESIS Leibniz Institute for the  
Social Sciences



**ASSOCIATE PROFESSOR  
CLEMENS STACHL**  
University of St. Gallen



**PROF. DR. BELLA  
STRUMINSKAYA**  
Utrecht University



**CHAN ZHANG, PHD**  
Zhejiang University

# DGOF BEST PAPER AWARD 2023

**FRIDAY, 22 SEPTEMBER 2023:  
10:45 AM - 11:15 AM (AWARD CEREMONY)**

The German Society for Online Research (DGOF) annually recognizes outstanding scientific contributions in online research through the DGOF Best Paper Award for a researcher or group of researchers.

The prize is awarded to a paper that provides a fundamental scientific contribution to the advancement of the methods of online research. Both theoretical/conceptual and empirical/methodological papers are considered for the award.

The award is worth 500 Euro and will be presented at the annual GOR conference. An abstract (and, if available, a preprint) of the award-winning paper will be posted to the DGOF website ([www.dgof.de](http://www.dgof.de)). To be considered for the award, papers must have been published in an outlet that uses a peer-review process (e.g., peer-reviewed journal, full papers in peer-reviewed conference proceedings, refereed book chapter) at the time of submission. Papers written in German or English and published not earlier than 2022 (if the paper was published online-first, then the online-first publication date counts) were eligible to be submitted for the DGOF Best Paper Award 2023.

### WINNER:

Kim et al. (2022). Universal adaptability: Target-independent inference that competes with propensity scoring. PNAS, 119(4):e2108097119. Accessible on: <https://doi.org/10.1073/pnas.2108097119>

Sponsored by University of Bielefeld, Faculty of Sociology

**UNIVERSITÄT  
BIELEFELD**  
Faculty of Sociology

THURSDAY, 21/SEPTEMBER/2023

# A1: SURVEY MEASUREMENT

(SPONSORED BY GESIS)

SESSION CHAIR: **ALMUTH LIETZ**Deutsches Zentrum für Integrations- und  
Migrationsforschung, Germany

## OPEN-ENDED SURVEY QUESTIONS: A COMPARISON OF INFORMATION CONTENT IN TEXT AND AUDIO RESPONSE FORMATS

**LANDESVATTER, CAMILLE; BAUER, PAUL C.**

MZES, University of Mannheim, Germany

**RELEVANCE & RESEARCH QUESTION** Open-ended survey questions provide a valuable and rich source of data in addition to closed-ended questions. However, their use presents challenges to respondents, such as increased response burden and a potential decrease in response quality. Therefore, it is important for survey researchers to design open-ended questions that result in informative answers.

With the rise of online surveys and recent technological advancements, spoken response formats that enable respondents to record their answers using speech input have become more feasible. This study investigates the impact of asking respondents to answer open-ended questions by audio compared to the traditional text-entry format. By comparing the two formats, we aim to understand how spoken responses impact answer length, quality, and overall information content.

**METHODS & DATA** We use a U.S. sample (N=1,500) and questions adapted from popular social survey programs, such as the „most important problem“ question. By experimentally varying the response format, we examine which format (e.g., text versus audio) leads to answers with a higher amount of information. We measure information content in several ways. First, we consider response length, which has been shown in previous research to vary for the formats. However, we argue that longer answers may not always indicate higher information content. Therefore, we also investigate the content and complexity of answers, irrespective of their length. In this second step, information content is measured via the number of topics derived from topic models and response entropy.

**RESULTS** Our results show that oral responses are on average longer across all nine questions examined in our study. Furthermore, our results also support the notion that audio answers tend to be slightly

more informative in terms of both the number of topics covered and response entropy.

**ADDED VALUE** Research on text- and voice-based response formats is very novel. We contribute to previous accounts by incorporating additional measures of information content and utilizing state-of-the-art speech-to-text technologies, providing valuable insights for the design of online self-administered surveys.

## ATTENTION PLEASE! COMPARING DIFFERENT WAYS OF PRESENTING AN INSTRUCTION MANIPULATION CHECK IN A PROBABILITY-BASED ONLINE PANEL

**HÖHNE, JAN KAREM {1}; RETTIG, TOBIAS {2};  
REVILLA, MELANIE {3}**

1: DZHW, University of Hannover, Germany

2: University of Mannheim, Germany

3: Institut Batcelona d'Estudis Internacionals, Spain

**RELEVANCE & RESEARCH QUESTION** Self-administered web surveys are a common method for measuring respondents' attitudes. Such web surveys are time- and cost-effective for researchers and allow respondents to participate with almost no time and situational restrictions. However, research has shown that respondents are frequently inattentive during participation (e.g., they do not read survey questions and instructions carefully), which potentially compromises data quality. Although many researchers recommend the inclusion of so-called "Instruction Manipulation Checks" (IMCs) asking respondents to prove their attention by engaging in a specific – for web surveys uncommon – action, empirically proven best practices are rare. Such best practices are key since respondents may fail IMCs due to misunderstanding and not just due to inattention. In this study, we attempt to close this gap by addressing the following research question: How to best implement IMCs in a probability-based online panel?

**METHODS & DATA** We conducted a survey experiment in the probability-based German Internet Panel (N = 4,348) in July 2020. Specifically, we randomly assigned respondents to one out of seven experimental groups differing with respect to text presentation (on one page or on several pages), question presentation (on the same page as the text or on a separate page), and IMC placement (first paragraph of the text or later paragraphs). To pass the IMC, respondents were asked to click on the German Internet Panel logo instead of the "next" button. Regardless of passing or failing the IMC, all respondents received a subsequent open question asking about the content of the previous page(s).

**RESULTS** The preliminary results indicate that the passing rate is lowest when presenting the entire text on one page (about 40%). Text without a question on the same page performs better in terms of passing rate. Placing the IMC in the first or later paragraphs of the text does not affect the passing rate. Non-compliance is low (about 2%).

**ADDED VALUE** This study adds to the state of research on gathering information about respondents' attentiveness during web survey participation. It provides empirical-driven recommendations for implementing IMCs and contributes to the discussion of non-compliance with IMCs.

## MEASUREMENT QUALITY OF A MULTI-ITEM SCALE IN PLAIN LANGUAGE

KUNZ, TANJA (1); GUMMER, TOBIAS (1,2); NEUERT, CORNELIA (1)

1: GESIS – Leibniz Institute for the Social Sciences, Germany

2: University of Mannheim, Germany

**RELEVANCE & RESEARCH QUESTION** There is general agreement that survey questions should be easy for all respondents to understand and answer in order to ensure good measurement quality. This is especially important in web surveys where there is no interviewer to assist the respondents. Translating survey questions into 'plain language' is one way to make them easier to understand. However, using different language variants raises doubts about whether the survey instruments are still comparable, whether the 'plain language' version is more straightforward for respondents to answer, and thus leads to better measurement quality. Consequently, we investigated the two research questions of whether (1) a multi-item scale in 'plain language' leads to less undesirable response behavior than the original version in 'standard language,' while (2) both language versions have similar scale properties.

**METHODS & DATA** With a between-subjects experimental design, we compared two versions of the Need to Evaluate Scale—the original version in 'standard language' and a translated version in 'plain language'—embedded in a web survey with participants of an online access panel in Germany (n=3,256). Measurement quality was assessed based on various response quality indicators (i.e., item nonresponse, response differentiation, acquiescence, extremity, midpoint responding, response times) and scale properties (i.e., dimensionality, internal consistency reliability, mean responses).

**RESULTS** Regarding our response quality indicators, the multi-item scale in 'plain language' showed higher response differentiation, fewer midpoint responses, and shorter response times than the original version in 'standard language.' Furthermore, we found comparable scale properties in both language versions.

**ADDED VALUE** Overall, the study provides promising results showing that questions in 'plain language' can measure the same content while reducing undesired response behavior in web surveys.

## B1: TEXT AS DATA

SESSION CHAIR: FRANZISKA KERN

infas 360 GmbH, Germany

## LARGE LANGUAGE MODELS FOR ASPECT-BASED SENTIMENT ANALYSIS

SIMMERING, PAUL FERDINAND; HUOVIALA, PAAVO

Q / Agentur für Forschung GmbH, Germany

**RELEVANCE & RESEARCH QUESTION** Large language models (LLMs) like GPT-4 offer unprecedented text processing capabilities. As general models, they can fulfill a wide range of roles, including those of more specialized models. We investigated how well GPT-3.5 and 4 perform for aspect-based sentiment analysis (ABSA). ABSA is used for providing insights into digitized texts, such as product reviews or forum discussions, and is therefore a key capability for market research and computational social sciences.

**METHODS & DATA** We assess performance of GPT-3.5 and 4 both quantitatively and qualitatively. We evaluate performance on the gold standard benchmark dataset SemEval2014, consisting of human annotated laptop and restaurant reviews. Model performance is measured on a joint aspect term extraction and polarity classification task. We vary the prompt and the number of examples used and investigate the cost-accuracy tradeoff. We manually classify the errors made by the model and characterize its strengths and weaknesses.

**RESULTS** Given 10 examples, GPT-4 outperforms BERT-based models trained on the full dataset, but does not reach the state of the art performance achieved by trained T5 models. The choice of prompt is crucial for performance and adding more examples improves performance further, however driving up the number of input tokens and therefore cost in the process. We discuss solutions such as bundling multiple prediction tasks into one prompt. GPT-4's errors are typically related to the idiosyncrasies of the benchmark dataset and extensive labeling rules. It struggles to pick up on the nuances of labeling rules, instead occasionally delivering more commonsense labels. While such errors hamper benchmark performance, they should not necessarily discourage from using LLMs in real-world applications of ABSA or similar tasks.

**ADDED VALUE** This study provides market researchers evidence on the capabilities of LLMs for ABSA. It also provides practical hints for prompt engineering and the cost-accuracy tradeoffs involved when using LLMs for structured extraction and classification tasks. By extension, it also helps with placing few-shot use of LLMs in contrast with finetuned models.

## STEREOTYPES IN SOCIAL MEDIA DATA: REGIONALIZATION WITH WORD EMBEDDING MODELS

**KNAUFF, STEFAN**

Bielefeld University, Germany

**RELEVANCE & RESEARCH QUESTION** Stereotyping is a common social practice that can lead to negative effects on all levels of society, from individual people to societies at large (Link and Phelan 2001). Research can contribute to a better understanding of stereotyping and stereotypes themselves, the contexts in which they emerge as well as the contexts in which they are used. Research findings may help to develop measures to counteract causes and negative effects of stereotyping. Within this presentation, we share first findings of our ongoing research project that operationalizes regional differences concerning stereotypes within a dataset of over 2 billion German tweets.

**METHODS & DATA** Different stereotypes can be clustered using the social dimensions of competence and warmth based on the Stereotype Content Model (Fiske et al. 2002). Charlesworth et al. (2021) have used the Word-Embedding Association Test (WEAT) to analyze gender stereotypes across several corpora of textual data. Building on the methods employed by Charlesworth et al. (2021), we use the stereotype content model by Fiske et al. (2002) as theoretical framing to uncover regional differences concerning different stereotypes. The empirical basis for the analysis is a corpus of about 2 billion German tweets, that was geo-localized using the free form location strings within the Twitter profiles, according to the method developed by Nguyen et al. (2022). For the analysis, we trained separate word embedding models for each analyzed region.

**RESULTS** Our first results show considerable regional differences regarding stereotyping practices on Twitter. Furthermore, using regionalized indicators from official statistics and regionalized survey data, we demonstrate how the found regional differences in stereotyping can be contextualized with data from additional data sources.

**ADDED VALUE** Social media posts are digital trace data (Lazer et al. 2009), providing different insights than traditional empirical social science methods, such as surveys, observations or interviews. While the empirical example for this presentation is Twitter, the presented methodological and theoretical approaches can be applied to other data sources as well.

## STATISTICAL ANALYSIS OF WEB BROWSING DATA: A GUIDE

**CLEMM VON HOHENBERG, BERNHARD (1); STIER, SEBASTIAN (1); CARDENAL, ANA-SOFIA (2); GUESS, ANDREW M. (3); MENCHEN-TREVINO, ERICKA (4); WOJCIESZAK, MAGDALENA (5)**

1: GESIS – Leibniz Institute for the Social Sciences, Germany

2: Universitat Oberta de Catalunya, Spain

3: Princeton University, USA

4: UC Davis, USA

5: American University, USA

**RELEVANCE & RESEARCH QUESTION** The use of digital trace data – data that emerge from people's interactions with digital systems – has become an important resource for researchers across disciplines. An important subset of trace data concerns people's activity when seeking out and consuming online information. Such data hold great potential for the social sciences. Yet, although recent literature provides some guidance when it comes to the collection of individual-level browsing data, there is no hands-on guide to help researchers avoid the pitfalls of processing and analysing them. Our overarching research question is how to avoid such problems and make the best of browsing data.

**METHODS & DATA** In individual-level browsing data, the unit of observation is a „visit“ to a URL by an individual connected to a timestamp. We structure our recommendations how to deal with such data along four stages: (1) pre-processing (2) filtering data (3) classifying data (4) modelling. To inform our recommendations at each of these steps, we make use of 9 different data sets from 7 different countries (US, UK, NL, DE, ES, FR, PL) collected with three different tools (Web Historian, Wakoopa, RealityMine) and across several years. We compare methods within data sets, across data sets, and make use of simulations.

**RESULTS** This paper yields many results, i.e. recommendations. We will just give two examples in this abstract. First, consider the necessity for some browsing data to create a „duration“ variable for a URL visit. We simulate different ways that can be used to create such a duration variable and derive recommendation for cutoffs. A second example concerns the classification of visits via domains: what is more fruitful, assembling extensive list of domains matching the target variable, or using automatic classification tools such as Web Shrinker? Again, we test two approaches against each other and describe benefits and shortcomings of each.

**ADDED VALUE** To our knowledge, this is the first project to assemble specific, data-driven recommendations (and ask some open questions), informed by researchers from several different teams, using a wide range of data sets.

# C1: MEDIA CONSUMPTION HABITS

[SPONSORED BY APROXIMA]

**SESSION CHAIR: BERNHARD CLEMM VON HOHENBERG**

GESIS – Leibniz Institute for the Social Sciences,  
Germany

## AN EXPERIMENTAL STUDY OF ONLINE INFORMATION SEEKING ON POLICY JUDGEMENTS

**ULLOA, ROBERTO (1,2); KACPERSKI, CELINA (1);  
KULSHRESTHA, JUHI (1,3); SPITZ, ANDREAS (1);  
BONNAY, DENIS (3); SELB, PETER (1)**

1: Konstanz University, Germany

2: Aalto University, Finland

3: Université Paris Nanterre, France

**RELEVANCE & RESEARCH QUESTION:** The internet provides an unprecedented platform for actively gathering political information, and the ambivalent influence that complex online experiences have on public opinion has sparked academic scrutiny. The plethora of correlative data sources showcases that the effects of internet and social media use on changes in political participation, attitudes, and voter behaviour are mixed but often significant [Zhuravskaya et al., 2020].

For example, the ongoing debate on echo chambers and polarization and whether or how they affect political decision-making is mainly supported by correlative or observational data (for a review, see Ross Arguedas et al., 2022; Terren & Borge-Bravo, 2021). Experimental studies on online information seeking that support causal conclusions are rare and generally conducted with curated search results, demonstrating patterns of selective exposure [Knobloch-Westerwick et al., 2015; Sude & Knobloch-Westerwick, 2022].

Our study advances prior literature by employing an experimental design in combination with unobtrusively collected behavioral web browsing data to investigate whether access to information from actively self-selected search queries and strategies in a real online environment (as opposed to curated or pre-packed information interventions) can affect political preferences.

**METHODS & DATA:** In a randomized controlled trial, across three data collection waves focusing on different policy topics, 1112 participants were recruited and randomly assigned to one of three groups: two intervention groups, receiving instructions to search for information related

to a specific policy, one intervention group was financially incentivized; and the control group, which did not receive any special instructions to search. Participants were queried on their political preferences twice, before the search task and 24 hours later, resulting in a between-within mixed design. The participants' online behaviour was tracked to measure and compare the policy-related online visits between the groups.

**RESULTS:** Preliminary results indicate that participants in the intervention groups (compared to the control group) show increased information seeking behaviour and knowledge, and only partial evidence of increased changes to their policy judgments.

**ADDED VALUE:** Research combining experimental interventions with behavioral data collected in realistic online environments via webtracking is exceedingly rare, but uniquely allows us to draw causal conclusions about the effects of online information seeking.

## MEDIA CONSUMPTION PATTERNS DURING THE RUSSIAN-UKRAINIAN WAR: THE ROLE OF PROXIMITY, MEDIA TRUST, AND INTEREST IN POLITICS IN GERMANY AND ISRAEL

**ELISHAR-MALKA, VERED (1); UNKEL, JULIAN (2);  
ARIEL, YARON (1); WEIMANN-SAKS, DANA (1)**

1: Academic College of Emek Yezreel, Israel

2: LMU Munich, Germany

**RELEVANCE & RESEARCH QUESTION** This study investigates news consumption patterns over legacy and social media in Germany and Israel during the first two months of the Russian invasion of Ukraine in February 2022. Consumption patterns will be analyzed as a function of individuals' objective and subjective proximity to Ukraine, interest in politics, and perceptions of the media, including trust, cynicism, and skepticism towards it.

**METHODS & DATA** A structured questionnaire was administered to nationally representative samples in Germany (n=1,310) and Israel (n=509) in May 2022. Data were analyzed using regression models to identify significant predictors of media consumption patterns in both countries.

**RESULTS** The consumption of legacy media in both countries was significantly predicted by greater objective and subjective proximity to Ukraine, increased political interest, and older age. Media trust and cynicism significantly predicted legacy media use in Germany but not in Israel, while media skepticism predicted greater legacy media use in Israel but not Germany. Social media use in Germany was significantly predicted by greater objective proximity, media cynicism, political interest, younger age, and being female. In Israel, greater objective and subjective proximity, lower media trust, and greater media skepticism were associated with increased social media use. Furthermore, the results reveal distinct differences in the factors predicting media consumption



between Germany and Israel, emphasizing the importance of considering cultural and regional contexts when examining media usage.

**ADDED VALUE** This study contributes to understanding how individuals' objective and subjective ties with a foreign country motivate them to consume mainstream and social media news about significant events. The findings highlight the significance of considering proximity and media system perceptions when attempting to comprehend how media are used to obtain information about crucial events such as the Russian-Ukrainian War. In addition, the findings support the importance of media trust as a predictor of people's tendency to rely on mainstream media rather than social media for information and news updates on significant events.

## THE PRIVACY PARADOX IN WHATSAPP: EXAMINING THE RELATIONSHIPS BETWEEN USAGE, CONCERNS, AND PROTECTION BEHAVIORS

LEVY, EILAT CHEN; ARIEL, YARON

Academic College of Emek Yezreel, Israel

**RELEVANCE & RESEARCH QUESTION** The widespread use of instant messaging apps such as WhatsApp has led to concerns about privacy and protecting personal information. This study aimed to explore the relationship between WhatsApp usage, privacy concerns, and privacy protection behaviors among Israeli WhatsApp users using the concept of the „privacy paradox“ (Kokolakis, 2017; Wirth et al., 2022).

**METHODS & DATA** A survey was conducted among 505 Israelis to gather data on their WhatsApp usage and perceptions of privacy. The sample was obtained from an online panel representative of the Israeli population based on the Central Bureau of Statistics data. The questionnaire items, mostly on a Likert scale, covered various topics related to WhatsApp usage, including frequency of use, understanding of privacy settings, and demographic information.

**RESULTS** The study found that 87% of respondents reported daily usage of WhatsApp, and 49% reported usage every hour. WhatsApp was used for various purposes, including interpersonal chats, group chats, and sharing multimedia content. The results also indicated that 40% of respondents had not changed their „last seen“ privacy setting, and 70% had not changed their profile picture privacy setting. There was no significant correlation found between the frequency of WhatsApp usage and privacy protection behaviors ( $r=.037, p>.05$ ); however, a significant positive correlation was observed between the frequency of usage and privacy concerns ( $r=.117, p<.05$ ). Privacy concerns were related to privacy protection behaviors ( $r=.525, p<.001$ ) and served as a mediator in the relationship between WhatsApp usage and privacy protection behaviors ( $z=2.424, p<.05$ ).

**ADDED VALUE** This study provides valuable insights into the use and perceptions of privacy on WhatsApp in Israel. The results highlight the need

for policies and practices to protect user privacy on instant messaging apps, especially for interpersonal communication. In addition, the findings contribute to the growing literature on IM apps and privacy by identifying the factors that influence users' attitudes and behaviors toward privacy. Further research is needed to examine these relationships in other populations and to explore the implications for privacy protection in the digital age.

## T1: GOR THESIS AWARD 2023 COMPETITION: PHD

(SPONSORED BY TIVIAN)

SESSION CHAIR: OLAF WENZEL

Wenzel Marktforschung, Germany

## APPROACHES TO DEALING WITH SURVEY ERRORS IN ONLINE PANEL RESEARCH

KOCAR, SEBASTIAN

University of Tasmania, Australia

Web surveys including online panel surveys are now routinely used for collecting survey data, and the subject matter of this thesis pertains to survey research conducted via online panels. To increase the quality of online panels in the era of nonresponse, arguably more methodological research is needed, and that is the focus of this research. The thesis and its chapters aim to answer the following overarching research question: What are the most prevailing survey errors in online panel research (with the largest impact on data quality/accuracy), and what are the most suitable approaches to dealing with them, including those for identification, reduction, correction, and balancing of survey errors?

To investigate approaches to dealing with survey errors, the Total Survey Error (TSE) paradigm as a conceptual framework is applied, and both errors of representation and errors of measurement are the subject of this research. Rather than providing an exhaustive analysis of all errors in online panel surveys, this thesis centres on examining the errors that have the greatest impact on total survey error in both online panel surveys and web surveys. The nine thematic chapters of this thesis concentrate on novel or refined techniques for identifying, comparing, and mitigating errors, biases, and sources of errors in online panel research. Besides, one of the notable contributions of this thesis is a review and discussion of a range of data sources, including primary and secondary data, experimental data, cross-sectional and longitudinal survey data, paradata and qualitative interview data. Also, a range of analytic

methods, including panel data analysis, meta-analysis, qualitative data analysis, and weighting and matching methods, are applied (and some of them compared) in the study of survey errors. The other theoretical and practical contributions, presented within three groups, are related to the investigation of individual types of survey errors and approaches to mitigate them, in online panel research.

First, worldwide probability-based online panels are identified, and their methodological approaches to recruitment and data collection thoroughly reviewed and compared as part of a meta-analysis. The presented evidence shows high levels of heterogeneity in both recruitment rates and recruitment solutions. Furthermore, it clarifies the variability in recruitment rates that can be attributed to factors like year of recruitment and panel's geographic location, as well as recruitment characteristics such as mode and incentives use. The other studies on errors of representation in this thesis present evidence on how online panel paradata can be effectively transformed and used to identify about three in four nonrespondents in a subsequent panel wave, and answer the question of why people participate in online panel surveys. That includes presenting evidence on how social-psychological theories can explain survey participation in a longitudinal design. The study which utilizes a predominantly qualitative approach, which is a relatively new methodology in the study of survey errors, identifies a range of motivations and barriers that are specific to the panel nature, and explores how these factors interact in the decision-making processes of panellists.

Second, two studies focus on measurement error in probability-based online panel research due to mixing modes. The study on measurement mode effects shows how measurement error is present in the case of a lack of measurement equivalence between modes, although the impact on the results is mostly relatively minor. However, due to the quasi-experimental survey design utilized, the analysis also controls for two potentially competing sources of differences in responses across different modes: mode effects and mode self-selection effects. Thus, this study also presents evidence on how applying matching methods (like coarsened exact matching) quite effectively controls for self-selection bias due to non-random assignment of online panellists to modes. Furthermore, the study on individual-level measurement mode effects presents a newly identified source of measurement error in online panel survey – panel measurement mode effects. The findings indicate that panellists alternating between online and telephone modes can have an impact on response stability, as well as potentially leading to social desirability and recency biases. The study also conceptualizes and showcases how panel conditioning can be a factor of two measurement aspects: the frequency of repeated measurements and the time gap between measurements. These results are later related to a trade-off between representation and measurement bias.

Third, the thesis studies two cost- and time-efficient approaches to online data collection – nonprobability online panels and a fairly new combination of random digit dialling (RDD), text message invitations, and web-push principles. The study on nonprobability panels, which are generally considered as less accurate but cheaper than probability-based panels, investigates post-survey adjustment methodology to improve inference in nonprobability samples. It presents evidence on how the quality of post-survey adjustments is mostly dependent on the avail-

ability of relevant high-quality covariates which are associated with either representation bias in nonprobability samples and/or outcome variables. Finally, the study on RDD text-to-web approach shows very low response rates, and outlines effective solutions to increase response. It also presents evidence on the fairly high accuracy of the proposed approach, which seems to be feasible for continuing recruitment to a probability-based online panel as well.

Due to a range of relevant phenomena investigated in this thesis, it can be potentially considered as the most comprehensive recent overview of the state-of-the-art of online panel research. This includes the investigation of the cost dimension of probability-based online panel research. For example, it is noteworthy that organizations managing these panels are increasingly choosing more affordable strategies for recruitment, such as moving from telephone to postal recruitment, as well as carrying out end-of-survey recruitment. The studies presented in this thesis shed light on several novel and existing cost-saving solutions for recruitment and panel maintenance, including the utilization of online panel paradata and effective communication techniques (e.g., reminders, prenotifications, and recruitment material content) to boost response rates. In addition to theoretical contributions, such as the introduction of panel measurement mode effects as a new source of measurement error in the theory of response process, the thesis offers practical insights. Notably, it discusses the requirements of collecting data from the offline population in probability-based online panel research, which are challenged from various perspectives.

## RESPONDENT AND RESPONSE BEHAVIOR IN ONLINE PANEL SURVEYS

RETTIG, TOBIAS

University of Mannheim, Germany

**RELEVANCE & RESEARCH QUESTION** In the field of survey research, the prevalence, popularity, and number of online surveys and panels has dramatically increased over the past decades. Whenever researchers draw conclusions based on survey data, they need to be sure that their data are accurate and of high quality. Researchers expect respondents to carefully consider each question and give a truthful response that reflects how they feel about the topic at hand; their “true” answer. Any violation of this principle may lead to responses that inaccurately capture the information researchers sought to measure (i.e., measurement error) and thus, researchers may draw inaccurate or incorrect conclusions. A wealth of survey-methodological literature has been devoted to measurement error over the past decades. This thesis adds to the body of literature on measurement error by expanding the conceptual understanding and the empirical evidence on two potential sources of error in the response process: Respondent inattention (respondents who do not read instructions, questions, or response options and instead select their response arbitrarily) and memory effects (respondents who remember their responses to previous iterations of repeated survey questions and are influenced by this information in their later answers). Each chapter of the cumulative dissertation focuses on a different aspect of these two sources of measurement error: The first article



experimentally investigates whether respondents read experimental treatments depending on text length. The second article provides a conceptual framework of how memory of previous responses may influence respondents' cognitive response process, an overview of the existing literature on memory effects and adjacent concepts such as question order effects, and integrates memory effects into the wider survey-methodological literature. The third article investigates respondents' ability to remember previous responses with the same survey after about 20 minutes across different types of questions (on attitudes, behaviors, and beliefs), thus providing insights for survey designs that incorporate repeated questions after a short time (e.g., pretest-posttest designs, MTMM). The fourth article investigates memory effects in the longer term by studying respondents' ability to remember their responses from a previous panel wave after four months, thus providing insights for measurement repetitions in longitudinal designs.

**METHODS & DATA** This thesis uses data from multiple survey experiments implemented in the German Internet Panel (GIP), a probability-based online panel of the German population ages 16 to 75 years at the time of recruitment. To investigate respondent attention, respondents in one panel wave (N=4,028) were randomly assigned to receive a treatment text of varying length with a hidden instruction to click the survey logo instead of answering the question (instruction only, one, two, or four paragraphs). Passing rates of this attention check, their relation to text length, respondent characteristics, and response speed are examined. To investigate short-term memory effects, half of all respondents in one wave (N=1,858) were randomly selected and received two questions at the beginning of the questionnaire. Question type was varied between questions on attitudes, behaviors, or beliefs. At the end of the questionnaire, respondents were asked to indicate whether they could remember their answer, what their answer had been, and how certain they felt about correctly recalling their answer. Using correct repetitions of the original answer despite self-reporting not to remember it as a baseline for correct repetitions not due to memory (i.e., stable underlying information or chance), the presence of memory effects and differences across question types and respondent groups are investigated. To investigate memory effects in the longer term, these follow-up questions were two panel waves (4 months) later again presented to all respondents who participated in both waves (N=3,809).

**RESULTS** Results show a link between text length and whether respondents read the text with passing rates ranging from 79.5% for the shortest to 41.1% for the longest text condition. Attention levels are higher for higher age, female respondents, higher education, freshly recruited respondents, smartphone respondents, and differ across the field period and daytime. Response speed is correlated with attention but not suitable as a standalone indicator. A systematic misclassification of fast readers (in turn correlated with age and education) as inattentive is likely. Respondents remember their previous responses after 20 minutes in most cases. Memory differs across question types. Extreme responses are more likely remembered. Younger, higher educated, and female respondents are more likely to remember their responses. Evidence for memory effects is present after 4 months (i.e., not all correct repetitions of previous responses are explained by stable information or chance). Memory differs across question types. However, incorrect repetitions are most commonly off by only one scale point. The

group of respondents who remember their responses is small and little evidence suggests they are systematically different from respondents who do not remember, the only significant difference being gender.

**ADDED VALUE** The thesis contributes to the existing literature in several ways. Conceptually, it provides a framework for understanding memory effects and their role in respondents' cognitive response process and integrates memory effects into the wider survey-methodological literature. The empirical results also have concrete implications for practitioners: A considerable number of respondents will not read excessively long texts and receipt of the intended treatment systematically varies across socio-demographic groups. Comparisons across these groups and removal of inattentive respondents are thus problematic and likely to introduce bias. Instead, texts should be kept short, and the demands of the survey adjusted to fit what respondents are able and willing to provide. Response time as a standalone attention indicator is likely to systematically misclassify respondents and should thus not be used. Respondents are likely to remember their response within one survey. Memory differs across question types, socio-demographic groups, and possibly attitude strength. Different questions and groups of respondents may thus be differently affected. Researchers should thus use any designs with measurement repetitions after a short time with caution (e.g., test-retest, pretest-posttest, MTMM). Longer-term, memory effects are still present for some respondents after 4 months but the group is small and similar to unaffected respondents. Measurement repetitions in longitudinal settings may thus have negligible measurement error due to memory effects with four or more months between repetitions.

## ONLINE PRODUCT TESTING FOR GENERATING ONLINE PRODUCT REVIEWS – AN ANALYSIS OF THE EFFECTS ON REVIEW RATING AND REVIEW QUALITY

**TESCH, TABEA**

Royal Canin (Mars Incorporated), Germany

**RELEVANCE & RESEARCH QUESTION** Online product reviews are of great importance to companies. Positive reviews increase customers' willingness to buy a product (Marchand et al. 2017) and their willingness to pay a higher price (Kübler et al. 2018). Moreover, readers of online reviews perceive high-quality reviews as more helpful and hence are more likely to follow their recommendation (Lu et al. 2018). Therefore, online reviews can directly affect a company's performance.

A frequently used approach to encourage customers to provide online reviews is product testing (Chae et al. 2017). With product testing programs customers are invited to test a product and then to describe their impression of the product in an online review. Many companies typically give a product away free of charge, or offer it at a reduced price, in exchange for the promise of writing an online review. Providers of online product testing programs are manufacturers (e.g., Philips, „Frag Team

Clean“ by Henkel and „Vocalpoint“ by P&G), retailers (e.g., Amazon, dm, Rewe, Media Markt and Tchibo) and agencies (e.g., Konsumgöttinnen, Kjero, Home Tester Club, oh! of the day, Toluna and Trnd).

While product testing programs inevitably lead to a higher number of online reviews for the test product, it remains uncertain whether these reviews will be positive and of high quality. Therefore, the dissertation aims to answer the following research questions: Does product testing influence reviewing behavior in terms of review rating and review quality? Do context factors like the product price, product complexity, and the number of previously published reviews on the product have an impact on these effects? What underlying psychological mechanisms explain the effect of product testing on reviewing behavior?

**METHODS & DATA** Four studies were conducted to analyze reviewing behavior in product testing programs. In Studies 1 and 2, qualitative insights into the feelings and perceptions of product testers were gained by administering an open-ended questionnaire to 100 participants from various product testing programs in the United Kingdom (Study 1), and by conducting in-depth interviews with 12 experts, namely, Amazon Vine product testers from Germany (Study 2). Study 1 data were collected online, and participants provided written responses; Study 2 data were gathered through video conferences, recorded, and transcribed.

In Study 3, the theoretically derived psychological mechanisms and the Study 1 and 2 findings related to product testers' reviewing behavior were tested with an experimental approach. More specifically, it was examined how perceived inequity and perceived pressure mediate the relationship between product testing and review rating and the relationship between product testing and review quality. A posttest-only control group design (Campbell and Stanley 1963) was used and the type of review manipulated (product testing versus non-product testing) by randomly assigning participants to one of two groups. Three hundred participants were recruited from Prolific, a well-established platform for online research (Paharia 2020). To test the hypotheses the PROCESS procedure (Hayes 2018) was employed.

With Study 4, the Study 3 findings were expanded with field data from Amazon Vine, one of the world's largest product testing programs, by investigating the potential moderating effects of product price, product complexity, and previously published review volume. Using a customized Python-based web data crawler, publicly available review data – 207,254 online reviews written by 400 Amazon Vine reviewers – on the German version of Amazon's web site (Amazon.de) were retrieved. Relying on a multilevel analysis, reviews written by customers enrolled in the Vine program (who received products free of charge) were examined and compared with reviews of other products that they purchased. All reviews were published in German between September 2000 and March 2020. The featured products spanned 19 categories, such as books, electronics, and toys. As online reviews are nested within reviewers, multilevel analysis was performed.

**RESULTS** Product testing programs do not necessarily prompt better quality or higher review ratings; rather, the effect depends on the context, as established by product prices, product complexity, and previously published review volume. Customers taking part in prod-

uct testing programs offer more positive ratings for higher priced and more extensively reviewed products. Moreover, when the test product is high priced or complex, participants in the program offer higher quality reviews. In contrast, if many reviews already are available, product testers tend to devote less effort and thus produce a review of lower quality. These findings offer important implications for marketing theory and practice.

Two opposing psychological mechanisms help to explain product testers' reviewing behavior: perceived outcome-to-input ratio and perceived pressure. On the one hand, the perceived positive outcome-to-input ratio in case of higher-priced products, as explained by equity theory (Adams 1963), increases the effect of product testing programs on review rating and quality. On the other hand, a feeling of restricted behavioral freedom, as explained by the theory of psychological reactance (Brehm 1966), may be triggered. This sentiment can reduce review ratings and review quality, potentially cancelling out the positive effects of product testing programs via perceived inequity.

**ADDED VALUE** When designing product testing programs, companies might seek three distinct goals: increase the number of reviews (volume), increase product ratings, or increase review quality. The dissertation offers guidelines for how managers can leverage product testing programs to achieve their specific goals.

The theoretically derived and experimentally confirmed psychological mechanisms contribute to marketing theory and offer further managerial insights. Both perceived inequity and perceived pressure can be leveraged with appropriate program designs.

# A2:RESPONDENT BEHAVIOR AND DATA QUALITY

[SPONSORED BY GESIS]

**SESSION CHAIR: TANJA KUNZ**

GESIS - Leibniz Institute for the Social Sciences,  
Germany

## ATTITUDES OVER TIME – DOES REPEATED INTERVIEWING TRIGGER REFLECTION PROCESSES IN RESPONDENTS?

**KRAEMER, FABIENNE (1); LUGTIG, PETER (2);  
STRUMINSKAYA, BELLA (2); SILBER, HENNING (1),  
SAND, MATTHIAS (1); BOSNJAK, MICHAEL (3); WEISS, BERND (1)**

1: GESIS – Leibniz Institute for the Social Sciences, Germany

2: Utrecht University, The Netherlands

3: Trier University, Germany

**RELEVANCE AND RESEARCH QUESTION** Longitudinal surveys are widely acknowledged for their possibility to track attitudes over time and thus, to study change and stability of social patterns. However, previous research suggests that repeated survey participation itself affects respondents' attitudes over the course of a study by raising awareness and triggering reflection processes on surveyed topics (cognitive stimulus model). Former studies have provided first evidence by documenting more stable and reliable survey responses and increased opinionation among experienced panelists. However, most studies were non-experimental, and the underlying mechanisms of observed attitude changes remain largely unknown.

**METHODS AND DATA** We investigate attitude change over time due to repeated interviewing by using a longitudinal survey experiment comprising six panel waves. We manipulated the frequency of receiving identical question content (target questions) and assigned respondents to one of three groups: One group received the target questions in each wave (fully conditioned), the second group received the target questions in the last three waves (medium conditioned), and the control group received the target questions only in the last wave of the study (unconditioned). The experiment was conducted within a German non-probability ( $n = 1,946$ ) and a probability-based panel study ( $n = 4,660$ ), resulting in 2x3 experimental groups in total. Our analysis focusses on between-group and within-group comparisons of abortion attitudes measured by a 4-item battery. Additionally, we use measures on self-reported knowledgeability about the topic of abortion and attitude strength to further investigate underlying mechanisms of change.

**RESULTS** Preliminary results of a multiple-group confirmatory factor analysis show no significant differences in factor loadings and mean-structures between the six groups, indicating no effect of conditioning frequency on abortion attitudes. The structural equation model including knowledgeability and attitude strength shows significant effects of attitude strength on abortion attitudes across groups. The results of additional within-group analyses will be presented at the conference.

**ADDED VALUE** Altogether, our experimental study contributes to the understanding of what causes attitude change in panel studies by examining different levels of exposure to survey content and including measures of attitude strength and knowledgeability as indicators of triggered reflection processes.

## SURVEY PROFESSIONALISM: NEW EVIDENCE FROM TRACE DATA

**CLEMM VON HOHENBERG, BERNHARD (1);  
VENTURA, TIAGO (2); WOJCIESZAK, MAGDALENA (3);  
NAGLER, JONATHAN (2); MENCHEN-TREVINO, ERICKA (4)**

1: GESIS – Leibniz Institute for the Social Sciences, Germany

2: New York University, USA

3: UC Davis, USA

4: American University, USA

**RELEVANCE & RESEARCH QUESTION** Opt-in panels have become an important resource for low-cost survey research in the social sciences – but the sheer number of platforms offering financial compensation makes it attractive to become a „survey professional“, and there is a concern that survey results are skewed by the presence of such professionals. However, there is little concrete evidence about how widespread survey professionalism really is, and whether it indeed has any adverse consequences. Using three survey samples from the US, each of which is linked to individual browsing records, we ask: How much browsing of survey respondents consist if survey taking (RQ1)? How do survey professionals differ from non-professionals in terms of socio-demographics and political variables (RQ2)? Are survey professionals more careless in responding than non-professionals, measured by the stability of responses across waves (RQ3)?

**METHODS & DATA** We rely on three US samples from between 2018 and 2019: A sample recruited in 2018 through ads on Facebook; a sample acquired in 2019 for a three-wave survey from Lucid; sample recruited in 2019 from Yougov. All three data sets combine survey responses with individual-level records of browsing behavior before the survey. To decide whether web visit counts as survey taking, we combine three approaches. First, we rely on a report of questionnaire software published by Bevec and Vehovar (2021). Second, assuming that the names of web addresses provide clues about their content, we classified all url hosts that contained the word „survey“ as survey taking. Third, we had research assistants manually code the 500 most frequently visited URL hosts from each data set for whether people are taking surveys on them.

**RESULTS** Depending on the data set, survey taking ranges from 15 to 45 % of all browsing (RQ1). Analyses for the other research questions are being run.

**ADDED VALUE** While some work has shed light on the problem of survey professionalism via self-reports, to our knowledge, this is the first paper exploring the issue with behavioral data. Our first results raise serious concerns about data quality of online samples.

## THE IMPACT OF RESPONDENT EXPERIENCE ON RESEARCH QUALITY: UNDERSTANDING THE IMPORTANCE OF QUALITY SURVEY DESIGN AND PANEL ECOSYSTEMS

WITTMANN, EVA; CARRE, CECILE  
Ipsos, Germany

**RELEVANCE & RESEARCH QUESTION** Many current discussions about quality in research center around identifying and removing respondents who provide low-quality responses. While this is an important aspect of ensuring research quality, it overlooks the fact that some respondents may be well-intentioned but still provide unsatisfactory responses due to factors such as suboptimal survey design or panel experience.

In recent times, there has been a consensus in the industry that a respondent-centric approach is critical for the future of online research. Rather than solely focusing on penalizing certain behaviors, we advocate for a framework that identifies the drivers and motivations behind respondent behavior. Our goal is to not only address the behavior of so-called fraudsters but also to improve the behavior of respondents who can provide better responses through improvements to our questionnaires and ecosystems.

Our presentation will cover the following topics:

- Segmentation of different types of respondents and strategies for improving the behavior of those who provide suboptimal responses
- An outlook of how panel ecosystems influence certain types of behavior and suggestions for improving them in the future

**METHODS & DATA** We are showing learnings generated through several meta-analysis run across the Ipsos Global Online business.

**RESULTS** to be finalised in August 2023

**ADDED VALUE** We think a discussion needs to be sparked on how we can use respondent resources in a more sustainable way to future proof online research. We will show the approaches and thinking we at Ipsos use to incorporate respondent centricity into our quality systems.

## B2: ML AND AI IN SURVEYS

**SESSION CHAIR:** PAUL FERDINAND SIMMERING  
Q Agentur für Forschung GmbH, Germany

## AUTOMATED SPLIT QUESTIONNAIRE DESIGN: THE WAY FORWARD IN SURVEY RESEARCH?

WEITZEL, DANIEL (2,1); TSCHIATSCHEK, SEBASTIAN (1);  
RITTEL, SIMON (1); PFAFF, KATHARINA (1); KRITZINGER, SYLVIA (1)  
1: University of Vienna, Austria  
2: Colorado State University, USA

**RELEVANCE & RESEARCH QUESTION** Ever decreasing response rates in face-to-face, paper, or telephone surveys as well as the relatively low costs of online surveys have led to a widespread adoption of the online survey mode. This new mode, however, faces its own challenges, particularly when working with offline-recruited samples. One difficulty researchers face is obtaining representative samples, another challenge is the mode-specific need for shorter surveys.

**METHODS & DATA** To address these challenges, we propose and evaluate a split questionnaire design to overcome the survey length constraints, extending a proposal of Axenfeld et al. [2022]. We introduce a machine learning approach that can automatically generate n split questionnaire designs that maximize across-split imputability of survey responses. Based on existing surveys and corresponding responses, our approach generates survey design suggestions by allocating questions into n different questionnaire splits. Questions are assigned to each split such that across-split information gain for omitted responses is maximized.

**RESULTS** First evidence suggests that the approach increases the amount of information researchers can obtain from the same sample size through an automated process that learns from previous surveys. Eventually, we also test to which extent we obtain representative samples by applying this approach.

**ADDED VALUE** Our approach allows us to conduct shorter questionnaires which may increase the quality of responses in a saturated "market of respondents".

## STANDARDIZED ANNOTATIONS FOR SURVEY DATASETS: ENABLING AUTOMATED QUALITY ASSURANCE AND EVALUATION

RATHJE, MARTIN (1); HUXOLL, JOHANNES (1);  
MCDONNELL, SOPHIA (1); WITTENBURG, GEORG (2)

1: Kantar Public, Germany

2: Inspirient GmbH, Germany

**RELEVANCE & RESEARCH QUESTION** To reliably automate processing of survey datasets – as we introduced in last year's GOR – it is critical to reduce the ambiguity inherently contained in raw survey data. This, for example, includes the question whether a dimension is a meta, socio-demographic or a response variable; or the selection of which time dimension can be used for wave comparisons. To address this problem, we propose a set of standardized data annotations that enable automated processing across heterogeneous datasets. Our research question is thus to what extent annotations may facilitate more efficient and accurate survey data processing.

**METHODS & DATA** After introducing Inspirient's AI system for automated survey data processing at Kantar Public last year, we evaluated which steps of automated processing are most susceptible to ambiguities, thus obstructing efficient processing and leading to costly manual readjustments. We distilled these experiences into a set of data annotations that address these ambiguities and formalized them in DIN SPEC 32792. The updated AI system now incorporates these annotations in its internal reasoning process, and we evaluated their effectiveness in multiple real-world survey data processing tasks.

**RESULTS** Our results show that the proposed annotations can help facilitate fully automated processing of survey data without the need for iterative manual readjustments, e.g., for quality assurance and analytical evaluation, while maintaining quality and consistency of results. Examples included in our talk cover basic and advanced use, incl. socio-demographic / response annotations to generate contingency tables as well as dependent / independent variable annotations to automate multivariate regression analyses.

**ADDED VALUE** Our work provides a foundation upon which more efficient automation / AI initiatives in the field of social research / survey analytics may be built. We illustrate the usefulness of this foundation by applying the proposed annotations to the automated process that Kantar Public rolled out last year, reconfirming gains in efficiency and completeness of results. Furthermore, we note that these annotations also reduce ambiguity in data interpretation among researchers, facilitating cross-team collaboration and archival / reuse of historic datasets.

## USING AI TO ENHANCE QUALITATIVE

MILOSEVIC, LARA-ELENA

Bonsai GmbH, Germany

**RELEVANCE & RESEARCH QUESTION** The study was conducted to understand how the generation under 30 (Gen Z) perceives the topic of female period and the growing number of explicit displays in advertisement. We further wanted to understand how AI can help to detect and overcome socially desirable answers in this research and how the decoding of non-verbal communication can further enriches the quality of the outcome. We want to learn how in how far artificial intelligence can become a substitute for qualitative in-depth research in the near future.

**METHODS & DATA** In 16 in-depth interviews we confronted the target (GenZ) with (explicit) imagery of female periods in advertisements. Projective techniques and collages were used to approach unconscious behaviour and emotions. The interviews were reviewed by an artificial intelligence that decoded emotions using facial coding & voice AI. The technology derives human emotions from micro expressions and voice tonality. It tells how confident a person is in what they are saying. A low confident score indicates insecurity and/or socially desirable answers. The results were put against conventional language analytics to see in how the learnings are detectable in participants way of speaking.

**RESULTS** The explicit portrayal of female periods is highly emotional across the targets. The hypothesis that the younger generation is more open in regards to female period does not hold true. Emotion AI was able to underpin the findings with clear indications on emotions like happiness, surprise, disgust or neutral and helped detecting socially desired answers. While the wording appeared open and positive, the emotion AI showed a dropping confidence score. Yet, this also becomes visible in (classic) language analytics. The growing discomfort shows in a more hesitant articulation and increased use of passive or a third-person narrative.

**ADDED VALUE** We want to showcase how emotion AI can benefit qualitative research in adding data and proof pre-existing hypothesis. How it helps to spot, analyse and even control socially desirable answering behaviour. But also, how at the end of the day, it is still the combination of human and artificial intelligence that is the silver bullet.



# T2: GOR THESIS AWARD 2023 COMPETITION: BACHELOR/ MASTER

[SPONSORED BY TIVIAN]

SESSION CHAIR: **OLAF WENZEL**

Wenzel Marktforschung, Germany

## HOW TO COMMUNICATE IN A CRISIS? THE INFLUENCE OF LOCAL GOVERNMENTS' HUMOROUS CRISIS RESPONSE STRATEGY AND CRISIS RESPONSIBILITY ON TRUST, EMOTIONS AND BEHAVIORAL INTENTIONS

**HÄMPKE, JANNA**

University of Münster, Germany;

Ludwig Maximilian University of Munich, Germany

**RELEVANCE & RESEARCH QUESTION** When choosing how to communicate on social media during a crisis, local governments should consider target group-specific communication as well as their own crisis responsibility. This study is the first to investigate how a local government's humorously framed response strategy to a low-severity crisis on social media influences people's emotions, trust in the local government and their crisis-related behavioral intentions, specifically when considering the government's crisis responsibility. According to Situational Crisis Communication Theory (SCCT, Coombs, 2007), central factors affecting stakeholders' emotional, attitudinal and behavioral responses to a crisis and the organization involved are the level of the organization's crisis responsibility and its crisis response strategy, which should match its crisis responsibility (e.g., Coombs, 2007). When the attributed crisis responsibility is high, the crisis can reduce stakeholders' trust in the organization, decrease supportive behavior and elicit negative emotions towards the organization, which mediate between attributed crisis responsibility and people's trust in an organization and behavioral intentions (e.g., Coombs, 2007; McDonald et al., 2010).

To decrease negative affect and increase positive affect, several authors have suggested to use humor in crisis response strategies in low-severity crises (Kim et al., 2016; Vigso, 2013). Humor can lower stress levels and the intensity of and the focus on negative emotions.

Instead, it elicits feelings of happiness or cheerfulness (Martin, 2007) enhancing stakeholders' attitudinal and behavioral reactions toward an organization. Online crisis communication researchers have already found the positive impact of humor on the relationship between an organization and its stakeholders during or after a crisis (e.g., Vigso, 2013). However, results are limited to private organizations and give first hints that inappropriate humor may threaten people's trust in an organization and discourage them from taking appropriate actions (Bitterly, 2022). Hence, it is necessary to extend existing results to public organizations and to further investigate the conditions under which humorous crisis communication might be useful.

This study contributes to this topic by examining the effect of local governments' crisis responsibility and their humorously framed crisis response strategies posted on Twitter to a low-severity crisis on trust in the local government, positive and negative affect and behavioral intentions towards the public organization, namely the public's willingness to follow local governments advice, to share their crisis communication with others and to seek for crisis-related information. In addition, it is tested whether experienced affect mediates the relationship between crisis responsibility or humorously framed crisis response strategy and trust or behavioral intentions. Furthermore, I question whether the influence of humorously framed crisis response strategies on trust in the local government and on positive and negative affect is dependent upon whether the local government is responsible for the low-severity crisis.

**METHODS & DATA** A total of 517 German participants (aged between 18 and 74 years) took part in an online experiment in which they were confronted with three randomly presented fictive crisis scenarios where the local government's attributed crisis responsibility (high vs. low) as well as the framing of their crisis response strategy (humorous vs. rational) were systematically varied between subjects. Two pre-studies ( $n_1 = 10$ ,  $n_2 = 13$ ) were conducted to select the scenarios and the respective crisis responses. A series of MANOVAs, follow-up ANOVAs and mediation analyses were examined to check the hypotheses. In exploratory analyses I also tested the predictive power and moderating role of demographics and participants' sense of humor and disposition to trust and the crisis scenarios themselves as well as participants' liking of the tweets in dependence of both experimental factors.

**RESULTS** First, crisis responsibility had a significant negative effect on negative affect, trust and behavioral intentions, but not on positive affect. Negative (but not positive) affect mediated the effect of crisis responsibility on trust and some behavioral intentions variables. Second, a negative effect of humorously framed strategies on trust and positive affect (but not on negative affect and behavioral intentions) was found. Positive (but not negative affect) mediated the effect of framing on trust (but not on behavioral intentions).

The crisis responsibility x framing interaction was not significant. Exploratory analyses indicated that individual characteristics and the scenarios themselves moderated the effect of crisis responsibility and framing on trust. Additionally, it was shown that participants liked the humorous framings less than the rational ones.

**ADDED VALUE** This study presents the first experiment on the effectiveness of humor in local governments' crisis communication on social media, its conditions and working mechanisms. Furthermore, it contributes to a growing body of evidence for the SCCT [Coombs, 2007], suggesting that an organization's crisis responsibility is one of the most important factors in crisis communication due to its negative effect on negative affect, trust and behavioral intentions. In contrast to most findings in crisis communication research in private sectors, the results mainly argue against the use of humor in crisis communication on social media due to its negative effect on trust and positive affect, and missing effect on behavioral intentions.

Therefore, independent from the level of crisis responsibility humorous communication cannot compensate for the perceived negative impact of this responsibility. However, exploratory analyses indicate that humorous crisis communication on social media might be appropriate under specific circumstances in low severity crises if both target group and framework conditions (e.g., crisis characteristics) are suitable. Future studies on possible moderating and mediating variables, such as perceived crisis severity, are necessary.

## HOMOPHILY IN ONLINE FRIEND SELECTION. EMPIRICAL INVESTIGATION OF FRIEND SELECTION IN THE APP FRIENDSUP

**RIEGE, RAPHAEL**

Kassel University, Germany

**RELEVANCE & RESEARCH QUESTION** Homophily is a basic organizing principle in network formation. This means that people tend to prefer people who are similar to themselves. This thesis investigates the role of similarity in the process of selecting friends online and contributes to the larger body of research on homophily and online social connection.

The impact of homophily on social connections and outcomes is an important area of study, as it has the potential to reproduce and reinforce existing social inequalities.

**RQ 1:** In which characteristics is homophily evident in online friendship choices?

**RQ 2:** How is friendship choice influenced by digital infrastructures?

**METHODS & DATA** The data analyzed was real-world selections from the friendsUp app, a platform that allows female users to connect with potential friends (ff 6 million cases). The study analyzes the relationship between similarity and the likelihood of two users connecting inside the application.

The similarity dimensions analyzed were sociodemographic, cultural, and visual dimensions such as age, race, occupation, self-description, and face similarity. To examine written descriptions, text mining ap-

proaches were used. For analyzing face similarity, an AI-assisted method for face recognition was used.

**RESULTS** In summary, the analysis of the friendsUp app data confirms the existence of homophily effects, where users tend to seek friends with similar demographic, cultural, and visual characteristics. Specifically, age, motherhood, profile description, occupational status, race, and physical appearance affect the probability of initial contact and response. The analysis also shows that the number and order in which profiles are displayed, has a significant effect on the probability of selection.

**ADDED VALUE** Homophily also exists in online social networks and contributes to segregation of social groups. Physical appearance, race, and age or occupational status influence friendship formation in online social networks, which can lead to discrimination or exclusion. Understanding the processes of friendship formation in online social networks can inform the design of social networking applications, promote healthy and inclusive online social interactions, and mitigate the potential negative effects of homophily in digital friendships.

## DATA LINKAGE TO VALIDATE AND CALIBRATE TRAFFIC ESTIMATIONS ON A NATIONWIDE SCALE: A FRAMEWORK FOR OFFICIAL STATISTICS

**BOSTANCI, INAN ERAN (1); GOOTZEN, YVONNE (2); LUGTIG, PETER (3)**

1: Zuse Institute Berlin, Germany

2: Statistics Netherlands, The Netherlands

3: Utrecht University, The Netherlands

**RELEVANCE & RESEARCH QUESTION** Traffic estimation is an important tool in official statistics and can inform policy makers in their decision-making process for regional planning. Due to the complexity of traffic estimation and limits in data access, traffic demand modeling often focuses on smaller geographic areas and short-term forecasts. Statistics Netherlands (CBS) developed a framework to estimate traffic counts, linking administrative, survey and infrastructure data. This framework allows for long-term predictions on a nationwide scale. However, it relies on assumptions about the population and traffic estimates cannot be confirmed with the data.

**METHODS & DATA** This paper shows how the framework can be extended by including traffic loop sensor data in the linkage process for calibration and validation. Sensor data provides observed vehicle counts for a sample of road segments on the road network. It therefore reveals the accuracy of expected counts on these segments.

In the next step, geographical and road network information is used to predict the accuracy on unobserved road segments. Multiple calibration models are developed and compared. Finally, estimations are calibrated for the entire Dutch road network.



**RESULTS** Our results show that the extended framework significantly improves the accuracy of traffic estimation on a nationwide scale. Specifically, estimations produced by the calibration models significantly outperform the current method and produce valid estimations for official statistics and policy-making.

**ADDED VALUE** The paper demonstrates how data linkage can be performed to produce valid nationwide traffic estimation for official statistics and policy-making. The proposed framework can be used, e.g. to estimate the local effects of planned housing developments or large factories on traffic demand of specific roads. Suggestions for future research are given.

## POSTERS

### ELECTRONIC WORD-OF-MOUTH MARKETING – HOW INFLUENCERS, CUSTOMER REVIEWS AND SPONSORSHIPS AFFECT CONSUMERS' CREDIBILITY ASSESSMENTS AND PURCHASE INTENTIONS

**BOSAU, CHRISTIAN; SCHEPPE, JOANA**

Rheinische Fachhochschule Köln gGmbH, Germany

**RELEVANCE & RESEARCH QUESTION** Since the establishment of social media, word-of-mouth marketing has taken on a new dimension, and more recommendations are being shared on social networks. In addition to classic customer reviews, influencer marketing generates a new area in the electronic word of mouth. Companies are taking advantage of this new opportunity to promote their brands and provide them for free with the aim of product recommendations. Therefore, sponsored recommendations are particularly important in influencer marketing, yet in classic customer reviews too, where recommendations are sometimes based on a sponsored product.

**METHODS & DATA** This paper investigates the question how the identity of the author of a product recommendation (influencer vs. customer) as well as a possible sponsorship of a recommendation has an impact on the credibility and purchase intention of consumers.

With the help of an 2x2 experiment, the product recommendation of an average customer on Amazon.de is compared with a product recommendation of an influencer on Instagram. It also examines the influence of a sponsored recommendation compared to a non-sponsored and intrinsically motivated recommendation as well as the interaction effect of both factors. Based on an online survey, the data of 179 probands were analyzed using a MANOVA.

**RESULTS** The results of the study show no significant difference between the product recommendation of an average customer and the product recommendation of an influencer regarding credibility and purchase intention. However, the study shows that sponsored recommendations are judged to be less credible than un-sponsored recommendations ( $F(1,175) = 130.062, p = .000, \eta^2 = .43$ , leading to less purchase intention  $F(1,175) = 85.567, p = .000, \eta^2 = .33$ . Though, the sponsorship of a recommendation does not moderate the effect of the respective product recommendation.

**ADDED VALUE** The paper provides a contribution to the research of digital recommendations in social media and their effect on consumers, as well as valuable advice for companies that would like to promote their products through word-of-mouth marketing. Overall, the study shows that both customer reviews and product recommendations from influencers are attractive marketing tools for companies, however only in the un-sponsored way.

### FROM TEXT TO VOICE: INNOVATING THE FINAL COMMENT QUESTION IN A WEB SURVEY

**CLAASSEN, JOSHUA**

DZHW, Germany

**RELEVANCE & RESEARCH QUESTION** Self-administered web surveys mostly include closed questions with pre-defined answer lists. However, survey designers frequently implement open questions at the end of web surveys that allow respondents to express themselves freely and beyond previous questions. Such "final comment questions" (FCQs) are intended to collect respondents' elaborations of the web survey including its questions and shed light on general reactions, topic coverage, potential critique, methodological problems, and technological issues. One problem associated with FCQs is that they frequently result in short or no answer at all because typing in answers is burdensome. This especially applies to web surveys completed on smartphones with a virtual on-screen keypad. Building on technological advancements and an increase of mobile devices in the form of smartphones, I address the following research question: Do text and voice answers to a FCQ differ in terms of item non-response, word length, and response quality?

**METHODS & DATA** I analyze data from a smartphone survey experiment conducted in the German Forsa Omninet Panel ( $N = 1,001$ ) in November 2021. Respondents were randomly assigned to final comment questions with requests for text or voice answers. Voice answers were recorded on iOS and Android smartphones using the open-source tool SurveyVoice (SVoice). The open answers were manually coded to gain further insights on the quality of respondents' final comments.

**RESULTS** My results indicate that item non-response is high (about 50% in the text and voice conditions) but does not differ between text and voice answers. Instead, item non-response is significantly related to respondents' age and interest in the topic of the survey. However, voice answers are on average about three times longer than their text coun-

terparts and contain more topics. In addition, the request for voice answers results in a higher share of substantive answers than the request for text answers.

**ADDED VALUE** This study demonstrates that FCQs with requests for voice answers have the great potential to extend the existing methodological toolkit. In this way, it contributes to the development of respondent-centered web surveys that potentially reduce respondent burden and increase the quality of answers.

## SPARK THE FIRE: THE EFFECT OF CHANGES IN TRUST AND WELL-BEING ON POLITICAL PARTICIPATION AMID THE COVID-19 PANDEMIC

CONSOLINI, MICHELE; MASCHERINI, MASSIMILIANO  
Eurofound, Ireland

### RELEVANCE & RESEARCH QUESTION

Well-being; vaccine; political participation  
COVID-19 had pervasive and uneven effects on people's well-being, as erosion in mental and financial health. Moreover, anti-vaccination campaigns proliferated online, calling for distrust and political activism [Mascherini & Nivakoski, 2022]. Previous literature shows a causal relationship between an erosion in well-being and protest participation, whereas improved well-being stimulates voting and institutional participation [Flavin & Keane, 2012].

Thus, our research question is: "to what extent does a change in well-being and trust impact institutional and protest participation and voting intentions during the pandemic?"

### METHODS & DATA

e-survey series; online panel  
Weighted logistic regressions were carried out. The dynamic models availed of change variables across rounds among panellists. We predicted political participation and intention to vote by change in financial and mental well-being, change in trust in institutions, preferred media outlet, vaccination rate, alongside demographics.

The "Living, Working and Covid-19" online survey series was used. Five survey rounds have been completed across all EU Member States, with the latest carried out in spring 2022. Panellists who completed the survey in spring 2021 and spring 2022 were considered (12,419 respondents).

### RESULTS

trust and discontent; news source; social media  
Wealthier and vaccinated respondents who grew more dissatisfied with democracy were more likely to engage in institutional participation, such as meetings of political parties or unions. We found similar results for intention to vote. Conversely, erosions in wealth and health, along-

side not being vaccinated, signalled a higher likelihood of participating in protests. Thus, the erosion of well-being, more than trust in institution, figured as a determining factor for the choice of political participation citizens engage in.

### ADDED VALUE

Panel; causality; pan-European  
The panel structure allowed us to capture opinions on trust in institutions and changing well-being amid a volatile period, thus showing key developments in political attitudes. Moreover, the paper reinforced the causal links between well-being and political participation at the European level. Different citizens' needs imply different ways of participating in the democratic process. This incentivates both practitioners and policymakers to promote different political activities and perhaps helps in preventing extreme or violent protests.

## THE MORE THE MERRIER? GROUP SIZE AND THE STATUS OF ONLINE GROUPS ON REDDIT

DILLHÖFER, JAN; KLEINIEWIESE, JULIA  
MZES, University of Mannheim, Germany

**RELEVANCE & RESEARCH QUESTION** The impact of group size on a group's subjective social status is an emerging topic in social psychology research. This mechanism is significant given the projected increase in diversity within Western societies, resulting in a potential decrease in the current majority population. To accurately predict future societal dynamics, it is crucial to examine environments that reflect similar characteristics to human interactions with high levels of diversity. With the internet offering a diverse range of content created by individuals from various sociodemographic and ethnic backgrounds but also with various interests, social psychology theories are being utilized to explore the formation of social status in online communities. Hence, this study aims to investigate the impact of group size on the social status of online communities.

**METHODS & DATA** We argue that subreddits are integral to Reddit users' experiences, creating attachments similar to offline groups. Thus, we conducted a survey on Reddit using the Macarthur Scale of Subjective Social Status to measure social status of subreddits. Our study combines survey responses and digital trace data, including individual behaviour and subreddit indicators. The approximation of group size is based on the unique comments made in a subreddit in one month. Our research design theoretically and methodologically illustrates the intersection of internet group membership and individual behaviour, which can help to shed light on the communication on social media sites, like Reddit.

**RESULTS** Our findings indicate a reliable link between group size and community social status, as shown by a statistically significant positive relationship observed through a generalized additive model. These results challenge the Optimal Distinctiveness Theory, which predicts that neither too small nor too large groups promote positive group membership.

**ADDED VALUE** Our research demonstrates that social comparisons on Reddit are based on group characteristics rather than individual features. We define groups as a set of internet users with common interests. Eventually, our findings contribute to ongoing discussions about filter bubbles and echo chambers by examining how exposure to information (e.g. the attention of posts) leads to how different groups and subsequently individuals belonging to these groups are perceived and feel valued.

## MEASURING ATTITUDES TOWARDS SURVEYS IN MIXED-DEVICE SURVEYS: DOES IT MATTER HOW WE PRESENT THE SCALE?

**EULER, THORSTEN (1); SCHWABE, ULRIKE (1); FIEDLER, ISABELLE (1); JUNGEMANN, NIKLAS (2)**

1: DZHW, Germany

2: University of Kassel, Germany

**RELEVANCE & RESEARCH QUESTION** In general, measurement error negatively affects data quality (Grooves et al., 2009). As online surveys are usually mixed-device surveys (Lugt & Toepoel, 2015), questions are differently presented for mobile and non-mobile respondents. Mobile devices provide less display space. We investigate whether different modes of presentation, namely matrix versus carousel, affect the measurement of attitudes towards surveys. We choose a scale on attitudes towards surveys, as investigating survey climate and non-response trends has received special importance recently (Silber et al., 2023).

**METHODS & DATA** We realised a survey experiment with split-half design in a survey for higher education graduates in 2019 (gross sample  $n=3,345$ ). About 45 percent of respondents participated via a mobile device. As an instance of complex questionnaire designs, we use a nine-item battery (endpoint labelled 7-point-Likert scale) of questions addressing attitudes towards surveys (SAS, de Leuw et al., 2021). We differentiate three groups: For non-mobiles, our control group ( $n=861$ ), all items are displayed as a single matrix question using the full screen. Mobile users are randomly split in two groups, our experimental groups: One group ( $n=251$ ) received the nine items as a matrix question as well, the other group ( $n=258$ ) got the item grids split into single items (so-called vertical scaled carousel) with automatic forwarding.

**RESULTS** To investigate whether data quality depends on the mode of representation, we first examine descriptives for distributions, response times, straightlining and item nonresponse for all three groups. In a second step, we check for measurement equivalence between the control and the two experimental groups. Our results indicate that there are significant differences in response behaviour on some items depending on the respective presentation of items. Mobile respondents with matrix questions are more likely to refuse the battery completely and show a higher variance when giving valid answers. As such, the carousel presents itself as more applicable alternative for mobile devices, even though it takes more time.

**ADDED VALUE** Overall, our results have direct implications for designing mixed-device surveys: We recommend a responsive design implementing the vertical scaled carousel. However, we reflect on the generalizability of our findings.

## SANREMO FESTIVAL 2023, THE ITALIAN MEDIA “LOVE-BRAND”

**FACHIN, PAMELA**

Rai Pubblicità, Italy

**RELEVANCE & RESEARCH QUESTION** Sanremo Festival is the most important media event of the year, able to monopolise the audience attention cross generations, including GenZ. Festival 2023 was more multi-platform than ever.

Thanks to the project “From the stage to the city”, Sanremo was animated by sponsored events on field ensuring additional opportunities to see for brands.

It was strategic for Rai Pubblicità to implement an articulated set of surveys to monitor the positive impact of brand integration activities exclusively created for partners.

**METHODS & DATA** The set of surveys was articulated in levels:

- Pre&Post event (editorial overview + brand partner activities focus with CAWI/CAMI methodology, 6000 interviews in total, nat. rep.)
- Ongoing on field directly on Sanremo area thanks to tablet with pre implemented online survey mixed with ethnographic observation (CAWI + Personal interview digitalized + Ethnographic observation, 5 days, 1000 interviews, gender and age balanced)

In order to guarantee the solidity and quality of the data collected, the total sample in each phase was divided in dedicated homogeneous legs (Pre-Post: A & B); On field (a matrix designed around the 5 main locations hosted the activities of the brand sponsors put under control the interviews). In addition to this, it was implemented a digital and social focus with customized project set on a Consumer Intelligence platform able to analyse mentions, interactions, engagement, sentiment, emoji, key words on the event and on the brands involved.

**RESULTS** The analysis of the results showed that Sanremo can generate high interest and attention curve on the editorial and commercials contents: before the event start, during and post event.

Sanremo is “The Place to Be”: a live event with strong appeal, able to give visibility to social issues (sustainability, inclusion, gender gap, etc.), to involve all generations and to provide enjoyable moments to share with others (physically and digitally). The research set prove that Sanremo is the optimal context for brand and commercial initiatives.

**ADDED VALUE** The survey showed that in recent years Sanremo is a true love-brand media: the media communication and brand integration activities were cutting-edge, innovative and engaging.

## DESTRUCTIVE EFFECTS OF DISINFORMATION ON ELECTION PROCESS IN EU AND UKRAINE. HOW WE CAN FIND THE TRUTH?

GOROKHOVA, TETIANA

Centre for Advanced Internet Studies, Germany

**RELEVANCE & RESEARCH QUESTION** Today the world community faces a new challenge - the threat to the information security of individual countries and their citizens, who, under the conditions of information illiteracy and incompetence, can freely succumb to hostile information influence and threaten state sovereignty and territorial integrity. The aim of the research is to identify key areas for improving the legal and organizational model of protection of the electoral process from the destructive effects of disinformation, as well as to create a basis for broad dialogue between the state and the professional community in finding solutions for effective information protection. The study asked the following questions: 1. What is the current experience of the EU and Ukraine in counteracting destructive, disinformation actions during election campaigns? 2. To what extent are the respective efforts of state bodies correlated with the measures of civil society, media, and large Internet resources?

**METHODS & DATA** The study used methods of quantitative and qualitative content analysis. The similarity in the topics covered during the chosen period, the way of coverage and accents, the use of emotions, evaluative judgments, and other manipulative technologies, etc. on news sites, and government sites that are most popular in Europe and Ukraine. With the help of this analysis, it is possible to determine the rhetoric that is picked up on these platforms, and how it influences public opinion in the electoral process.

**RESULTS** In the context of the research study was analysed the level of interference in electoral processes in EU countries and Ukraine using digital technologies and manipulative tools over the past 10 years. It also analyses the level of participation of civil society organisations in the fight against manipulation and digital propaganda in comparison with state institutions in the 2019 elections in Ukraine.

**ADDED VALUE** Studying and analysing the possibilities of information technology in the constantly evolving electoral processes will help to better understand and analyse the events taking place in Ukraine, the European Union countries and the world and is an important condition for developing modern adequate information security strategies.

## ARTIFICIAL INTELLIGENCE, UNBIASED OPINIONS? ASSESSING GPT'S SUITABILITY FOR ESTIMATING PUBLIC OPINION IN MULTI-PARTY SYSTEMS

VON DER HEYDE, LEAH ROSA (1); WENZ, ALEXANDER (2);  
HAENSCH, ANNA-CAROLINA (1)

1: LMU Munich, Germany

2: University of Mannheim, Germany

**RELEVANCE & RESEARCH QUESTION** The recent development and release of widely accessible large language models (LLMs) appear to have the potential for a much more efficient collection of public opinion data: Argyle et al. (2023) prompted the LLM GPT-3 to elicit responses to survey questions from the American National Election Study (ANES) and found that the generated responses, on aggregate, closely matched the survey data. However, several factors might affect the generalizability of these findings beyond the US, including (1) each society's particular relationships between party systems, political attitudes and behavior, and socio-demographic factors and (2) the amount of country-native language training data available for developing LLMs. Therefore, we ask: Can LLMs be used for estimating public opinion outside of the US?

**METHODS & DATA** We apply the methodology employed by Argyle et al. (2023) to the German context: We prompt GPT-3 with personas matching the individual demographic characteristics of the 2017 post-election cross-sections of the German Longitudinal Election Study (GLES). From the resulting vote choice probabilities, we compare the average party vote shares to those indicated by the GLES survey data. We also assess whether GPT-3 is able to make accurate estimates of vote choice for different population subgroups and compare the impact of predictor variables on the GPT response generation and the GLES vote.

**RESULTS** Based on the prompt design and model parameters we present, we find that GPT-3 is not as successful in replicating the GLES data as it was in replicating the ANES data. Preliminary results show that GPT-3, on average, does not correctly predict citizens' vote choice, exhibiting a bias towards left and green parties, and making better predictions for some subgroups than others.

**ADDED VALUE** By assessing the specific conditions under which LLMs can be leveraged for studying public opinion in different countries and political systems, this study contributes to the growing body of research about how LLMs can be used for social science research. We conclude that GPT cannot replace traditional surveys at this point.

## HOW DO DIFFERENT ADVERTISING STRATEGIES ON SOCIAL NETWORKING SITES AFFECT PERFORMANCE AND SAMPLE COMPOSITION

HEBEL, ANNA

GESIS – Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** Survey recruitment through social networking sites such as Facebook or Instagram is becoming more and more popular due to its powerful targeting options [focusing on demographics, interests, or behaviours]. Potential survey respondents see an advertisement for a survey, click on a link, and are taken to the survey's landing page outside of Meta (Facebook and Instagram). However, despite its growing popularity, little is known about how the advertising influences the recruitment process and subsequently the sample's composition. For example, which images appeal more to women or men, or old or young users? Which images lead to the most link clicks and completed surveys? The research question is how do different advertising strategies affect performance and sample composition?

**METHODS & DATA** The survey takes place from mid-May to mid-June 2023. It uses Meta's "charge per link click" option, which invites potential respondents to visit the survey's landing page. Since this study targets the overall set of Facebook and Instagram user, targeting criteria are very broad, that is, two age categories and both genders, which results in four ad sets. The design of the advertisements is based on four different survey motivations with eight different Image/Text combinations. Ad performance is measured by counting clicks and completed surveys. In addition, sample composition is examined for all advertisements to find out which groups of people feel invited by which advertising images.

**RESULTS** The study will be fielded in May/June 2023. Eight different advertising images are compared and analysed to determine which images perform best overall (most link clicks and completed surveys) and which are particularly appealing to specific groups of people. The study will be completed by mid-June 2023, and basic preliminary analyses will be performed by August.

**ADDED VALUE** This research helps to improve advertisements on meta and it shows what impact the ads have on respondent participation and associated selectivity.

## CAN A TARGETED REFUSAL CONVERSION STRATEGY INCREASE DATA QUALITY IN A WEB-SURVEY EXPERIMENT?

LAUPPER, ELLEN (1, 2); REIPS, ULF-DIETRICH (2)

1: Swiss Federal University for Vocational Education and Training, Switzerland

2: University of Konstanz, Germany

**RELEVANCE & RESEARCH QUESTION** Unlike face-to-face surveys, where interviewers can convert refusals by directly addressing expressed concerns about participating in the survey, web surveys must rely on indirect refusal aversion strategies such as personalization, reminders, and incentives. Given the lower response rates for web surveys and concerns about the potentially lower data quality provided by late respondents, we test a feasible strategy to motivate potential refusals by addressing their concerns in a targeted manner. Our study addresses the question of how and when late respondents and nonrespondents to a business web survey can not only be persuaded to participate, but also be motivated to provide high data quality. To this end, an automated refusal conversion strategy originally proposed by Lewis et al. (2019) is adapted, extended, and experimentally tested.

**METHODS & DATA** A 3 (time of intervention during field period) x 3 (type of appeal: targeted, general, main reason) factorial experimental survey design was chosen. The 3000 participants will randomly be assigned to the experimental conditions, with one third each at the three different times during the survey period (i.e., instead of the first and second reminder and after the end of the fielding period). They will receive a special reminder letter with a link to a web survey in which they are asked to indicate their main reason for not participating so far. After having completed this „one-question web survey“, they will randomly receive either a „targeted appeal“ that addresses the reason stated for non-participation, a „general appeal“ or a „main reason appeal.“ To complete the survey, they must then indicate whether they want to proceed immediately to the main survey or postpone the decision.

**RESULTS** Because data collection will not be completed until mid-September, preliminary results will be available for the first and second reminder. For these, initial comparisons of response rate and data quality indicators between the three experimental groups will be reported.

**ADDED VALUE** This study experimentally tests a viable strategy for converting refusals into responses for web surveys and provides valuable insights into the relationships among reasons for nonresponse, timing of interventions, and data quality.

## CAN BRANDS SAVE THE WORLD? INVESTIGATING THE INFLUENCE OF SLOGANS ON CONSUMER BEHAVIOR: A CASE STUDY OF HABIT CHANGING BRANDS

PREUSS, FREYA; DUST, LISA

Facts and Stories GmbH; Germany

**RELEVANCE & RESEARCH QUESTION** Consumer habits play a crucial role in shaping behavior, yet the study of habit formation and its relationship with brands remains relatively unexplored. This research focuses on „habit changing brands,“ using the example of a liquid hand soap from a natural cosmetics brand that tries to encourage its users to save water. The aim is to understand how brands can contribute to

changing consumer behavior regarding the purchase, use, and disposal of products, particularly in the context of promoting sustainability. By employing targeted communication strategies such as slogans, consumers can be informed about their contribution to a larger goal. This study seeks to investigate the conditions under which habit formation occurs or fails to occur.

**METHODS & DATA** The study assesses the brand slogan's impact on a test group's water conservation during handwashing, considering factors like faucet type and ecological attitudes. The research initially involved 87 participants, almost evenly divided into those using the branded (test group) and unbranded product (control group). They completed online surveys over about 10 weeks. Statistical analysis, including dependent t-tests and Pearson product-moment correlations, will be employed to explore mean differences and relationships within and between the two groups.

**RESULTS** This work is still in progress, and the results will be submitted at a later date.

**ADDED VALUE** This study offers several contributions to the field of online consumer research. First, it consolidates research findings on the effectiveness of slogans and brand messaging. Second, it uncovers challenges associated with behavior change factors. Third, it addresses the research questions using study results, shedding light on the theories of habit acquisition and the duration of habituation. By examining the specific case of habit-changing brands and their impact on water-saving behaviors, this research provides valuable insights into the potential of brands to influence sustainable consumer habits.

## MEASURING ATTITUDES TOWARDS SURVEYS: A VALIDATION STUDY

**SCHWABE, ULRIKE; EULER, THORSTEN; FIEDLER, ISABELLE; JUNGERMANN, NIKLAS**  
DZHW, Germany

**RELEVANCE & RESEARCH QUESTION** Besides others, general attitudes towards surveys are part of respondent's motivation for survey participation. These attitudes do predict participant's willingness to perform supportively during (online) surveys (de Leeuw et al., 2017; Jungermann et al., 2019). Therefore, different attempts have been made to measure these general attitudes adequately. Most recently, the Survey Attitude Scale (SAS) as proposed by de Leeuw and colleagues (2019) comprises three dimensions: survey enjoyment, survey value, and survey burden. However, other research proposes additional dimensions such as survey reliability, survey privacy and survey intentions (Gengler et al., 2019; Looseveldt and Storms, 2008). Aiming to validate the SAS, we compare convergent and discriminant validity of different dimensions of attitudes towards surveys.

**METHODS & DATA** We included different items from the literature in a survey for higher education graduates in 2019 (gross sample n=3,345). Being randomly assigned, some respondents answered the items pro-

posed by de Leeuw and colleagues (2019) (SAS). The others answered items that capture the dimensions proposed by Looseveldt and Storms (2008) and Gengler et al. (2019). We apply confirmatory factor analyses (CFA) to replicate the original scales and compare their fit indices and Cronbach's Alpha for reliability. We then examine convergent and discriminant validity by correlating different dimensions. Finally, we run exploratory factor analyses (EFA) on the entire item set to assess potential for a more convincing scale.

**RESULTS** First, CFA results indicate replicability of the proposed scales with our sample. However, the SAS model performs less effectively than the alternate model, yet the SAS dimensions demonstrate higher reliability scores. Secondly, as expected, we find high correlations between those items which intend to measure similar dimensions such as survey value and survey reliability or survey burden and survey privacy. Finally, EFA results do not support alternative dimensions in our data superior to existing scales.

**ADDED VALUE** The recommended scales are promising to measure attitudes towards surveys. Despite some weaknesses, the SAS instrument is valid and efficiently. As our validation study is based on a sample of highly qualified graduates, generalizing our results to the entire population should be taken with caution.

## AI-ENABLED DATA ANALYSIS QUALITY: ADDRESSING A KNOWLEDGE GAP

**WETZELHÜTTER, DANIELA (1); PRANDNER, DIMITRI (2)**

1: University of Applied Sciences Upper Austria, Austria  
2: Johannes Kepler University Linz, Austria

**RELEVANCE & RESEARCH QUESTION** During quantitative data analysis various errors can occur, e.g. use of inappropriate analysis methods, mishandling missing data or overfitting overly complex model. It can be assumed that, apart from deliberate deception, a lack of knowledge in particular increases the risk of erroneous results. And this is exacerbated by new AI-supported possibilities for data analysis. This poses a significant challenge, as one may now be enabled to conduct simple as well as complex data analysis, with only limited knowledge. Consequently, our research question is: What are the pitfalls of using AI-based data analysis?

**METHODS & DATA** Our approach illustrates the potential for error in AI-based data analysis: descriptive statistics, factor and regression analysis. We use published replication datasets to do so and compare the published results with AI-based ones. We use AI-tools to generate the required syntax for SPSS, Stata and R, before executing the code and providing the AI with results to interpret. We use replication data available at AUSSDA (<https://data.aussda.at/dataverse/AUSSDA/?q=replication>), which provides us with both data and a correct syntax.

**RESULTS** While tools like Chat GPT can solve simple calculations and equations correctly, this is not the case when it comes to data analy-



sis. (1) Descriptive data analysis tends to include mistakes, reporting such results may have severe consequences. (2) However, the use of a syntax generated by Chat-GPT (e.g., for conducting exploratory factor analysis) seems more promising - but it requires expertise (to adapt e.g. Missing value handling, criteria, extraction, rotation). Though, the information needed can be requested to overcome knowledge gaps. However, the more complex the procedure being considered for application, the higher the risk that the syntax will not work.

**ADDED VALUE** Because AI-supported programming for data analysis offers valuable opportunities, we illustrate that such option need to be treated with utmost care. However, our results are limited due to the continuous, rapid development of AI-tools for data analysis.

Excerpt Literatur: Burger, B., Kanbach, D. K., Kraus, S., Breier, M., & Corvello, V. [2023]. On the use of AI-based tools like ChatGPT to support management research. *European Journal of Innovation Management*, 26(7), 233-241.

## DIGITAL NUDGING AGAINST HATE SPEECH ON INSTAGRAM

WÖLK, MICHAEL

FH Wiener Neustadt, Austria

**RELEVANCE & RESEARCH QUESTION** The individual freedom to communicate one's own expressions of opinion widely on Instagram can lead to abusive use in the form of so-called hate speech. This can take the form of personal insults, threats or defamation and be directed against individuals or entire social groups. (cf. Zhang & Luo, 2019) Hate speech on the Internet is most strongly perceived by 14 to 24-year-olds (cf. Landesanstalt für Medien NRW, 2018). Those affected by Hate Speech can thereby struggle with psychological consequences (cf. Geschke, Klaffen, Quent & Richter, 2019) and partially or completely withdraw from social groups (cf. Schmier et al., 2020). This leads to the research question: „How does the effectiveness of the two nudging types Social Norms and Warnings in the digital setting against Hate Speech on Instagram differ among 16- to 24-year-old Austrians with regard to the perceived feeling of safety?“

**METHODS & DATA** To answer the research question, an experimental design in the form of a quantitative online survey was conducted. The survey was designed as within subject design - the questionnaire contains two graphically presented stimuli, which are presented in randomized order in the survey tool Qualtrics. A dependent sample t-test was used as the analysis method. A sample of  $n = 235$  was recruited from the calculated population of 892,170 Austrian Instagram users using a convenience sampling method.

**RESULTS** The results of the study show that the sample rated the Warning Nudge as more effective than the Social Norm Nudge. The estimation of effectiveness was measured by four dependent variables, with three of these showing a significant result at the t-test. The dependent

variables were Estimation of Nudge Effectiveness ( $t(235) = 3.543$ ,  $p = <0.001$ ), Estimation of Own Sense of Safety ( $t(235) = 3.368$ ,  $p = <0.001$ ), Estimation of Fear of Posting ( $t(235) = -1.204$ ,  $p = 0.230$ ), and Assumed Limitation of Instagram Use ( $t(235) = 2.792$ ,  $p = 0.006$ ).

**ADDED VALUE** The findings should contribute to the preservation of democratic debate culture and diversity of opinion in order to counteract negative trends such as silencing (intimidation tactics to limit public discussions).

## A3.1: RECRUITMENT PROCESSES FOR ONLINE SURVEYS

(SPONSORED BY GESIS)

SESSION CHAIR: JOSHUA CLAASSEN

DZHW, Germany

### THE IMPACT OF AD IMAGES ON SURVEY RECRUITMENT AND SAMPLE COMPOSITION IN FACEBOOK SURVEYS

ZINDEL, ZAZA (1); KÜHNE, SIMON (1); ZAGHENI, EMILIO (2); PERROTTA, DANIELA (2)

1: Bielefeld University, Germany

2: Max-Planck-Institute for Demographic Research, Germany

**RELEVANCE & RESEARCH QUESTION** This study examines the impact of ad images on the composition of resulting samples in social media recruited surveys. While Facebook is a cost-effective tool for recruiting survey participants, studies have noted the high selectivity of resulting samples. Therefore, it is crucial to investigate who self-selects into a study advertised on Facebook. The study hypothesizes that strongly themed, topic-related advertisements are less expensive and result in a larger number of completed interviews than neutral ads. However, thematic ads lead to a more politically polarized sample than neutral ads.

**METHODS & DATA** We experimentally manipulated Facebook ads' images and assessed their effects on the recruitment process and composition of survey respondents. The study was conducted in Germany and the U.S., using 15 different Facebook ad images to recruit participants for a survey on climate change and immigration. Ads with neutral images were used as the baseline condition, while ads indicating the survey



topic were used as the treatment. Further, an online access panel was used as a baseline comparison.

**RESULTS** Thematic ads related to the survey topic are less expensive and result in a larger number of completed interviews than neutral ads. The findings furthermore suggest that using images related to the survey topic significantly affects the sample's composition. Specifically, thematic ads resulted in a more politically polarized sample compared to neutral ads. Further results will be available at the conference.

**ADDED VALUE** Overall, this study highlights the importance of carefully choosing Facebook ad images to maximize recruitment success while minimizing sample bias. It provides valuable insights into the impact of Facebook ad images on the recruitment process and sample composition in social media recruited surveys.

## ATTRITION AND TYPE OF PANEL RECRUITMENT: COMPARING PANEL ATTRITION BETWEEN A PROBABILITY AND NON-PROBABILITY RECRUITED SAMPLE OF SWEDISH CITIZENS

**LINDGREN, ELINA; CASSEL, FELIX**  
University of Gothenburg, Sweden

**RELEVANCE & RESEARCH QUESTION** Random sampling procedures can help increase representativeness in cross-sectional surveys, but the extent to which representativeness is maintained in panel surveys is less known. Previous research indicates that panel attrition is related to demographic characteristics. If attrition is similarly demographically biased for probability and non-probability samples, the higher representativeness in the probability-based samples should sustain. If, however, randomly recruited panelists have a greater tendency to drop out over time than non-probability members, differences in representativeness between the sample types could even out over time. Given that probability samples often cost more resources there is reason to investigate how much gain such sampling procedures provide compared to non-probability sampling for longitudinal surveys.

We address this question by asking:

How representative is a probability and non-probability sample on demographic factors in the first panel wave? How do demographic representativeness in each sample type evolve over time? How do response patterns on outcome variables differ between the sample types, and how do the differences evolve over time?

**METHODS & DATA** Data comes from four waves of the Knowledge Resistance Panel (n=10,948) collected annually since 2020. The primary intention of the panel is to study knowledge on issues that enjoy academic and expert consensus. Attrition is studied by between-wave panel termination as well as partials, breakoffs and item-missing within waves between the samples. How attrition varies between sociodemographic factors, knowledge and media consumption items is also studied.

**RESULTS** We find that panel termination was larger for the probability sample between the first two waves, but these differences shrank in the latter waves. While differences in demographic composition between the two samples was significant in the first wave, they also partly evened out over time. Differences in item missing, partials and breakoffs were statistically insignificant between the two samples in all waves.

**ADDED VALUE** We demonstrate the need to take differences in panel recruitment into account when studying panel attrition and generalizing findings from panel surveys. Full results from the fourth and final panel wave, conducted in spring 2023, will be presented at the conference.

## COMPARING SOCIAL MEDIA SAMPLES IN AFRICA AGAINST THE DEMOGRAPHIC HEALTH SURVEY

**ROHR, BJÖRN (1); SILBER, HENNING (1); FELDERER, BARBARA (1); WEISS, BERND (1); PÖTZSCHKE, STEFFEN (1); PRIEBE, JAN (2)**

1: GESIS – Leibniz Institute for the Social Sciences, Germany  
2: Bernhardt Nocht Institute for Tropical Medicine, Germany

**RELEVANCE & RESEARCH QUESTION** Sampling respondents via ads on social media (e.g., Meta advertisement) is a promising new approach to reach respondents both in hard-to-sample regions and at a cross-national level. To be successful, Meta optimizes their advertisement to reach the most clicks on a provided link (e.g., the survey link). The optimizing algorithm includes the layout of the ad and user characteristics. When setting up an advertisement campaign, the selection of various targeting strategies is possible. One option is targeting all Meta users of a selected target group with a single ad-set or using multiple ad-sets for demographic subgroups to ensure demographic variation. Yet, the impact of such variations of the targeting strategy on the sample composition is not well understood. Behind this background our research question is [a] whether social media sampling can result in samples that are similar to population benchmarks and [b] whether the using multiple ad-sets for demographic subgroups improves the sample accuracy.

**METHODS & DATA** In this study, we evaluate the sample composition of surveys we conducted using Meta ads in Sub-Saharan Africa. Our target sample consisted of Facebook users from three African countries (Ghana n = 527, Kenya n = 2,843, and South Africa n = 313). We quasi-experimentally varied two targeting strategies over the course of three weeks. First, our demographic targeting strategy grouped the countries by region, age, and gender, whereas the second sampling strategy did not implement demographic targeting. As benchmarks, we use socio-demographics and health-related variables such as HIV knowledge from the Demographics and Health Survey (DHS).

**RESULTS** Results showed that the social media sample was very different from the DHS benchmarks, especially regarding demographic information, such as gender and education. We further found that different targeting strategies did not result in significant variation with respect to the accuracy compared to the benchmarks in any country, neither for univariate nor bivariate comparisons.

**ADDED VALUE** The study results inform researchers on how a Meta survey of African countries compares to probability surveys and provides insights into the advantages and disadvantages of employing demographic targeting as a sampling strategy for social media recruitment.

## A3.2: MIXED MODE AND MODE TRANSITION I

[SPONSORED BY GESIS]

**SESSION CHAIR: GRACE CHANG**

University of Southampton, United Kingdom

### DOES INCLUDING INTERNET USERS IN THE MAIL MODE IMPROVE THE DATA QUALITY OF A PROBABILITY-BASED MIXED-MODE PANEL?

**BRETSCHI, DAVID {1}; DAIKELER, JESSICA {2}; NEUMANN, ANNA {2}**

1: BKA, Germany

2: GESIS - Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** Previous research on probability-based online panels shows that providing the offline-population with an alternative survey mode reduces coverage bias. However, little is known about whether offering an offline mode to internet users who are unwilling to participate in panel surveys online pays off in lower nonresponse bias.

**METHODS & DATA** This study uses data from the GESIS Panel, a probability-based web and mail panel, to investigate how including internet users in the mail mode affects nonresponse bias of population and model estimates.

**RESULTS** The results show that internet users included in the mail mode differ from non-internet users assigned to the mail mode and panelists responding in the web mode in demographic variables and internet usage after the recruitment. However, excluding internet users in the mail mode from the data set rarely introduces a higher bias in estimates of demographic characteristics compared to data from a reference sample. Further analyses examine how mean and model estimates of three studies published with GESIS Panel data would have been affected by excluding internet users from the sample who are unwilling to provide survey data online. Here the findings show potential nonresponse bias

in several estimates of means from the model variables. However, the model estimates of both studies are largely the same after removing those cases from the analyses.

**ADDED VALUE** Therefore, authors' conclusions would likely have remained unchanged if internet users had not been included in the mail mode.

### THE PUBLIC OPINION SURVEY OF THE CITY OF STUTTGART. METHODOLOGICAL DEVELOPMENTS AND THEIR IMPLEMENTATION

**SCHÖB, ANKE**

City of Stuttgart, Germany

**RELEVANCE & RESEARCH QUESTION** The Statistical Office of the state capital Stuttgart is the competence center for surveys in the city administration. In addition to answering questionnaires on paper, respondents have been able to take part in the surveys online since 2008. For the Stuttgart Citizens' Survey, which has been conducted every two years since 1995, this means an expansion of the mail option including a web option (mixed-mode design). In order to respond to technological change, the design of the survey has changed repeatedly in the process. This process continues to this day: While in the early years of the use of the online option, the distribution of the Internet (2008: 65%), the access to the Internet (modem, ISDN, broadband Internet) and the computing power were in the foreground, today the question arises to the used device to answer the questionnaire online. More recent studies (Millar/Dillman 2011; Lesser et al. 2016; Mauz et al. 2018) no longer necessarily indicate an advantage (higher response rates) of the mail option over a sequential push-to-web option, so that this year's Stuttgart Citizens' Survey will use this approach for the first time. The research question is, what are the consequences of switching to a push-to-web approach? Can we expect a higher response rate? Does the demographic structure of the sample change compared to the previous year's survey? Do online respondents differ from respondents who completed the questionnaire by mail?

**METHODS & DATA** The data used are the surveys from 1995 to 2023. For the comparison of the demographic structure data of 2021 and 2023 are used. To compare the onliners and offliners, the data from 2023 is used.

**RESULTS** The first results of the population survey 2023 will be available in August 2023.

**ADDED VALUE** The study contributes to the question of whether a shift to a push-to-web approach is associated with advantages or whether disadvantages can be identified.

# B3: SMARTPHONES, SENSORS, AND GEODATA

SESSION CHAIR: **JOSS ROSSMANN**

GESIS – Leibniz Institute for the Social Sciences,  
Germany

## HOW LONG CAN WE ASK RESPONDENTS TO PARTICIPATE IN AN APP-BASED TRAVEL DIARY STUDY?

**REMMERSWAAL, DANIELLE (1,2); SCHOUTEN, BARRY (1, 2);  
LUGTIG, PETER (1); STRUMINSKAYA, BELLA (1)**

1: Utrecht University, The Netherlands

2: Statistics Netherlands, The Netherlands

**RELEVANCE & RESEARCH QUESTION** Smartphone apps as a survey mode offers promising features to collect both self-report measures and digital behavioural data via smartphone sensors. The use of passive measurements on smartphones can be beneficial for response rates as it can reduce the burden for participants by replacing traditional survey questions, and for data quality as it can reduce recall bias. For diary studies that usually have low response rates and high recall bias, smartphone sensor data can be especially beneficial. In our study we test in an experimental setting whether the length of the diary survey and the extent of participant involvement are of influence on response burden in the app by analyzing nonresponse patterns and paradata.

**METHODS & DATA** In this study we report on two experiments carried out between November-2022 and February-2023 among 2544 individuals from a cross-sectional sample of the Dutch population. We asked individuals to use a smartphone app to collect their travel data over a 1-day or 7-days period (experiment 1). An alternative online questionnaire was offered. The app records travel behaviour using location sensors and compiles a diary which respondents can edit and enrich. Two versions of the app have been used: in the first app version the editing possibilities are extensive (time editing, adding stops/trips manually), in the second the edit possibilities are limited to deleting stops/trips (experiment 2). We assess whether registration and app-activity are related to individual characteristics with registry data.

**RESULTS** App registration rates are higher for the 7-day group (14.5%) than for the 1-day group (10.5%). Dropout is substantial for all participants, and is only partly explainable by technical issues. 1-day partici-

pants and limited-editing participants both spend less time in the app, and rated the app more user-friendly and less time-consuming.

**ADDED VALUE** With the use of app-collected paradata collected by the app we are able to assess more detailed how long people spend on the study and what they are doing. This enables us to improve our understanding of response burden in app studies. Additionally, our analysis can contribute to the design of future app-based diary studies.

## MAINTAIN THE FOREST – GUIDANCE OF VISITORS BY SMARTPHONE-BASED POSITION DATA

**KERN, FRANZISKA; DICK, JULIA**

infas 360 GmbH, Germany

**RELEVANCE & RESEARCH QUESTION** Especially during the Covid-pandemic more and more people sought relaxation and exercise outdoors in forests. But there are sensitive areas that must be protected from masses of visitors and different groups of interests using forests. How can visitors be guided through forests while protecting certain spaces and at the same time enable economic purposes like wood harvesting?

**METHODS & DATA** This research question has been addressed by using smartphone-based position data combined with small area methods and the database of infas 360. The input data were anonymised smartphone data from several apps (e. g. weather, navigation etc.) provided by a location intelligence company. To obtain a representation of visitor frequencies on the very fine-area hexagon structures that is as unbiased as possible, a small-area estimation was used in a next step. In this estimation model, frequencies per category (summer weekday/weekend, winter weekday/weekend) served as the dependent variable. The microgeographic and sociodemographic input data from infas 360 were used as independent variables.

**RESULTS** The coefficients generated from the analysis were used to estimate a frequency for each hexagon. These results were then converted to a scale of 1 to 100. An index over 30 was defined as being a highly frequented area in the forest. A few forestry offices were asked to evaluate the result. They concluded that the outcome was similar to their subjective experience and that the data at hand were very useful.

**ADDED VALUE** This approach shows an innovative way how to estimate visitors on a microgeographical level. In contrast to conventional methods, this approach is significantly cheaper and less complicated, but just as precise. Instead of many costly visitor counts at different time points, only the input data, microgeographic data and a statistical calculation procedure are now required to obtain valid estimates. This method can be applied to a wide variety of scenarios and can also provide great added value e.g. in urban or landscape planning or in the evaluation of infrastructure.

## LINKING REAL ESTATE ONLINE DATA WITH OFFICIAL STATISTICS TO ANALYZE DIFFERENCES IN RENT PRICE

SCHMIDT, MANUELA

University of Bonn, Germany

**RELEVANCE & RESEARCH QUESTION** ImmobilienScout24 is the largest internet platform for real estate in Germany, with a self-reported market share of about 50% of all real estate objects offered for sale or rent in Germany. Private and commercial real estate owners advertise their objects for a fee. In an online questionnaire, they report on their object, e.g., its address, offer price, size, and some optional information on various characteristics that determine the value of a property. The research question "which apartment characteristics lead to higher rent prices, and how do these depend on the region" can be addressed by complementing these variables with information on the regions the objects are in.

**METHODS & DATA** The data is provided by ImmobilienScout24 via the research data center (FDZ) Ruhr at the RWI and includes residential real estate. Each object is georeferenced and matched to administrative regions, municipalities (Gemeinden) and districts (Kreise), as defined by the Federal Statistical Office of Germany. For this paper, all apartments for rent between 2016 and 2021 in the federal state Nordrhein-Westfalen are included in the analyses, as georeferencing is most reliable from 2016, when the website enforced mandatory reporting on the objects' addresses. Using the information on administrative regions, it is possible to link this data with official data on the municipalities and districts, thus providing information on the population, the sociodemographic composition, and the economy of the region an object is located in. Through a multilevel analysis it can be observed how much of the variance on rent prices is attributable to the apartments themselves and how much depends on the region and its characteristics? And do these effects interact?

**RESULTS** Results show how the effects of apartment characteristics on the rent price change when region characteristics are taken into account.

**ADDED VALUE** In this presentation, I will discuss challenges with data linkage and present selected results to highlight how geo-referenced online data can be enriched using federal statistics.

## C3: EXTREMISM, HATE SPEECH AND POLITICAL ENGAGEMENT

(SPONSORED BY APROXIMA)

SESSION CHAIR: FELIX GAISBAUER

Weizenbaum-Institut e.V., Germany

### ONLINE VS. OFFLINE POLITICAL ENGAGEMENT IN IRREGULAR ELECTION CAMPAIGNS: THE ROLE OF POLITICAL TRUST AND DESIRE FOR CHANGE IN SHAPING ENGAGEMENT PATTERNS

ARIEL, YARON; ELISHAR-MALKA, VERED; WEIMANN-SAKS, DANA

Academic College of Emek Yezreel, Israel

**RELEVANCE & RESEARCH QUESTION** Between 2019 and 2021, Israel held four consecutive general elections. One of the most notable aspects of the fourth round was the unprecedented scale of grassroots oppositional activities, including mass rallies at symbolic locations. This study examined online and offline political engagement predictors, asking: under conditions of political instability, which type of political engagement will be more prevalent? Specifically, it focused on the potential impact of voters' trust in politicians and the political system and a desire for political change on engagement patterns.

**METHODS & DATA** A representative sample of 550 participants was surveyed. The sample was selected from an online panel through a quota sampling method, considering the demographic variables of age, gender, and education. Variables included trust in the political system, trust in politicians, desire for political change, and online/offline engagement.

**RESULTS** Results showed a statistically significant difference ( $t(df=504) = 14.45, p < .01$ ) between offline and online political engagement. In addition, there was a significant correlation between voters' trust in the political system and both offline ( $r = .31, p < .01$ ) and online ( $r = .25, p < .01$ ) political engagement. Similarly, a significant correlation was found between voters' trust in politicians, offline ( $r = .36, p < .01$ ), and online ( $r = .34, p < .01$ ) political engagement. No significant correlation was found between the desire for political change and offline political engagement ( $r = -.07, p > .05$ ). In contrast, a significant negative correlation was found between this variable and online political engagement ( $r = -.10, p < .05$ ). Multiple linear regression ( $R^2 = 12.4\%$ ,  $F(4, 500) = 16.81$ ,

$p < .001$ ) of offline political engagement indicated that the desire for political change, trust in politicians, and trust in the political system were significant predictors. However, regression for online political engagement ( $R^2 = 11.9\%$ ,  $F(4, 500) = 16.05$ ,  $p < .001$ ) found that only trust in politicians was a significant predictor.

**ADDED VALUE** Findings suggest that online and offline political engagement are motivated by different sources as two distinct types of political behavior. The desire for change is an influential predictor, especially offline political participation. Several factors may explain this phenomenon, including the relatively high motivation required for offline engagement. This kind of activity also requires voters' trust in the political "rules of the game," as our findings indicate.

## SENTIMENT AND HATE SPEECH ON SOCIAL MEDIA. AN ANALYSIS OF THE RELEVANCE OF GENDER, AGE AND MIGRATION BACKGROUND FOR THE EMERGENCE OF NEW COMMUNICATION STRUCTURES WITHIN THE YOUTUBE COMMUNITY

**BUDER, CLAUDIA; PHILIPP, AARON;  
VERWIEBE, ROLAND; WEISSMANN, SARAH**  
University of Potsdam, Germany

**RELEVANCE & RESEARCH QUESTION** Negative feedback, such as "hate" comments, and different forms of discrimination are daily occurrences for most YouTubers and other content creators (Park et al. 2021). The received negative feedback has an influence on the content creator's sense of self (Jones 2015), their interactions and activities on the platform. An increasing amount of international studies investigate this topic in recent years (Waqas et al. 2019; Salian & Ghosh 2022). They show that while the majority of content creators experience a variety of online hate or harassment, gender, age, and size of audience seem to be noteworthy at-risk factors for different forms of online harassment (Thomas et al., 2022).

In this field, the case of German-speaking countries remains a research gap especially for YouTube as a platform. Therefore, we analyze differences in frequency and content patterns of sentiment and hate speech, their change over time and along important socio-structural characteristics (gender, age, migration background) as well as the channel topic and channel size.

**METHODS & DATA** We use a random sample of  $N = 5,000$  YouTubers from a total survey of YouTube channels registered in Germany, Austria and Switzerland ( $N = 120,000$ ). We annotated relevant sociodemographic variables such as age, migration background, gender, and educational background via a structured classification survey (Seewann et al., 2022). We then apply different sentiment analysis algorithms (including ML) and hate-speech-detection on comments selected from the videos of the channels. In addition, we estimate differences in sentiment

and hate-speech in dependence of sociodemographic variables with regression analysis and use cluster analysis to study different types of hate-receiving-channels.

**RESULTS** First results show effects of gender and migration background on sentiment and hate-speech. Channels hosted by males receive less often negative sentiment than females and content creators with migration background are confronted with a higher amount of hate-speech. In addition, successful content creators receive more negative sentiment and hate speech is relatively common for those creators who engage in the area of knowledge, society, and politics.

**ADDED VALUE** We fill the lack of missing sociodemographic variables in social media data, especially YouTube.

## DID COVID-19 INCREASE RADICALIZATION POTENTIAL AMONG MISOGYNIST EXTREMISTS?

**COUFAL, LINDA (1); WEDEL, LION (2)**

1: Faculty of Social Sciences, Charles University, Czech Republic  
2: Freie Universität Berlin, Germany

**RELEVANCE & RESEARCH QUESTION** Incels (involuntary celibates) who base their identity on the inability to form romantic relationships have recently been recognized as an important contemporary phenomenon among extremist communities online. Covid-19 has been shown to increase traffic on Incel forums. From this, researchers derive an increased potential for radicalization. This research analyses those claims in-depth, posing the research question: did covid-19 increase the threat of radicalization among Incels?

**METHODS & DATA** We conduct our analysis on the publicly visible posts of the incels.is forum. Using piecewise regression, we detected four breakpoints that mark the moments in the time series when the trend changes significantly. Their dates mark the transition periods between the pre-lockdown, lockdown, and post-lockdown times. We further observe if the pandemic influenced forum users' communication regarding signs of extremism. We operationalize extremism own to incels as having three dimensions: toxicity, misogyny, and topic. We compare these three measures between the periods suggested by the piecewise regression.

**RESULTS** Based on the forum's activity data, we conclude that the beginning of the pandemic had the most decisive impact on the forum's activity. This is perhaps because lockdown measures worldwide started to differ later in the pandemic. From the topic distribution, we can observe that radicalization-related topics such as racism and rape decreased during the initial worldwide lockdowns and later increased to even higher levels than before the pandemic. We also observe that neither toxicity nor misogyny significantly changed during the period of interest.

**ADDED VALUE** In contrast to previous research, we analysed incels.is for a longer period (2019-2022) - thereby, we can identify the beginning

and end of the Covid-19-caused activity spike. This allows a comparison of pre-lockdown, lockdown, and post-lockdown periods to investigate the initial effect of the lockdown and the longer-lasting impact of Covid-19 on the forum. We employ topic modeling, misogyny, and toxicity scores to quantify the effect of the pandemic on this prominent online hub of radicalization. This contrasts previous research which draws conclusions mainly on keyword analysis and impression reading of selected posts.

## D3: INNOVATION IN PRACTICE (1): TRANSFORMING MR WITH DATA-DRIVEN TOOLS

**SESSION CHAIR: STEFAN OGLESBY**  
data IQ AG, Switzerland

### AUTOMATED EXPERT LISTENING: DATA-DRIVEN TREND RESEARCH WITH SONAR

**HÜCK, STEFFEN**  
TD Reply, Germany

As McKinsey put it in a 2021 piece, the “ability to learn, innovate, and seek good ideas regardless of their origin” is a trait shared by the most successful and future-ready companies in the world. This underscores the importance of trend research. However, in practice, trend research is often unsystematic and relies on subjective opinions of self-proclaimed trend gurus rather than solid data.

Fortunately, with the advent of new data-driven and AI-powered technologies, trend research can now be more fact-based and relevant for business success. Our innovative trend research platform, SONAR, developed in collaboration with the German Institute for Artificial Intelligence (DFKI), utilizes these cutting-edge technologies. SONAR is based on an Expert Listening logic, making it the first specialized expert listening tool that automatically extracts relevant information from expert media. Expert media refers to the specific expert sources where future trends usually first appear, including academic publications, patent databases, expert blog articles, specialist media, and science-oriented mass media.

At the heart of SONAR is a continuously growing expert media database that currently includes over 80 million articles. Through self-learning text mining algorithms and natural language processing (NLP), SONAR identifies trends and compares them with each other. Trends can be sorted by publication volume, year-over-year volume growth, or month-over-month volume growth, and can also be compared using filters such as media type and geographic origin. One of SONAR's key strengths is its ability to provide a historical perspective on trend development, as its database spans back to 2013. This long-term view is of great significance when forecasting future trends, as it allows for superior capabilities compared to more recently developed tools.

Today, we successfully use SONAR to support Fortune Global 500 companies in quantifying, analyzing and comparing relevant trends in a fact-based rather than in a subjective, gut-based manner.

We are confident that expert listening tools like SONAR can elevate the prestige of innovation units within companies and empower them to shape their strategic direction more effectively. Together with other innovative data-driven approaches such as search and social listening, expert listening can play an integral role in fostering data-driven decision-making.

### THE FUTURE OF THE DATA ECONOMY: MAKING BIG-TECH DATA ACCESSIBLE TO MARKET RESEARCHERS

**ARTS, JOHN**  
Rita Data, The Netherlands

In today's rapidly evolving data landscape, the rise of Zero-party data emerges as a game-changer for researchers and businesses alike. Join us for a comprehensive dive into the new data economy and its transformative implications for customer insights.

The democratization of consumer data has shattered old barriers. Gone are the days when the data monopolies of big tech firms dictated the tempo. Now, businesses of all sizes can access this invaluable resource, equalizing opportunities and setting the stage for innovative consumer insights. At the heart of this transformation lies Zero-party data: the consciously and proactively shared information by consumers. Unlike its predecessors – first and third-party data – Zero-party data offers a window into the genuine preferences, intentions, and desires of consumers without the shadow of privacy concerns.

Our session won't just theorize; we'll provide tangible proof. Notable brands like HelloFresh and Skyscanner are already harnessing the power of Zero-party data through our platform, Rita. Their experiences serve as a testament to the efficacy and potential of leveraging this unique data type.

By diving deep into real-world applications and the success stories of these industry leaders, attendees will gain a concrete understanding of how to utilize Zero-party data effectively. This isn't just about understanding the new age of data; it's about witnessing its transformative power in action and learning how to adapt accordingly.



Embrace this opportunity to stay ahead of the curve. Discover how to harness the unparalleled potential of customer insights in the era of Zero-party data, and ensure your research methodologies are primed for the future. Navigate the data revolution confidently with evidence-based insights.

## THE FAST AND THE CURIOUS – UNLOCKING MOMENTS THAT MATTER WITH AUTOMATED SURVEY RESEARCH

TURNER, JAMES; PELHAM, TONY

Delineate, UK

**RELEVANCE & RESEARCH QUESTION** In an increasingly competitive marketplace, brands are striving for innovation and a distinct competitive advantage through a nuanced understanding of “category to moment fit” — the alignment between product offerings and specific consumer experiences or “moments that matter.” Understanding these key moments is crucial for product innovation and strategy, yet traditional survey methods often fall short in capturing them due to limitations in respondent recall. The research question driving this paper is: How can real-time survey sampling facilitate better category to moment fit by enhancing consumer insights captured close to the moments that matter?

**METHODS & DATA** We employed an innovative approach to survey research by using a dynamic survey script and real-time, automated online surveys with intra-day sampling periods. This methodology allows us to reach consumers at or near real-life moments, thereby tapping into their episodic memory rather than semantic memory recall to understand motivations and context. In one example project 5,000 completed surveys were completed with highly accurate interlocking samples. In this example, key moments such as advertising exposure, shopping missions and category consumption were clearly identified.

**RESULTS** Findings show a notable difference between real-time survey data and traditional wave-based survey data. Real-time surveys exhibit improved robustness and accuracy when compared to wave-based traditional survey. Furthermore, respondents in real-time research and episodic recall periods offered detail of their experiences, capturing the emotional need and context of the moments. This leads to more actionable insights, revealing specific touchpoints and opportunities for intervention or product innovation and strategy.

**ADDED VALUE** The study revolutionises consumer research by shifting the paradigm from generalized insights to more discreet, moment-specific consumer segments. By harnessing real-time survey sampling, brands can gain a granular understanding of how different consumer moments map to specific categories, thereby greatly enhancing market understanding. This nuanced data enables businesses to create tailored marketing strategies, customised product offerings, and highly targeted customer experiences, increasing the potential for competitive. Importantly, the real-time approach eliminates certain forms of recall bias, which have historically been a significant impediment to reliable market segmentation.

## A4.1: DATA QUALITY IN ONLINE SURVEYS

(SPONSORED BY GESIS)

SESSION CHAIR: TOBIAS RETTIG

University of Mannheim, Germany

### DATA QUALITY INDICATORS: SOME PRACTICAL RECOMMENDATION

BHAKTHA, NIVEDITA; SILBER, HENNING; LECHNER, CLEMENS

GESIS - Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** It is very common in survey research to collect data on multi-item scales or constructs with Likert or categorical response options. Low quality data (LQD) in such samples has been a well discussed issue but is gaining further traction due to the rise in online data collection. Low quality data can be defined as those responses where the respondents have not put in sufficient thought and effort into responding. Consequently, there has been a surge in literature on methods for identifying LQD, however, the methods and best practices are scattered across disciplines. In this study, we aim to demonstrate the use of data quality indicators used across different fields such as psychology, sociology, epidemiology, and public health on multi-item scales. We will discuss various indicators used, the correlation among them, and the best practices associated with flagging low quality data.

**METHODS & DATA** We use publicly available German Longitudinal Election Study cross section wave 21, pre-election data for the demonstration. We use functions and approaches provided in the dataquieR and careless packages in R for data quality analysis. Different rules of thumb and thresholds will be applied to the data quality indicators for flagging problematic data.

**RESULTS** We observe that many data quality indicators should be used to flag low quality data. Different indicators pinpoint different underlying reasons for problematic data. The correlation among different indicators are typically low. Different thresholds and cut-offs for the indicators flag different number of observations as problematic responses.

**ADDED VALUE** We have unified data quality practices from different fields. We have demonstrated the use of data quality indicators in identifying low quality data in multi-item largescale survey data. We have

taken up the discussion on setting thresholds for the indicators to flag low quality data. Finally, we provide recommendations on combining different indicators to flag potentially problematic data.

## FIELDING A LONG ONLINE SURVEY: EVIDENCE FROM THE FIRST GENERATIONS AND GENDER SURVEY (GGS) IN THE UK

MASLOVSKAYA, OLGA; CHANG, GRACE; PERELLI-HARRIS, BRIENNA  
University of Southampton, UK

**RELEVANCE & RESEARCH QUESTION** Our team has collected the first Generations and Gender Survey (GGS) in the UK. This survey used push-to-web design with online-only mode available for respondents. The approximate length of time for survey completion was specified as being around 50 minutes for the respondents. The length of the online surveys topic has recently attracted a lot of attention from survey methodologists as many high quality social surveys moved to online data collection or mixed-mode designs in the recent years. The rule of thumb until recently was not to have online surveys of the length exceeding 10-20 minutes. Recent experiments conducted by the European Social Survey (ESS) suggest that it is possible to conduct longer (35 minutes or even 55 minutes) online surveys without significant reduction in data quality. However, more evidence is needed to establish the optimal length of online social surveys.

**METHODS & DATA** In this paper, we will present the evidence for fielding a long online probability-based survey. We reflect on the challenges and opportunities of conducting a long probability-based online data collection in the UK by reporting on nonresponse, break-off rates, and quality of responses. We also investigate the de-briefing questions in which respondents were able to reflect on how they felt about the survey. We will compare paradata for the length of time it took respondents to complete the survey with the respondents' perception on how long the survey was when these paradata became available in April 2023.

**RESULTS** Preliminary results suggest that despite the fact that the UK GGS survey was long and complex, 82% of respondents found the survey "not at all difficult". High proportion of respondents felt that the survey was about as long as they expected (47%) with further 19% felt that the survey was shorter than they expected. Also, another positive outcome of the survey was that 82% of participants gave consent to be recontacted for the second wave of the UK GGS survey.

**ADDED VALUE** Our findings provide evidence for the optimal length for long and complex online social surveys and have important implications for survey practice.

## ABILITY TO IDENTIFY FAKERS IN ONLINE SURVEYS: COMPARISON OF BIDR.SHORT.24 AND MCSD-SF

VÉSTEINSDÓTTIR, VAKA; KRISTJANSDÓTTIR, INGUNN ROS;  
STEINGRIMSDÓTTIR, KATRIN S. J.  
University of Iceland, Iceland

**RELEVANCE & RESEARCH QUESTION** Socially desirable responding (SDR) is a common problem in self-report measures, as the tendency to present oneself favorably to others can influence the honesty of responders. One facet of SDR is faking, an intentional misrepresentation in self-report. There are, however, two kinds of faking: faking good and faking bad. Faking good involves deliberately presenting oneself favorably to others, whereas faking bad involves deliberately presenting oneself in an undesirable manner. There are several ways to detect faking, one of them being SDR scales. Two of the most widely used scales are the Marlowe-Crowne Social Desirability Scale (MCSDS) and the Balanced Inventory of Desirable Responding (BIDR). To evaluate which SDR measure is better suited to detect SDR, one can compare their ability to detect faking. A previous study comparing the ability of MCSDS and BIDR to detect faking found that MCSDS outperformed BIDR in identifying both types of faking. A limitation of the applicability of those results is the fact that the comparison was between full-length versions of the measures but short-form versions are usually more practical as they reduce response fatigue. For that reason, the current study compared two short-forms' ability to detect faking, MCSD-SF (short-form version of MCSDS) and IM-Short.24 (short-form of the IM subscale of BIDR).

**METHODS & DATA** Participants were recruited online through a probability-based panel. The final sample consisted of 106 men and 122 women, others chose not to answer the gender question. Participants were randomly assigned to one of three groups: 1) standard instructions, 2) fake good instructions, or 3) fake bad instructions and then asked to complete both the MCSD-SF and the IM-Short.24.

**RESULTS** Discriminant function analyses and receiver operating characteristic curve analyses showed that, overall, MCSD-SF outperformed IM-Short.24 in identifying faking good, while IM-Short.24 outperformed MCSD-SF in identifying faking bad.

**ADDED VALUE** These findings show a clear preference for the use of MCSD-SF for identifying fake good responses and IM-Short.24 for identifying fake bad responses, which should assist researchers in choosing an appropriate measure for their studies, as well as advance the use of SDR measures overall.

## A4.2: MIXED MODE AND MODE TRANSITION II

[SPONSORED BY GESIS]

**SESSION CHAIR: CARINA CORNESSE**

DIW Berlin & RISC, Germany

### DATA COLLECTION IN OFFICIAL STATISTICS: RESPONSIVE, SMART AND MIXED

**BLANKE, KAREN; KNAPP, DANIEL**

Federal Statistical Office, Germany

**RELEVANCE & RESEARCH QUESTION** The future of mixed- mode/devices data collection in Official Statistics and especially in social surveys is an ongoing discussion, both at international and national levels. With on-line data collection becoming the preferred option in recent years ("On-line First" strategy) and an ever-increasing number of mobile users in online surveys, the situation became even more complex. Consequently, Destatis has strengthened its effort to strive for a respondent-centered survey design: offering different modes/devices and using these to their full potential.

In order to meet these goals, both a clear vision for future data collection as well as a thorough redesign of our online surveys, following a fully responsive approach, have been decided on. Both with the aim to achieve a survey interface that not only accommodates mobile users, but generally increases the user-friendliness and provides a positive user experience, thereby increasing respondents' motivation to participate and thus to increase response rates.

**METHODS & DATA** In order to accomplish this redesign of our online questionnaires, several iterative functionality and usability tests are being conducted at Destatis, featuring cognitive qualitative interviews.

**RESULTS** Final results of the qualitative pretests are still work in progress and will be handed in at a later date. The presentation will also briefly depict an overview of the data collection vision, including the introduction of mobile apps as a novel mode of data collection (Time-Use-Survey 2022, Household Budget Survey in 2023).

**ADDED VALUE** This presentation will demonstrate our data collection strategy in social surveys, the implementation of apps (advantages and disadvantages) and show how to overcome design challenges in fully responsive online surveys, in order to meet the reality we're facing, thus serving as input for other survey researchers encountering similar challenges.

### INTRODUCING WEB IN A TELEPHONE EMPLOYEE SURVEY: EFFECTS ON COVERAGE, NONRESPONSE AND COSTS

**MACKEBEN, JAN; SAKSHAUG, JOE**

Institut für Arbeitsmarkt- und Berufsforschung, Germany

Policy decisions in business and economic fields are often informed by surveys of employees. Many employee surveys use costly interviewer administered modes to reach this special population. However, certain employee subgroups may be especially hard to reach using these modes. Thus, besides high administration costs, nonresponse bias is a concern. To reduce costs and potential nonresponse bias, some employee surveys have introduced web as part of a sequential mixed-mode design.

However, the impact of introducing web on response rates, nonresponse bias, and costs in employee surveys is understudied. The present study addresses this research gap by analyzing a mode design experiment in which employees selected for a national survey in Germany were randomly assigned to a single-mode telephone design or a sequential web-telephone mixed-mode design. The study revealed four main findings.

First, introducing the web mode significantly increased the response rate compared to the single-mode design. Second, despite the higher response rate, aggregate nonresponse bias was higher in the mixed-mode design than in the single-mode design. Third, the likelihood of web participation varied across certain employee subgroups, including occupation type and employment contract. Lastly, potential cost savings were evident under the mixed-mode design.

# B4: COMBINING SELF-REPORTS WITH DIGITAL BEHAVIORAL DATA

**SESSION CHAIR: STEFAN OGLESBY**  
data IQ AG, Switzerland

## STUDYING ONLINE PORNOGRAPHY USE WITH A COMBINATION OF WEB TRACKING AND SURVEY DATA: RESULTS FROM A GERMAN ONLINE SAMPLE

**DOLAY, ORKAN (1); SIEGERS, PASCAL (3);  
VON ANDRIAN-WERBURG, MAXIMILIAN (2);  
BREUER, JOHANNES (3); RATHE, CLEMENS (1);  
NOWAK, HOLGER (1)**

1: Bilendi & respondi, Germany  
2: University of Würzburg, Germany  
3: GESIS - Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** Research about the use and effects of pornography has almost exclusively relied on self-reported data from survey studies. Following a call to use more reliable data sources, we conducted a series of analyses using web tracking data in combination with surveys to replicate questionnaire-based findings about patterns and predictors of online pornography use (OPU). To gain deeper insights into usage types and their correlates, we extend this first analysis based on top-level domains by adding more fine-grained data on content categories of online pornography (OP) based on classifications using full URLs from the web tracking.

**METHODS & DATA** Bilendi & respondi maintain a panel where participants have agreed to install a software metering their internet use. The data include information about which websites they visit, when and for how long. For our first analyses, we used tracking data on the top-domain level from  $N = 3018$  panelists collected between June 2018 and June 2019 and combined them with data from an online survey among a subset of panelists ( $n = 1315$ ). For the follow-up analyses, we used content classifications based on full URLs using text mining methods. These classifications are used for in-depth investigation into usage types and their sociodemographic and attitudinal correlates.

**RESULTS** Overall, the results from our initial analyses confirm existing findings from questionnaire-based research related to sex and age dif-

ferences. Our data also shows temporal patterns, indicating that – for most users – online pornography use constitutes a leisure time activity that competes with other spare time activities. Regarding predictors of OPU, we found that while gender role attitudes, sexism, and social dominance orientation are not associated with OPU, we do find small differences between members of different religious communities.

**ADDED VALUE** Using web tracking data, we can assess the reliability and robustness of findings based on self-report data which can be biased when studying OPU. Overall, our results indicate that previous results from self-report studies largely hold when using tracking data. Additional analyses based on classifications of full URLs allow us to get a better understanding of specific content preferences and their predictors.

## MEASURING FACEBOOK USE: THE ACCURACY OF SELF-REPORTED DATA VS. DIGITAL TRACE DATA

**PANKOWSKA, PAULINA KAROLINA (1); KEUSCH, FLORIAN (2);  
BACH, RUBEN (2); CERNAT, ALEXANDRU (3)**

1: Utrecht University, The Netherlands  
2: Mannheim University, Germany  
3: University of Manchester, UK

**RELEVANCE & RESEARCH QUESTION** The growing availability of digital trace data in recent years has prompted social and behavioral scientists to rely more on these data sources in their research. Information obtained from these data is increasingly used in current studies to replace surveys, or as a benchmark to validate and assess the quality of survey-based measures. These studies often rely on the unrealistic assumption that log-data are error-free.

However, much like any other data source, digital traces are likely to be subject to non-negligible error. Research related to social media and its impacts is a domain in which these concerns are particularly prominent. Therefore, in this paper, we examine the quality of self-reported and digital-trace measures of Facebook use simultaneously, while allowing each of the sources to contain measurement error.

**METHODS & DATA** To assess and compare the nature and incidence of the error in these two sources, we apply hidden Markov models to a sample from a non-probability online panel in Germany. The self-reported measures are based on a longitudinal survey, the digital-trace measures are based on information from tracking apps that were installed on the respondents' smartphones and/or computers.

**RESULTS** Our results suggest that both sources measure Facebook use rather consistently for about two thirds of the sample. For the remaining one third of the respondents, we observe large inconsistencies between the survey and the digital-trace measures. We find evidence that some of the inconsistencies between survey and digital-trace measures are due to an underestimation of Facebook use for a sub-sample of respondents in the tracking data. This underestimation is likely to be due to incomplete tracking of use behavior across all of their relevant devices.

We further show that the contradictory results obtained when using the two sources to assess the effect of Facebook use on a variety of political outcomes can be reduced when the inconsistent group is removed from the analysis.

**ADDED VALUE** Our research shows the importance of accounting for measurement error in new, 'big-data' sources as well, as assuming these data to be error-free can lead to biased results and erroneous findings.

## D4.1: INNOVATION IN PRACTICE (2): OPPORTUNITIES AND CHALLENGES

**SESSION CHAIR: CATHLEEN STÜTZER**  
TU Dresden, Germany

### REPLICATING COGNITIVE BIASES: USING SOCIAL MEDIA AS A NOVEL PARTICIPANT POOL FOR RESEARCH

**BUBLITZ, ANDREA (1); UEBERWASSER, RAPHAEL (2)**

1: University of Zurich, Switzerland  
2: Boomerang Ideas AG, Switzerland

**RELEVANCE & RESEARCH QUESTION** A/B tests are crucial for companies to test new ideas. While big tech companies regularly conduct experiments, smaller companies often lack the means to establish the necessary infrastructure. Marketers increasingly rely on social media for A/B tests but research has revealed that these tests are biased due to algorithmic optimization. To lower barriers to market entry, Boomerang Ideas offers a self-service to conduct unbiased A/B tests on social media. The first survey question is displayed as an ad on social media. Participants who click on the ad are randomly assigned to a stimulus only after leaving the social media newsfeed. Even though Boomerang Ideas mitigates the biasing effect of social media algorithms, it remains an open question whether social media is a valid participant pool for market research.

Social media users suffer from an overload of information. Users often access social media on their smartphones further reducing available cognitive capacity. This project hence assesses the performance of A/B tests on social media by replicating a well-established cognitive bias, the anchoring effect.

**DATA & RESULTS** We conduct a survey of Swiss users on Facebook, Instagram, Snapchat, and LinkedIn (N = 221). The social media ad states that Germany now offers monthly passes for the regional public transport for 49 EUR/month and asks whether Switzerland needs a similar offering. Participants in the high (vs low) anchor condition are then told that the monthly pass for the Swiss public transport costs 340 (vs 75) CHF/month. We measure willingness to buy a pass for the regional public transport for 50 CHF/month from 1 (No, never) to 4 (Yes, definitely). According to the anchoring effect, we expect users in the high anchor condition to be more willing to buy. A t-test uncovers that participants in the high anchor condition are indeed more willing to buy ( $M_{low} = 2.9$ ,  $M_{high} = 3.3$ ;  $p < .001$ ). Randomization checks further support that there are no differences between treatment groups with regard to gender ( $p = .56$ ), age ( $p = .97$ ), language ( $p = .21$ ), and social media platform ( $p = .42$ ).

**ADDED VALUE** This project uncovers the potential of social media as a new participant pool for market research. Using the A/B test tool by Boomerang Ideas the anchoring effect replicates on social media. Participants' demographic details further indicate that the Boomerang Ideas algorithm successfully counteracts social media algorithms yielding randomized treatment groups.

### TRUSTWORTHY ANALYTICS WITH GENERATIVE AI: FOUR USE CASES FOR CHATGPT/GPT-4

**WITTENBURG, GEORG**  
Inspirient GmbH, Germany

**RELEVANCE & RESEARCH QUESTION** Rigorous, well-documented data analysis is a corner stone of opinion research. ChatGPT and GPT-4 are the most recent additions to the toolset of every researcher/social scientist and are thought to have the potential to boost efficiencies by 50% and more – but they have trouble reliably adding more than three numbers, let alone apply complex statistical methods! This talk takes a use-case-driven approach to answering the following research questions: How can current Large Language Models (LLMs), like ChatGPT / GPT-4, be applied to rigorous analysis of survey data? Which use cases are feasible, and under which tradeoffs?

**METHODS & DATA** We've jointly identified four major use cases for LLMs applied to the analysis of raw survey data: 1) Natural-language data exploration, 2) desk research support, 3) generation of deliverables, and 4) LLM chatbots as client-facing deliverables. For each of these use cases, we've preprocessed samples of raw survey data in a fully automated process to allow current LLMs to be applied, as per the use case requirements. We've validated the textual results across teams to understand their respective potential and shortcomings, and consequently improved both our automated preprocessing as well as LLM prompting.

**RESULTS** The results across the four use cases are impressive: Given nothing more than raw survey data (in form of an SPSS or Microsoft Excel file), we were able to reliably use LLMs to produce relevant output for

each use case. The key aspect of our approach is the capability to link each LLM statement back to its original analytical results, thereby establishing a chain of trust into each statement of the LLM conversation that regular LLM use would lack otherwise.

**ADDED VALUE** At this time it's hard to fully estimate the automation potential&resulting efficiency gains for research agencies. We can however state with confidence, that the provided samples for all four use cases were generated within about one hour, for the given SPSS/Excel input data. In our talk, we'll demonstrate this process step-by-step, and together – based on a live demo – explore potential and limitations of this approach.

## WHEN DIGITAL NATIVES STRUGGLE WITH DIGITAL SPACES: HOW TO COMPENSATE FOR INSECURITIES IN ONLINE VIDEO MEETINGS

VON CORVIN, BARBARA (1); WAHRENBARGER, SYBILLE (2)

1: Human8, Belgium

2: Hochschule Fresenius, Germany

Als Folge der Kontaktbeschränkungen durch die Covid-19-Pandemie wurden Meetings und Fokusgruppen nicht mehr offline in Person, sondern als Online-Videomeetings durchgeführt. Dabei konnte in der Praxis ein vergleichsweise zurückhaltendes Kommunikationsverhalten der Altersgruppe der Gen Z [1995-2010] im Vergleich zur Gen Y [1981-1995] beobachtet werden sowie eine damit einhergehende weniger stark ausgeprägte Interaktionsfähigkeit. Der digitale Raum wird den Digital Natives damit offenbar zum Verhängnis.

In zwei darauffolgenden spezifisch ausgerichteten Studien sollte diese Vermutung genauer untersucht werden:

- Eine erste Studie im Sommersemester 2022 explorierte in 20 halbstrukturierten Interviews das Kommunikationsverhalten der Gen Z und Gen Y in Offline- und Online-Gruppenkonstellationen. Die Ergebnisse zeigten auf, dass sich viele Mitglieder der Gen Z in Online-Videomeetings unsicher fühlen, insbesondere durch eine verminderte gestische und mimische Unterstützung der Interaktion. Infolgedessen wurden die geringere Bestätigung der eigenen Äußerungen durch andere, fehlende Anzeichen für einen passenden Zeitpunkt eigener Beiträge und die nicht greifbare Einschätzung der eigenen Wirkung durch die Beteiligten thematisiert.
- Die gewonnenen Erkenntnisse bildeten die Basis für eine im Wintersemester 2022/2023 durchgeführte, tiefergehende Mixed-Method-Untersuchung des Kommunikationsverhaltens von 18- bis 27-jährigen. Dabei wurde Kommunikation sowohl in Online-Video-Meetings als auch im Rahmen von VR-Meetings [Metaverse] anhand von Befragungen, Beobachtungen und Messung biometrischer Daten (Eyetracking, EDA, EKG) analysiert.

In ersten Ergebnissen zeigte sich, dass auch im Metaverse Gestik und Mimik zwar nicht ausreichend übertragen werden, aber allein der erzeugte Eindruck der physischen Nähe im virtuellen Raum ein Gemeinschaftsgefühl zustande bringt, wodurch die Teilnehmenden sich deutlich wohler fühlen und weniger Hemmungen haben, ihre Meinung offen in der Gruppe zu teilen.

Aus den zahlreichen Einzelerkenntnissen lässt sich festhalten, dass die beobachteten Aspekte in der heutigen Zeit von höchster Brisanz sind, tiefergehend untersucht zu werden, um die soziale Interaktion in digitalen Räumen weiterhin optimieren zu können. Wir würden uns freuen einige interessante Punkte dieses Forschungsprojektes im Rahmen eines Vortrags vorstellen zu dürfen und mögliche Lösungsmöglichkeiten zu diskutieren.

FRIDAY, 22/SEPTEMBER/2023

## A5: INCREASING SURVEY PARTICIPATION (SPONSORED BY GESIS)

SESSION CHAIR: PETER LUGTIG

Utrecht University, The Netherlands

## IMPACT ON PREPAID AND POSTPAID INCENTIVE EFFECTIVENESS OF SURVEY MODE AND INCENTIVE TYPE

LIETZ, ALMUTH; KÖHLER, JONAS; LI, IDA; SIEGEL, MADELEINE

Deutsches Zentrum für Integrations- und Migrationsforschung (DeZIM), Germany

**RELEVANCE & RESEARCH QUESTION** Incentives in surveys are a frequently used instrument to increase response rates and have been widely researched. For postal surveys it was shown, that prepaid incentives are more effective than postpaid incentives, especially when cash is offered. However, in web and mixed mode surveys, prepaid incentives are not necessarily associated with higher response rates. Accordingly, we assume that the survey mode (postal or online), determines the outcome and has a moderating effect regarding the effectiveness of



prepaid compared to postpaid incentives. Furthermore, we investigate whether the results differ when vouchers are offered instead of cash.

**METHODS & DATA** To test our assumptions, we applied two experimental designs to two survey waves of the German DeZIM.panel: In the first experiment, that was included in the recruitment wave, we tested the effect of the cash incentive (€5 vs. €10) as well as the payment type (unconditional prepaid vs. conditional postpaid). In this experiment, people could participate both by paper and online questionnaires. In the second experiment, that was integrated in the first wave, we examined voucher incentives (€10) as well as the payment type (unconditional prepaid vs. conditional postpaid). Here, participants could only participate online. To evaluate the results of the experiments, we consider several measures of data quality: Response rates, willingness to participate again, and item non-response. Our data base - the DeZIM.panel - is a randomly drawn, offline recruited online access panel in Germany with an oversampling of specific immigrant cohorts. Since 2022, regular panel operation has taken place with four waves per year, supplemented by quick surveys on current topics.

**RESULTS** For both survey modes, prepaid incentives result in higher response rates, but a lower willingness to participate in the panel. In addition, prepaid incentives are superior in terms of response rates when vouchers are offered.

**ADDED VALUE** Our preliminary results indicate, that prepaid incentives may especially motivate for current participation but not for panel willingness. Postpaid incentives seem to have more binding power for panel participation and may be a less cost-intensive and practical alternative, especially for online access panels.

## QUID PRO QUOTA? THE EFFECTS OF INCENTIVES ON SURVEY PARTICIPATION

WEITZEL, DANIEL (1,2); PFAFF, KATHARINA (1); KRITZINGER, SYLVIA (1)

1: Universität Wien, Austria

2: Colorado State University, USA

**RELEVANCE & RESEARCH QUESTION** The use of computer-assisted web interviews (CAWI) and thus the reliance on online panels has increased over the past decades. While quite some research has already examined the factors impacting response rates in cross-sectional surveys, our research focuses on factors enhancing the setup of offline-recruited online panels - and thus mid- to long-term factors.

**METHODS & DATA** We examine four types of monetary incentivisation strategies and their effect on participation at online surveys and web panel recruitment. More specifically, we compare unconditional monetary incentives, conditional monetary incentives, conditional donations to a charity organization, and leaving respondents the choice between a monetary incentive or donation. By contrasting panel recruitment with panel participation in subsequent waves, we test amongst others whether these four incentive strategies differ in successfully recruiting online panelists in the

longer run. Drawing a sample stratified by region (NUTS3) from the Austrian Central Population Register (ZMR), a public register in which all persons registered in Austria are recorded, our analysis encompasses over 2.000 respondents across multiple survey waves fielded in 2022.

**RESULTS** Results show that the effect of incentives is heterogeneous in nature. While unconditional incentives increase survey participation in terms of response rate, other incentivisation strategies perform better at panel recruitment and thus survey participation in the longer run.

**ADDED VALUE** Prior research tends to focus on one particular outcome such as strategies to improve response rates. This paper shows that conflicting goals are likely to arise in survey design and that researchers need to be clear about the primary survey goal in order to develop the optimal design.

## B5: QR-CODE TECHNOLOGY

SESSION CHAIR: FLORIAN KEUSCH

University of Mannheim, Germany

### THE EFFECT OF QUICK RESPONSE (QR) CODES ON PANEL RECRUITMENT AND SURVEY RESPONSE RATES

WEISSENBILDER, MARCUS; LUNDMARK, SEBASTIAN; BERGQUIST, JULIA  
SOM-Institutet, Sweden

**RELEVANCE & RESEARCH QUESTION** Quick Response (QR) codes can visually represent a web address (URL) on printed materials and scanning it will redirect the user to the URL depicted. Printing QR-codes in addition to the URL may stimulate response rates if scanning it requires less effort than typing the URL. However, research has observed weak relationship between response rates and QR-codes, potentially because few respondents understand how to scan them. In Sweden, nine out of ten residents used QR-codes to access public services in 2021. Assessing the impact of QR-codes in Sweden should pose a strong case for finding a positive influence of QR-codes on response and panel subscription rates. We assess whether this is the case.

**METHODS & DATA** Study 1: 36,000 individuals from a simple random sample of the Swedish population were invited to join an online panel and provided a login and password to enter the profile questionnaire. One group received the invitation with only the URL to the login screen, one group got the URL and QR-code to the login screen, and one group got the URL but a QR-code that bypassed the login screen. Study 2:

38,000 individuals from a simple random sample of the Swedish population were invited to complete a cross-sectional mixed mode questionnaire. In addition to the paper-and-pencil version, respondents were offered to complete the questionnaire online. Respondents were randomly assigned to the same QR-code manipulations as in Study 1.

**RESULTS** Study 1 found that recruitment rates did not increase by offering the two types of QR-codes. Furthermore, QR-codes pushed respondents to complete the questionnaire on smartphones, albeit without affecting data quality negatively. Study 2 showed that response rates were not increased by offering the two types of QR-codes. Again, QR-codes pushed respondents to complete the questionnaire on smartphones, but this time negatively affected data quality.

**ADDED VALUE** We show that a majority of our respondents used the QR-codes to access our online questionnaires, but even in an optimal case such as Sweden, adding the codes did not increase response rates. In fact, adding them resulted in more breakoffs.

## PROMISES AND PERILS OF USING QUICK RESPONSE CODES IN PROBABILITY-BASED WEB PANEL SURVEYS

ROSSMANN, JOSS

GESIS - Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** Respondents for probability-based web surveys of the general population are often invited in personalized letters as complete lists of email addresses are not available in most settings. To ease access to the web questionnaire, survey operators increasingly use scannable quick response (QR) codes in mailings in addition to links and personal access codes, which must be entered manually by respondents. Research on the effects of implementing QR codes in survey mailings is very limited to date. An exception is a study by Lugtig and Luiten [2021] who found that using QR codes did not affect login or breakoff rates but increased the use of smartphones. The research objective of this contribution is deepening our understanding of the effects of QR codes on response and device usage in probability-based web surveys.

**METHODS & DATA** The data come from two probability-based samples of a German panel study: The first sample consists of experienced panel members who had mostly provided an email address (>80%), while the second sample includes freshly recruited respondents for whom email addresses were rarely available (35%). QR codes were introduced in addition to links and personalized access codes in postal invitations and reminder letters of two consecutive waves with 5,014 and 4,915 respondents in 2022. Panel members who provided a valid address also received emails with a personalized link.

**RESULTS** Preliminary analyses showed that respondents from the freshly recruited sample more often accessed the web questionnaires via QR

code than those of the experienced sample who more likely used the personalized link in email invitations and reminders. Further, analyses indicated that accessing the questionnaires via QR code or personalized links in emails was associated with an increased likelihood of using a smartphone. Using this type of mobile device increased the risk of survey breakoff, while QR code access did not directly affect dropout. However, there was little indication of increased survey response overall.

**ADDED VALUE** This contribution extends the knowledge on the promises and perils of QR codes in probability-based web surveys and supports practical survey design decisions.

## C5: POLITICAL ISSUE SALIENCE

[SPONSORED BY APROXIMA]

SESSION CHAIR: VERED ELISHAR-MALKA

Yezreel Valley College, Israel

## DOES LOCAL WEATHER AFFECT CLIMATE CHANGE SALIENCE? EVIDENCE FROM ONLINE ARTICLE TRACKING DATA IN GERMANY

HAGEMEISTER, FELIX (1); MÜLLER-HANSEN, SÖREN (2)

1: Süddeutsche Zeitung Digitale Medien, Germany

2: Süddeutsche Zeitung, Germany

**RELEVANCE & RESEARCH QUESTION** If climate change is a problem that needs to be addressed politically, it is hard to overstate the importance of public attention towards climate change. With increasing temperatures and more frequent abnormal weather events, the question arises how elastic issue salience of climate change is to local weather. Despite many studies, it remains unclear to which extent local weather influences issue salience of climate change (Howe et al. 2019). Being able to estimate this effect will help to gauge the extent to which public attention for climate change might follow a similar upward trend as temperatures and abnormal weather events.

**METHODS & DATA** We measure salience with novel online tracking data on climate-related article views of a leading German newspaper in different locations in Germany since 2021. First, we use text data methods to group articles in climate- and non-climate-related articles. Second, we aggregate group-specific articles views by day and location and calculate the share of climate-related articles amongst all articles published on a given day. We match the tracking data with publicly available weather data from German weather stations. Finally, we employ a two-

way fixed effects design to estimate the effect of increased temperature, precipitation, and abnormal weather on the share of climate-related article views. We complement our analysis with online survey data of subscribers to elicit motivations and effects of reading climate-related news articles.

**RESULTS** Results are forthcoming. The study is in its pre-registration phase with the Center for Open Science (OSF), so no analysis has been carried out yet.

**ADDED VALUE** Our contribution to the literature is threefold. First, we are the first to use novel fine-grained business data from a leading German digital news outlet to measure issue salience of climate change more objectively. Second, we go beyond mere correlations and provide an estimate of the relationship between weather changes and climate change salience that accounts not only for various observed differences between people and across time, but also for many unobserved potential confounders. Finally, we complement our analysis with our own survey to examine mechanisms.

## TWO YEARS OF TRENDING TOPICS: MEASURING ISSUE ALIGNMENT IN THE GERMAN TWITTERSPHERE

**POURNAKI, ARMIN {1,2,3}; GAISBAUER, FELIX {4,2}; OLBRICH, ECKEHARD {2}**

1: Max-Planck-Institut für Mathematik in den Naturwissenschaften, Germany

2: Laboratoire Lattice, CNRS & ENS-PSL & Université Sorbonne nouvelle, France

3: Sciences Po, médialab, France

4: Weizenbaum-Institut e.V., Germany

**RELEVANCE & RESEARCH QUESTION** Twitter has been extensively studied, but analyses that span up several discussion topics on the platform remain rare. The present work fills this research gap by providing a high-level view on public debate on the platform. It investigates the interaction structure as well as the content of trending topics (‘trends’) of German Twitter over two years (03/2021-06/2023).

**METHODS & DATA** We collected roughly 20 million tweets mentioning each day's most salient trends of German Twitter. We construct one retweet network per trend, which we partition into different clusters using stochastic blockmodeling. On that basis, we (i) group trends together thematically by computing a sentence-embedding-based topic model using BERTopic. (ii) We embed trends in a two-dimensional space, using a UMAP embedding of the Jaccard distance matrix (with respect to users tweeting about the trends). (iii) We examine issue alignment by quantifying the similarity of retweet network clusterings across the determined topics.

**RESULTS** (i) Topic modeling allows us to categorize the trends into larger topics, the most prevalent being German internal politics, the Russian aggression against Ukraine, sports, Covid, climate

change, and foreign politics. Via (ii), we can clearly distinguish clusters of thematically similar trends. The embedding shows that sports and gaming are made up of a user basis rather isolated from political trends. The latter form a coherent cluster that is ordered along a time axis, which suggests a high user turnover during the observed timespan. (iii) We measure average clustering distance between topics, which tells us how well two topics align users into similar opinion clusters. We see that Covid-related trends exhibit the highest alignment within a topic cluster, meaning that users' opinions on, e.g., vaccination, masks, or school closures are aligned. With respect to alignment across topics, we find that political topics such as the Russian aggression against Ukraine, climate change and gender identity form a cluster of aligned topics, whereas spontaneous events such as sports, TV shows or weather-related news exhibit weaker alignment.

**ADDED VALUE** This new type of analysis can improve our understanding of political polarization and ideological alignment on German social media.

## A6.1: RESPONDENT INTERACTION

(SPONSORED BY GESIS)

**SESSION CHAIR: FABIENNE KRAEMER**

GESIS Leibniz – Institute for the Social Sciences, Germany

## HOW EYE CONTACT WITH A ROBO-ADVISOR SHAPES INVESTMENT DECISIONS

**SCHALLNER, RENE; KAISER, CAROLIN; MANEWITSCH, VLADIMIR**

Nuremberg Institute for Market Decisions, Germany

**RELEVANCE & RESEARCH QUESTION** Current financial robo-advisors which provide consumers with investment recommendations, are predominantly realized as text-based websites. With the current rise of social robots as service providers, we ask the question of how varying degrees of human-like appearance and behaviour, especially eye contact, of a robo-advisor influences the trust and investment decisions of consumers.

**METHODS & DATA** In an online experiment, ca. 4500 participants took part in an online consultation with a randomly chosen financial advisor, either a human or a robot with varying degrees of human-likeness, with or without eye contact – or a textual advisor.

To create the consultation, we employed a professional actress and the social robot “Furhat” as financial advisors, recorded typical speaking segments, and merged them into a virtual advisory dialogue. After the consultation, participants had to choose the amount to invest and if they wanted their funds managed by a human or a machine, followed by a questionnaire asking for their experiences, reactions, and decisions, as well as their technical, financial, and socio-demographic background. The investment was simulated, and participants received a payout based on the invested amount and the investment performance. Standard ANOVA, ANOVA with posthoc Tukey-HSD/Games-Howell statistical analysis, and regression analysis were performed on the collected data.

**RESULTS** We found that robotic advisors are trusted less than humans but lead to significantly higher satisfaction and probability of choosing the machine-managed asset than text-based advisors. Eye contact significantly increased the trust, liking, satisfaction, and recommendation likelihood of the human and even the robotic advisors.

**ADDED VALUE** We show that human-like robotic advisors perform significantly better than textual ones and how anthropomorphism matters. We also demonstrate the beneficial effects of eye contact, even in robotic advisors. Our findings help creators of advisory offerings to provide better service with increased trust, customer satisfaction, and the willingness to follow advice.

We raise awareness of consumers that they are more likely to trust and follow advice if an advisor has human-like features. Understanding the possibilities and limitations of AI's influence on consumer decisions is important for discussions about the role of AI in society.

## THE SOLIDARITY GAME: EXPLORING THE METHODOLOGICAL CHALLENGES OF REAL-TIME RESPONDENT INTERACTIONS IN WEB SURVEYS

HÖHNE, JAN KAREM (1); GOERRES, ACHIM (2);  
KEMPER, JAKOB (2); TEPE, MARKUS (3)

1: DZHW, University of Hannover, Germany

2: University of Duisburg-Essen, Germany

3: University of Oldenburg, Germany

**RELEVANCE & RESEARCH QUESTION** Self-administered web surveys are a predominant methodology for measuring people's attitudes and behaviors in social science research. Although such surveys facilitate drawing large-scale (prob and non-prob) samples, they are frequently criticized for collecting indirect measures that do not sufficiently mirror people's real attitudes and behaviors. One way to overcome this methodological constraint is to adopt interactive, lab-based measurement techniques from behavioral research, such as the solidarity game. The solidarity game is a one-shot game with three players that is used to measure people's willingness to help others who are in a worse position

than they themselves (Selten and Ockenfels 1998). In this study, we address the following research question: What are the methodological challenges of incorporating the solidarity game in a self-administered web survey?

**METHODS & DATA** We conducted a cross-national web survey in Austria and Germany in March 2023 using the open-source research platform oTree. In each country, we drew a quota sample of about 1,100 participants and matched them into groups of three to play the solidarity game over ten rounds with monetary rewards (max. of 1€ per round). Before the start of the solidarity game, participants received a game introduction and were randomly assigned to one out of three experimental conditions: control (no reward information during rounds), limited transparency (reward information during rounds), and full transparency (reward and participant behavior information during rounds).

**RESULTS** The preliminary results indicate that the group matching remains a key problem since only about 74% of all respondents that are willing to participate can be successfully matched. In addition, about 28% of the successfully matched groups must be eliminated because of break-offs. This particularly applies to early rounds (rounds 1 to 3). These preliminary results hold for both Austria and Germany.

**ADDED VALUE** This study is a first attempt to introduce interactive measurement techniques from behavioral research to large-scale web surveys. It extends the existing methodological toolkit for measuring attitudes and behaviors and provides new avenues for innovative web survey design.

## THE DICE SURVEY: HOW ITEM ORDER IN A VALIDATED PSYCHOLOGICAL STRESS INSTRUMENT MAY AFFECT THE OUTCOME

CARLANDER, ANDERS; CASSEL, FELIX; ENSTRÖM, DANIEL;  
LUNDMARK, SEBASTIAN

SOM Institute, University of Gothenburg, Sweden

**RELEVANCE & RESEARCH QUESTION** Validated survey instruments are critical for medical diagnoses and are usually presented in a fixed order between items. However, question order bias may influence the validity of such instruments. Online research facilitate randomization more efficiently than previous survey modes. We ask how a stress instrument (PSS-10) performs under four different settings depending on item order (set vs. random) and when combined with questions regarding mundane habits (e.g., “How often during the last month have you commuted?”).

**METHODS & DATA** This pilot study consisted of a non-probability sample (N=2,624) from the Swedish Citizen Panel (62 % males; age 42.2 ± 7.9). Participants were randomly assigned to (1) PSS-10 with set single-item order with consecutive set single-item order on habits, (2) PSS-10 with random single-item order with consecutive random single-item order on

habits, (3) PSS-10 and habits mixed with random single item presentation, (4) PSS-10 and habits mixed with random matrix presentation.

**RESULTS** No significant differences on PSS-10 between groups ( $p=.479$ ) was found but significant differences ( $p < .001$ ) on time spent in seconds was found where the random matrix presentation took longest to answer ( $228.8 \pm 89.2$ ) compared with set single-item order ( $204.5 \pm 93.2$ ), random single-item order ( $201.5 \pm 81.28$ ) and mixed with random single item presentation ( $202.4 \pm 82.4$ ). No significant differences on number of missing items were found ( $p = .051$ ). The difference in psychometric performance between the four groups were marginal, McDonald's  $W$  ranged from .864 to .872 (random order performed better), Goodness-of-fit from a confirmatory factor analysis demonstrated an equally poor fit across the four groups ( $RMSEA \approx .10$ ,  $CFI \approx .90$ ,  $SRMR \approx .06$ ). A Graded Response Model indicated that the random matrix presentation possessed a more uniform function over different levels of the latent construct and that more information was retrieved.

**ADDED VALUE** Despite popular belief that fixed orders are required for validated psychometric scales, we demonstrate that item order bias in combination with uncorrelated questions may be negligible when using PSS-10. Replication and extension of these findings in combination with correlated (i.e., health-related) questions to test for local item dependence is warranted and data will be collected in June 2023.

## A6.2: BIAS ASSESSMENT AND CORRECTION

(SPONSORED BY GESIS)

**SESSION CHAIR: ELLEN LAUPPER**

Swiss Federal University for Vocational Education  
and Training, Switzerland

### THE RELATIONSHIP BETWEEN NON- RESPONSE RATES AND NONRESPONSE BIAS. A NETWORK ANALYSIS

**DICKSON, SHANNON; LUGTIG, PETER; STRUMINSKAYA, BELLA**

Utrecht University, The Netherlands

**RELEVANCE & RESEARCH QUESTION** Response rates are still often used as a quality metric for surveys. In order to achieve high response rates a lot of effort is used during survey fieldwork, such as the use of mixed-modes, reminders and incentives. While we know that some of these

survey design characteristics are successful at reducing nonresponse rates, we know little about whether they are successful at reducing non-response bias, let alone what specific combinations of survey design features work.

**METHODS & DATA** We report on an updated meta-analysis based on an original meta-analysis by Groves and Peytcheva (2006). The updated meta-analysis consists of 187 studies, and 3301 estimates of nonresponse bias, along with survey design features. We apply network models to examine the causal structure of (combinations of) survey design features predicting both nonresponse rates and nonresponse bias.

**RESULTS** We find the relation between nonresponse rates and bias to be weak ( $r=.2$ ) overall. Survey mode is a central node in our network models, confirming anecdotal evidence that the choice of survey mode is a driving factor of the use of other survey design features. When results are split by mode, we find that the network models look very different, implying that different combinations of survey design features are linked to higher rates of nonresponse rates. Survey design features are only predictive of nonresponse bias in mail surveys.

**ADDED VALUE** This paper adds to the body of evidence showing that response rates cannot be used as proxy information for the risk of non-response bias. We also show that characteristics of the survey that are often used to increase response rates are only related to nonresponse bias in mail surveys, and not in other modes. We will discuss implications for survey practice in relation to fieldwork costs and fieldwork efforts.

### IN THE PRACTITIONER'S SHOES: A COMPARISON OF CORRECTION METHODS FOR NON-PROBABILITY SAMPLES.

**ARLETTI, ALBERTO (1); PACCAGNELLA, OMAR (1);**

**BARTOLI, BEATRICE (2)**

1: Department of Statistical Science, University of Padova, Italy

2: Demetra Opinioni.net, Italy

**RELEVANCE & RESEARCH QUESTION** We compare empirical performance of different weighting techniques for bias correction in non-probability web panel samples on a prediction task: the outcome of the Italian political elections of September 2022. We contribute in answering the following questions: How can a practitioner use weighting methods to improve their estimation of a future event? When web panel data is collected, which correction technique is the most effective? Which one shows less variability among samples or employ methods?

**METHODS & DATA** We drawn five surveys (1, 2: web panel CAWI + CATI; 3, 4: web panel CAWI only; 5: social media; total,  $N = 17,857$ ; CATI,  $N = 549$ , social media,  $N = 1131$ ) to investigate voting behaviour in the weeks before the elections. Different weighting methods are used to compare our sample against the election results, used as ground truth. The cor-

rection techniques compared are: Raking, Inverse Probability Weighted Estimation (IPW), Multilevel Regression and Post-stratification (MRP), Doubly Robust Post-stratification (DRP) and Neural Network Modelling (NN).

**RESULTS** First, the sample composition is examined and compared to known population counts in multiple dimensions. The samples show systematic bias. For example, web surveys overrepresent populist parties and underrepresent left wing parties, while the opposite happens for telephone surveys. Secondly, we find that correction techniques vary in effectiveness depending on the variables chosen for correction, sample and target statistics.

We observe that 1) the effectiveness of all corrections methods correlate with the expectation of NN and that 2) MRP and DRP are the most reliable method, but suffer from variable misspecification, while NN does not while retaining good performance overall.

**ADDED VALUE** The contribution of this paper is two-fold. First, it presents differences in voting behaviour between the web samples and census, indicating how the web panel composition might be contributing to bias. Second, the present paper evaluates the predictive accuracy of web panels on election results using different correction methods. Other papers have performed similar analysis when a reference sample is available. To our knowledge, fewer researchers have performed such investigation in absence of a reference sample. We contribute to closing this knowledge gap with the present paper.

## ASSESSING SELF-SELECTION BIASES IN ONLINE SURVEYS: EVIDENCE FROM THE COVID-19 HEALTH BEHAVIOR SURVEY

**DONZOWA, JESSICA; PERROTTA, DANIELA; ZAGHENI, EMILIO**  
Max Planck Institute for Demographic Research, Germany

**RELEVANCE & RESEARCH QUESTION** During the COVID-19 pandemic, we have witnessed many primary survey collection efforts, especially relying on advertisements using social media recruitment. However, it remains unclear if surveys carried out this way may be biased due to the self-selection of participants depending on their interest in the study subject. In this study, we assess how the display of the survey topic in advertisements during recruitment affects responses.

**METHODS & DATA** We use the survey data collected via the “COVID-19 Health Behavior Survey” that ran between March and August 2020 in eight countries in Europe and North America (N = 120,036). Respondents’ recruitment took place via Facebook advertisements that differed in the image shown in the ad. We use regression analysis to investigate whether survey responses resulting from images with stronger (or weaker) relation to COVID-19 are associated with higher (or lower) compliance with recommended preventive behaviors, such as wearing a face mask, increasing hand hygiene, or perceived threat of the pandemic.

**RESULTS** In 18 out of the 32 models that we calculated, we observe no effect of the images on the outcomes. Where an effect was found its size was small. Overall, we found no clear association between higher topic salience and the adoption of protective behaviors. On the contrary, we see that the direction of the effect differs depending on the outcome and the country.

**ADDED VALUE** Our findings provide empirical evidence of the low impact of self-selection bias on survey responses, based on the images used for survey recruitment during the early stages of the pandemic. Thus, our results reduce concerns regarding one potential source of bias in unobserved characteristics.

## B6: ONLINE SEARCH DATA

**SESSION CHAIR: STEFAN KNAUFF**  
Bielefeld University, Germany

### WHAT DO RACISTS GOOGLE? A SYSTEMATIC APPROACH TO SELECTING AND VALIDATING GOOGLE TRENDS DATA TO STUDY XENOPHOBIC ATTITUDES

**MEHLTRETTER, JOHANNA (1); KEUSCH, FLORIAN (1); SAJONS, CHRISTOPH (2)**

1: University of Mannheim, Germany

2: Mannheim Centre of European Social Research, Germany

**RELEVANCE & RESEARCH QUESTION** Researchers increasingly use aggregate search data from Google Trends to measure attitudes. These data can be freely and immediately retrieved at much more fine-grained temporal and regional units than traditional survey data. Additionally, search data are expected to be less prone to social desirability bias as people feel anonymous when conducting a Google query. However, using Google Trends to measure attitudes raises concerns about validity as researchers cannot determine the intent behind a search. In this study, we examine how we can systematically select and validate keywords from Google Trends as an indicator for attitudes at the example of xenophobia.

**METHODS & DATA** We generated a list of search terms that potentially mirror xenophobia by extracting keywords from (1) survey items, (2) qualitative discourse analysis, and (3) operationalizations from previous studies. Next, we examined the validity of the search terms on this list by (1) testing their face validity using Google Trends’ related searches function, (2) analyzing their convergent validity by comparing the



trends of theoretically related keywords, [3] checking their criterion validity by correlating the Google Trends measure with both a well-established survey indicator of xenophobia and other external indicators such as voting behavior, and [4] studying their discriminant validity by comparing the results from step [3] to those from a set of random keywords.

**RESULTS** Based on the three steps of keyword selection, we retrieved 1748 keywords. Only 84 of these terms proved to have face validity according to step [1] of the validation approach. Another 33 keywords passed step [2] as they correlate with other keywords measuring the same construct. Four and three keywords respectively also correlated with a survey measure of xenophobia and right-wing votes proving [3] convergent validity. In the end, from our 1748 initial keywords, only one keyword, namely „refugees back“, passed all steps of the validation approach.

**ADDED VALUE** We provide a first systematic, stepwise approach to search term selection and validation of Google Trends as an indicator for attitudes. If researchers can complete these steps successfully, they can confidently use the keywords to measure their attitudes of interest.

## USING REGIONAL SHARE OF SEARCH TO BOOST BRAND GROWTH

**RECKMANN, TOBIAS**

TD Reply GmbH; Germany

**RELEVANCE & RESEARCH QUESTION** While many industry experts have acknowledged the virtues of Share of Search (SoS) as a reliable proxy for share of market, the metric remains rather under-appreciated in the marketing industry owing to the industry's continued skepticism towards digital data. As our case demonstrates, however, this skepticism is undeserved, and the usefulness of SoS goes far beyond being a proxy for market share. For a global dairy company looking to increase its category penetration in an Asian growth market, we leveraged the ability of SoS to account for regional differences down to city level. A brand growth tracking model was created that the client used to calibrate his marketing spend by prioritizing certain regions over others, leading to impressive results.

**METHODS & DATA** The brand growth tracking model is based on two key metrics, Category Health and Brand Momentum, combined in a matrix. The first key metric, Category Health, represents the category search interest per capita (net users) in percent for all defined regions during a given time period. Converted to a percentage, it reveals the exact state of category penetration per region. Brand Momentum on the other hand refers to the brand's share of category search in percent for all defined regions during a given time period. It is calculated by dividing the total number of searches for a brand by the total number of searches for all brands in the category – repeated for all regions.

**RESULTS** The matrix allowed our client to pinpoint exactly which region he should prioritize to increase his category penetration and to convert

the increased category interest into more sales. As a result, the client managed to realise a market share growth by 25% within six months by calibrating his marketing spend through regional prioritisation, without increasing his total spend.

**ADDED VALUE** This case demonstrates how the regional Share of Search approach can uncover enormous growth potential for brands that would otherwise likely remain hidden. It also raises an important question: How much more effective and less superfluous could advertising be by balancing out marketing activities according to regional requirements?

## C6: TRUST IN SCIENCE

(SPONSORED BY APROXIMA)

**SESSION CHAIR: ROLAND VERWIEBE**

University of Potsdam, Germany

## TRUST AND VACCINE HESITANCY IN EUROPE: UNVEILING THE CONNECTION THROUGH PANEL DATA ON COVID-19

**MASCHERINI, MASSIMILIANO; CONSOLINI, MICHELE;**

**NIVAKOSKI, SANNA**

EUROFOUND, Ireland

**RELEVANCE & RESEARCH QUESTION** The COVID-19 pandemic has led to a worldwide race to develop and administer vaccines. However, vaccine hesitancy has been identified as a significant barrier to achieving herd immunity. Vaccine hesitancy has been particularly pronounced, with significant variations among member states. This study aims to investigate the role of trust in institutions, interpersonal trust and trust in healthcare in vaccine hesitancy in the European Union, and how the COVID-19 vaccination campaign affected trust.

**METHODS & DATA** This study uses individual-level panel data collected by the Eurofound "Living, Working and COVID-19" e-survey in spring 2020/2021/2022. The sample includes more than 17,000 respondents from all EU member states for whom longitudinal data are available. We analyse the role of trust on vaccine hesitancy and whether individuals with low trust in institutions, interpersonal trust and trust in healthcare in 2020 were more likely to become vaccine-hesitant in 2021 and 2022. We also examine the impact of the COVID-19 vaccination campaign on trust and whether vaccine hesitancy was primarily caused by a pre-existing low level of trust or by the vaccination campaign itself.

**RESULTS** Our results show that individuals with low trust in institutions and in healthcare in 2020 were indeed more likely to become vaccine-hesitant in 2021 and 2022. However, the COVID-19 vaccination campaign also led to a significant decrease in trust among those vaccine hesitant who previously had a high level of trust in institutions. This suggests that the vaccination campaign acted as an aggregator of discontent, exacerbating pre-existing issues of low trust in institutions. Less clear is the role of interpersonal trust.

**ADDED VALUE** This study provides a unique perspective on vaccine hesitancy in the EU by using longitudinal data to examine the relationship between trust in institutions and vaccine hesitancy. To the best of our knowledge, this is the first study to investigate these questions using longitudinal data across all EU member states.

Our findings have important implications for policymakers and public health professionals seeking to promote vaccine acceptance and to increase trust in institutions in the context of the ongoing COVID-19 pandemic.

## PUBLIC TRUST IN SURVEY RESULT REPORTS: DOES METHODOLOGICAL INFORMATION ABOUT THE SELECTION PROCESS AND SURVEY MODE MATTER?

STADTMÜLLER, SVEN; SILBER, HENNING; TIROLF, NOAH  
GESIS - Leibniz Institute for the Social Sciences, Germany

**RELEVANCE & RESEARCH QUESTION** Online surveys and non-probabilistic methods for selecting and recruiting participants are playing an increasingly important role in survey practice. But are the results of such surveys less credible in the eyes of the public?

Based on a vignette experiment, the project investigates whether the reported selection process and survey mode affect the perceived trust in survey results distributed through the mass media. It also explores how many and which respondents take advantage of the opportunity to read additional information about the survey methodology.

**METHODS & DATA** We use data from a vignette experiment in which respondents were confronted with two short messages from a fictitious online magazine. Both messages report the results of a survey and provide basic information on the selection process and the survey mode.

We tested four combinations, namely (1) register sample and face-to-face survey (2) register sample and online survey with postal recruitment (3) telephone number sample and telephone survey (4) recruitment via websites of daily newspapers and online survey. In every second news report, more detailed information on the methodology of the survey was provided upon request. Respondents were recruited from a German commercial online access panel. The sample was quota-stratified by gender, age, and education level.

**RESULTS** For the full sample, perceived trust in the survey result did hardly differ among the four combinations of selection process and survey mode. However, if respondents were offered additional methodological information, this increased confidence in the survey result, especially for the probability surveys. Moreover, respondents who reported to have a profound prior knowledge of surveys were more likely to trust the probability surveys than the survey with self-recruitment.

**ADDED VALUE** Our research underlines the population's low level of competence in dealing with the quality of surveys and survey results. Given that the results of self-recruited surveys are increasingly reported in the media and at the same time have a great potential for bias, our findings are of high practical importance for public opinion research and political decision making.

## BEYOND LIKES: ENGAGING THE PUBLIC WITH MEDICAL CONTENT ON SOCIAL MEDIA IN SURGERY DEPARTMENTS

VASILIU, VLAD (1,3); SARAF, DANIEL (2); ARIEL, YARON (1)  
1: Academic College of Emek Yezreel, Israel  
2: The University of Alabama at Birmingham, USA  
3: Netanya Academic College, Israel

**RELEVANCE & RESEARCH QUESTION** The COVID-19 pandemic has exposed numerous issues in conveying reliable information from medical and scientific authorities to the public and the difficulty in controlling the spread of false and misleading information. This research examines whether content characteristics, particularly emotions, can increase public engagement with content published by medical professionals, specifically in the context of surgery departments. These departments often publish posts intended to inform and acknowledge, where the most effective form of engagement might not be the most engaging one but rather the one that helps create the most influential post by distributing it across platforms outside of Twitter. The study aims to understand the role of engagement in disseminating meaningful content and maximizing its reach and impact.

**METHODS & DATA** In this study, approximately 75,600 tweets were collected from physicians and specialists in 100 different surgery departments in the United States between 2018 and 2021. A weighted engagement index was calculated for each tweet based on the number of likes, comments, retweets (RT), and quotes. Concurrently, AI-based text analysis software (AI-NLP, VADER) assessed the emotional characteristics of each tweet (negative, neutral, and positive).

**RESULTS** The study found a negative correlation between several engagement indices and the use of negative emotions ( $r = -.01, -.009, -.013, p < .01$ ), but no clear positive correlation was found between engagement and the use of positive emotions ( $r = .003, .005, -.003, p > .05$ ). An interesting additional finding was that, while in many cases, „liking“ was found to be the most common engagement action on various social platforms, in this study, retweet (RT) action was found to be the leading engagement action, significantly surpassing other engagement actions.

**ADDED VALUE** This research highlights the importance of understanding emotions' role in engaging the public with medical content on social media platforms. In addition, the findings provide insights for medical professionals and communicators, emphasizing the need to consider emotional aspects when disseminating information. This could improve public understanding and response to critical health issues such as the COVID-19 pandemic.