



MARCH 31 - APRIL 2
BERLIN

CONNECTING PEOPLE:
INNOVATING MARKET
RESEARCH AND
SOCIAL DATA SCIENCE

THE 27TH GENERAL ONLINE RESEARCH CONFERENCE



WWW.GOR.DE



GOR
organized by DGOF

BELLA STRUMINSKAYA, SIMON KÜHNE, CARINA CORNESSE,
OTTO HELLWIG, STEFAN OGLESBY, YANNICK RIEDER (EDS.)

**27TH GENERAL ONLINE RESEARCH CONFERENCE PROCEEDINGS
BERLIN 2025**



IMPRESSUM:

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 978-3-9822985-7-3

BELLA STRUMINSKAYA, SIMON KÜHNE, CARINA CORNESSE, OTTO HELLWIG, STEFAN OGLESBY, YANNICK RIEDER (EDS.)

**GERMAN SOCIETY FOR ONLINE RESEARCH
DEUTSCHE GESELLSCHAFT FÜR ONLINE-FORSCHUNG (DGOF) E.V.
WWW.DGOF.DE**

EDITING: Anna Hristova, Anna Conrad **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 ORGANIZERS	
12	PROGRAMME OVERVIEW	
15	WORKSHOPS	
20	KEYNOTES	
22	GOR IMPACT AND INNOVATION AWARD	
26	DGOF BEST PAPER AWARD	
27	GOR POSTER AWARD	
36	GOR THESIS AWARD	
41	ABSTRACTS	
82	INVITED SESSIONS	



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



DR. CARINA CORNESSE
GESIS – Leibniz Institute for the Social
Sciences



PROF. DR. SIMON KÜHNE
Bielefeld University, Germany



PROF. DR. BELLA STRUMINSKAYA
Utrecht University, The Netherlands



DR. STEFAN OGLESBY
data IQ AG, Switzerland



YANNICK RIEDER
Janssen-Cilag GmbH, Germany

PROGRAMME COMMITTEE

PROF. DR. BELLA STRUMINSKAYA
(Programme Chair), DGOF Board & Utrecht University

PROF. DR. SIMON KÜHNE
(Vice Programme Chair), Bielefeld University

DR. CARINA CORNESSE
GESIS, Germany

DR. OTTO HELLWIG
DGOF Board & Bilendi & respondi

DR. STEFAN OGLESBY
DGOF Board & Data IQ AG

PROF. DR. FLORIAN KEUSCH
University of Mannheim

ZAZA ZINDEL
University of Bielefeld and
German Centre for Integration and Migration Research (DeZIM)

ANNA RYSINA
Verian

YANNICK RIEDER
(GOR Impact and Innovation Award 2025 competition), The Janssen
Pharmaceutical Companies of Johnson & Johnson

DR. OLAF WENZEL
(GOR Thesis Award 2025 competition), Wenzel Marktforschung

DR. FREDERIK FUNKE
(GOR Workshops), datenmethoden.de & LimeSurvey GmbH

CONFERENCE CHAIR

PROF. DR. BELLA STRUMINSKAYA
Programme Chair
DGOF Board & Utrecht University

PROF. DR. SIMON KÜHNE
Vice Programme Chair
DGOF Board & Bielefeld University

LOCAL PARTNER

Freie Universität Berlin

DGOF OFFICE

ANNA HRISTOVA, ANNA CONRAD

INTERNATIONAL BOARD

- DR. EVA AIZPURUA** National Centre for Social Research
- ALEXANDRA ASIMOV** GESIS – Leibniz Institute for the Social Sciences
- KATJA BIRKE** Produkt + Markt
- KAREN BLANKE** Destatis
- DR. CARINA CORNESSE** GESIS – Leibniz Institute for the Social Sciences
- PROF. MARCEL DAS** Centerdata
- PROF. EDITH DE LEEUW** Utrecht University
- LISA DUST** Kao Germany GmbH
- BEAT FISCHER** intervista AG
- DR. FREDERIK FUNKE** Dr. Funke – SPSS & R Trainings
- DR. GEORG-CHRISTOPH HAAS** Institute for Employment Research
- NIKLAS HAUPT** Mios GmbH
- FRANK HEUBLEIN** Frank Heublein M3 Services
- JULIA HOFFMANN** Heidelberg Materials AG
- PROF. ANNETTE HOXTELL** Fachhochschule Erfurt
- DR. WOJCIECH JABLONSKI**
Netherlands Interdisciplinary Demographic institute
- STEFAN KNAUFF** Bielefeld University
- DR. SEBASTIAN KOCAR** University College London
- FABIENNE KRAEMER** GESIS – Leibniz Institute for the Social Sciences
- DR. THOMAS KRÜGER** Umfragezentrum Bonn – Prof. Rudinger GmbH (uzbonn GmbH)
- DR. TANJA KUNZ** GESIS – Leibniz Institute for the Social Sciences
- ANDRÉ LANG** Insius
- PROF. BETTINA LANGFELDT** Kassel University
- ELLEN LAUPPER**
Swiss Federal Institute for Vocational Education and Training
- DANIIL LEBEDEV** GESIS – Leibniz Institute for the Social Sciences
- PROF. JI-PING LIN** Academia Sinica
- JAN MACKEBEN** Institut für Arbeitsmarkt- und Berufsforschung
- DR. OLGA MASLOVSKAYA** University of Southampton
- EMANUEL MAXL** Context-Research
- PROF. NATALJA MENOLD** Dresden University of Technology
- PROF. GUSTAVO S. MESCH** University of Haifa
- LONG NGUYEN** Bielefeld University
- MAREIKE OEHRL** september Strategie & Forschung GmbH
- DANIELLE REMMERSWAAL** Utrecht University & CBS
- DR. TOBIAS RETTIG** University of Mannheim
- FRANZISKA RIEDER** Omnicom Media Group Germany GmbH
- DR. EIKE MARK RINKE** University of Leeds
- DR. JOSS ROSSMANN** GESIS Leibniz Institute for the Social Sciences
- CAMILLA SALVATORE** Utrecht University
- GERO SERFAS** Q Agentur für Forschung
- DR. KAZUHIKO SHIBUYA** Alma Mater Europaea
- ROBIN SPICER** Bilendi GmbH
- DR. MARKUS STEINBRECHER**
Bundeswehr Center for Military History and Social Sciences
- PROF. BELLA STRUMINSKAYA** Utrecht University
- DR. TABEA TESCH** Royal Canin (Mars Inc.)
- DORIAN TSOLAK** Bielefeld University
- CHINTAN TURAKHIA** SSRS
- DR. ALEXANDER WENZ** University of Mannheim
- DR. OLAF WENZEL** Wenzel Marktforschung
- DR. GEORG WITTENBURG** Inspirient GmbH
- ANNE ZIMMER** RTL Deutschland GmbH
- ZAZA ZINDEL** Bielefeld University

DEAR GUESTS OF THE GOR 25!

CONNECTING PEOPLE

WELCOME TO THE 27TH EDITION OF THE GENERAL ONLINE RESEARCH CONFERENCE. WE ARE VERY HAPPY TO HOST THIS YEAR'S GOR IN COOPERATION WITH THE FREIE UNIVERSITÄT BERLIN.

This year, the GOR is returning to Berlin for the third time after 2017 and 2022. During our last visits, we were guests of the HTW Berlin, but this year we are pleased that the Freie Universität Berlin is opening its doors to us in the Henry Ford Building as our host.

Freie Universität Berlin (FU Berlin for short) was founded in 1948 as a university that could guarantee science and teaching freely and independently of political influence. The previous Berlin university, now Humboldt University, was located in the Soviet sector of the city and was under the control of the SED-dominated German Central Administration for People's Education and suffered from an increasing communist influence. The founding goals of the Free University of Berlin are also expressed in its Latin motto: Veritas – Iustitia – Libertas, the values of truth, justice and freedom. Since 2007, the FU Berlin has been the first Berlin university to be named a University of Excellence, and it regularly ranks among the world's top one hundred universities.

As in previous years, we have a great conference programme lined up for you including keynotes, discussions, presentations, awards, posters, workshops and much more. After a good discussion, we decided to dissolve the previous track structure of the GOR and to structure the presentations across thematic categories into sessions. In doing so, we want to break down the silos and network people even more closely. This is entirely in line with our new claim 'Connecting People: Innovating Market Research and Social Data Science'. You can choose between four parallel sessions and one invited session. In addition, there are three award competitions: i) the GOR Impact & Innovation Award 2025 for the most impactful market research case, ii) the GOR Thesis Award 2025 for the best thesis in online research, iii) the GOR Poster Award 2025 for the best poster of the conference. The DGOF Best Paper Award 2025 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 Tuesday, 1 April, Dr. Stephanie Eckman, Maryland University will talk about "Insights from Survey Methodology can Improve AI Models" and on Wednesday, 2 April, Prof. Dr. Brent Mittelstadt, Professor of Data Ethics and Policy and Director of Research at the Oxford Internet Institute (OII), University of Oxford (UK) will talk about "Do large language models have a duty to tell the truth?".

On the pre-conference day, Monday 31 March, five workshops will take place and on Monday evening we will meet for the GOR 25 Get Together with drinks and snacks. Tuesday evening is GOR party time.

We are particularly grateful for the enthusiastic support of and collaboration with our partners at Free University of Berlin: Prof. Dr. Stefan Liebig (Empirische Sozialstrukturanalyse, FU Berlin) and Dr. Carina Cornesse GESIS – Leibniz Institute for the Social Sciences. 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 2025!

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 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. SIMON KÜHNE is Professor of Applied Social Data Science at the Faculty of Sociology at Bielefeld University. His methodological research is focused on 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 and he is the Vice Programme Chair of the GOR 2025 Conference.

PROF. DR. BELLA STRUMINSKAYA is associate professor at the department of Methodology and Statistics at Utrecht University, The Netherlands, and a researcher at Statistics Netherlands. Her research focuses on innovations in data collection such as using apps, wearables, and sensors in surveys and official statistics, as well as data donation of digital trace data.

She is a country team lead and member of the Management Board of Survey of Health, Ageing and Retirement in Europe (SHARE), member of Methods Advisory Board of the European Social Survey (ESS) and member of the Scientific Advisory Board of Statistics Sweden. 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 25 conference.





DR. CARINA CORNESSE is Head of The Department Survey Design and Methodology at GESIS – Leibniz Institute for the Social Sciences. Previous to her current role, she was a guest professor of Social Stratification and Survey Methodology at Free University (FU) Berlin and held different positions at the German Institute for Economic Research, University of Bremen, and University of Mannheim.

Her research focuses on survey methodology, digital inequality, and online research. She is particularly interested in innovative methods for recruiting and surveying individuals and households as well as achieving inclusiveness and diversity in survey samples given unequal digital opportunities, habits, and preferences. She is a member of the DGOF Board as well as the chair of the DGOF Research Funding committee since 2023 and local host of GOR 2025.

DR. STEFAN OGLESBY, MBA 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.



YANNICK RIEDER is Emea Market Research Manager at Johnson & Johnson Innovative Medicine. His passion is implementing innovative approaches to gain in-depth customer and market insights and translating them into actionable strategies.

He has successfully done this for over 10 years on both the agency and corporate sides. Yannick is a DGOF Board Member since 2023

**DR. CARINA CORNESSE
AND PROF. DR. STEFAN LIEBIG
LOCAL HOSTS OF GOR 25
FREIE UNIVERSITÄT BERLIN**

DEAR PARTICIPANTS OF GOR 25,

It is our distinct pleasure to welcome you to the General Online Research (GOR) Conference 2025 at the Institute of Sociology at Freie Universität Berlin (FU Berlin). As one of Germany's leading research institutions, FU Berlin is a hub of innovation and interdisciplinary exploration, making it an ideal setting for advancing the frontiers of online research and its many applications. FU Berlin, established in 1948, has a long-standing reputation for fostering cutting-edge research and intellectual growth. Its global outlook and strong emphasis on collaboration mirror the core goals of GOR: connecting researchers and practitioners to share insights, debate methodologies, and push the boundaries of what is possible in the digital age.

In recent years, the Institute of Sociology has enhanced its emphasis on comparative and empirical research by engaging in outstanding national and international interdisciplinary research collaborations. Moreover, it has contributed to this focus through its English-language Master's program, European Societies, and its Bachelor's program, Global Societies. We are sure the exchange between the Institute and DGOF/GOR will be very productive.

The conference will be held in the iconic Henry Ford Building, one of FU Berlin's architectural landmarks. This historic venue is not only renowned for its distinctive design and historical significance but also for its state-of-the-art facilities that support dynamic academic exchange. With its spacious auditoriums and modern amenities, the Henry Ford Building sets the stage for engaging presentations, interactive workshops, and fruitful networking opportunities.

One of the defining features of GOR is its commitment to bringing together academic researchers and online research practitioners. This interdisciplinary exchange is crucial for translating theoretical advancements into practical applications and ensuring that empirical research benefits from real-world insights. For instance, web survey methodology has significantly evolved in recent years, adapting to the widespread use of smartphones and incorporating digital behav-



Dr. Carina Cornesse
GESIS – Leibniz
Institute for the
Social Sciences

Prof. Dr. Stefan Liebig
Institut für Soziologie
Empirische Sozial-
strukturanalyse

Foto:
berlin-event-foto.de/
Peter-Paul



ioral data to achieve more accurate and nuanced measurements. This confluence of academic rigor and industry innovation drives the development of new tools, techniques, and frameworks for understanding online behavior.

Since 1998, GOR is the place to showcase and shape these advancements in Germany. The conference highlights the expanding frontiers of online research, including the transformative possibilities of artificial intelligence and the challenges it brings. From understanding user behavior in digital environments to leveraging AI for societal benefits, GOR 25 will delve into the latest advancements, methodologies, and ethical considerations shaping our field.

In addition to thought-provoking sessions and panels, we encourage you to utilize breaks and networking events to exchange ideas and perspectives with colleagues. These interactions are key to exploring new research directions and forging interdisciplinary collaborations. Thank you for being part of GOR 25. We look forward to exploring new horizons together and driving innovation at the intersection of research and practice.

BEST REGARDS,

Dr. Carina Cornesse and Prof. Dr. Stefan Liebig

ORGANIZERS



DGOF

Freie Universität



Berlin

SPONSORS



annalect

**Bilendi
& respondi**



ERGO-Data GmbH®



EUROPEAN SURVEY RESEARCH ASSOCIATION



GESIS Leibniz Institute
for the Social Sciences



GIM | RELEVANCE COUNTS.

horizoom
people first.

inspirient
cognitive analytics

infas

norstat

produkt+markt
marketing research

reportbook

talk
ONLINE PANEL

xelper

MEDIA PARTNER

planung&analyse

marktforschung



COOPERATION PARTNER

succeet

insights | data | analytics

2026

18 – 19 March, 2026 @ RMCC
Wiesbaden/Frankfurt, Germany



w3m

19–27 May 2025

by **succeet**
meet & succeed

DATE: MONDAY, 31 MARCH 2025

- 09:00 - 10:00 AM** **BEGIN CHECK-IN** / Foyer EG
- 10:00 - 01:00 PM** **WORKSHOP 1** / Konferenzraum II
Web tracking - augmenting web surveys with app, URL, and search term data
oshua Claaßen, Jan Karem Höhne, DZHW, Leibniz University Hannover, Germany
- 10:00 - 01:00 PM** **WORKSHOP 2** / Konferenzraum III
Structured information extraction with LLMs
Paul Ferdinand Simmering, Q Agentur für Forschung GmbH, Germany
- 01:00 - 01:30 PM** **BREAK**
- 01:30 - 04:30 PM** **WORKSHOP 3** / Konferenzraum IIII
Working with research partners - strategies for effectively collaborating on a research project with external agencies
Mario Callegaro, Callegaro Research, United Kingdom and Monica Sanchez, Untold Research, United States
- 01:30 - 04:30 PM** **WORKSHOP 4** / Konferenzraum II
Collecting and analyzing smartphone (survey) data using the GESIS AppKit
Julian Kohne, Mareike Wieland, Lukas Rahnke-Otto, Maximilian Haag
GESIS – Leibniz Institute for the Social Sciences, Germany (all)
- 01:30 - 04:30 PM** **WORKSHOP 5 ONLINE**
Making inference from nonprobability online surveys
Camilla Salvatore, Utrecht University, The Netherlands
- 04:30 - 05:30 PM** **BREAK**
- 05:30 - 07:00 PM** **DGOF MEMBERS GENERAL MEETING** / Jung&Schönn
- 07:00 - 07:30 PM** **BREAK**
- 07:30 - 08:30 PM** **EARLY CAREER SPEED NETWORKING EVENT** / Jung&Schönn
- 08:30 - 11:59 PM** **GOR 25 GET TOGETHER** / Location: Jung&Schönn

DATE: TUESDAY, 01 APRIL 2025

- 08:00 - 09:00 AM** BEGIN CHECK-IN / Foyer EG
- 09:00 - 10:15 AM** 1. GOR 25 OPENING & KEYNOTE 1 / Max-Kade-Auditorium
High Quality Training Data for AI Models: Lessons from 20 years in Surveys,
by Dr. Stephanie Eckman
- 10:15 - 10:45 AM** BREAK
- 10:45 - 11:45 AM** 2.1: INNOVATIONS IN MARKET RESEARCH / Hörsaal A
- 10:45 - 11:45 AM** 2.2: WEB TRACKING / Hörsaal B
- 10:45 - 11:45 AM** 2.3: VOTING BEHAVIOR AND INFORMATION SOURCES / Hörsaal C
- 10:45 - 11:45 AM** GOR THESIS AWARD I: BACHELOR/MASTER / Hörsaal D
- 10:45 - 11:45 AM** INVITED SESSION I: SESSION BEI MARKFORSCHUNG.DE / Max- Kade- Auditorium
- 11:45 - 12:00 PM** BREAK
- 12:00 - 01:15 PM** 3.1: VIRTUAL INTERVIEWING / Hörsaal A
- 12:00 - 01:15 PM** 3.2: RESPONDENT ENGAGEMENT AND ATTRITION / Hörsaal B
- 12:00 - 01:15 PM** 3.3: DATA DONATION / Hörsaal C
- 12:00 - 01:15 PM** 3.4: DIGITAL BATTLEFIELDS / Hörsaal D
- 12:00 - 01:15 PM** GOR THESIS AWARD II: PHD / Max-Kade-Auditorium
- 01:15 - 02:30 PM** LUNCH-BREAK
- 02:30 - 03:30 PM** 4.1: POSTER SESSION / Foyer OG
- 02:30 - 03:30 PM** 4.2: POSTER SESSION / Foyer OG
- 02:30 - 03:30 PM** 4.3: POSTER SESSION / Foyer OG
- 02:30 - 03:30 PM** 4.4: POSTER SESSION / Foyer OG
- 02:30 - 03:30 PM** 4.5: POSTER SESSION / Foyer OG
- 03:30 - 03:45 PM** BREAK
- 03:45 - 04:45 PM** 5.1: INNOVATIONS IN SAMPLING / Hörsaal A
- 03:45 - 04:45 PM** 5.2: RESPONDENT NUDGING AND INCENTIVES / Hörsaal B
- 03:45 - 04:45 PM** 5.3: LLMS AND SYNTHETIC SURVEY DATA / Hörsaal C
- 03:45 - 04:45 PM** 5.4: SOCIAL MEDIA INFLUENCE IN WAR TIMES / Hörsaal D
- 03:45 - 04:45 PM** I: IMPACT & INNOVATION AWARD I / Max-Kade-Auditorium
- 04:45 - 05:00 PM** BREAK
- 05:00 - 06:00 PM** 6.1: INCREASING SURVEY RESPONSE / Hörsaal A
- 05:00 - 06:00 PM** 6.2: MIXED MODE AND MODE TRANSITIONS / Hörsaal B
- 05:00 - 06:00 PM** 6.3: SURVEY INNOVATIONS / Hörsaal C
- 05:00 - 06:00 PM** 6.4: LLMS AS ANALYSIS TOOLS / Hörsaal D
- 05:00 - 06:00 PM** II: IMPACT & INNOVATION AWARD II / Max-Kade-Auditorium
- 08:30 - 11:59 PM** GOR 25 PARTY / Beate Uwe

DATE: WEDNESDAY, 02 APRIL 2025

- 08:30 - 09:00 AM** BEGIN CHECK-IN / Foyer EG
- 09:00 - 10:00 AM** 7.1: SURVEY RECRUITMENT / Hörsaal A
- 09:00 - 10:00 AM** 7.2: DIGITAL BEHAVIOR AND DIGITAL TRACES / Hörsaal B
- 09:00 - 10:00 AM** 7.3: QUESTIONNAIRE DESIGN / Hörsaal C
- 09:00 - 10:00 AM** INVITED SESSION II: INNOVATION IN PRACTICE: SMART SURVEY TECHNIQUES / Hörsaal D
- 10:00 - 10:30 AM** BREAK
- 10:30 - 11:15 AM** 8: KEYNOTE 2 / Max-Kade-Auditorium
Do large language models have a duty to tell the truth?,
by Prof. Brent Mittelstadt, University of Oxford, United Kingdom
- 11:15 - 11:45 AM** 9: GOR AWARD CEREMONY / Max-Kade-Auditorium
- 11:45 - 12:00 PM** BREAK
- 12:00 - 01:15 AM** 10.1: RELUCTANT RESPONDENTS AND ITEM NONRESPONSE / Hörsaal A
- 12:00 - 01:15 AM** 10.2: AI AND AUTOMATION IN (SURVEY) / Hörsaal B
- 12:00 - 01:15 AM** 10.3: EXPLORING REPRESENTATION IN SOCIAL MEDIA / Hörsaal C
- 12:00 - 01:15 AM** INVITED SESSION III: DGOF KI (AI) FORUM: INSPIRATION SESSION (SESSION HELD IN ENGLISH)
OPPORTUNITIES & CHALLENGES IN APPLYING AI TO MARKET RESEARCH / Location: Hörsaal D
- 01:15 - 02:30 PM** LUNCH-BREAK
- 02:30 - 03:45 PM** 11.1: NONRESPONSE BIAS AND CORRECTION / Hörsaal A
- 02:30 - 03:45 PM** 11.2: BOTS, AVATARS, AND ONLINE LABS / Hörsaal B
- 02:30 - 03:45 PM** 11.3: SOCIAL MEDIA SURVEYS AND RECRUITMENT / Hörsaal C
- 02:30 - 03:45 PM** INVITED SESSION IV: DGOF KI (AI) FORUM: WORLD CAFÉ (SESSION HELD IN GERMAN)
OPPORTUNITIES & CHALLENGES IN APPLYING AI TO MARKET RESEARCH / Hörsaal D
- 03:45 - 04:00 PM** BREAK
- 04:00 - 05:00 PM** 12.1: REDUCING NONRESPONSE / Hörsaal A
- 04:00 - 05:00 PM** 12.2: APP-BASED DIARY STUDIES / Hörsaal B
- 04:00 - 05:00 PM** 12.3: REMINDERS AND TARGETING / Hörsaal C

MONDAY, 31 MARCH 2025
10:00 - 01:00 PM
KONFERENZRAUM II

WORKSHOP 1

PRESENTATIONS

WEB TRACKING – AUGMENTING WEB SURVEYS WITH APP, URL AND SEARCH TERM DATA

JOSHUA CLAASSEN, JAN KAREM HÖHNE

DZHW, Leibniz University Hannover, Germany

WORKSHOPSLOT 2,5h

TARGET GROUPS Researchers and practitioners with a general methodological interest in web-based surveys and digital trace data

IS THE WORKSHOP GEARED AT AN EXCLUSIVELY GERMAN OR AN INTERNATIONAL AUDIENCE? International

WORKSHOP LANGUAGE English

DESCRIPTION OF THE CONTENT OF THE WORKSHOP Web surveys frequently run short to accurately measure digital behavior because they are prone to recall error (i.e., biased recalling and reporting of past behavior) and social desirability bias (i.e., misreporting of behavior to comply with social norms and values). New advances in the collection of digital trace (or web tracking) data make it possible to directly measure digital behavior in the form of browser logs (e.g., visited URLs and search terms) and apps (e.g., duration and frequency of their use). Building on these advances, we will introduce participants to web surveys augmented with web tracking data. In this course, we initially give a thorough overview of the manifold new measurement opportunities provided by web tracking.

In addition, participants obtain comprehensive insights into the collection, processing, analysis, and error sources of web tracking data as well as its application to substantive research. Importantly, the course includes applied web tracking data exercises in which participants learn how to ...

... operationalize and collect web tracking data.

... work with and process web tracking data.

... analyze and extract information from web tracking data.

GOALS OF THE WORKSHOP At the end of the workshop, participants will be able to 1) independently conceptualize the collection of web tracking data, 2) decide on best practices when it comes to data handling and analysis, and 3) critically reflect upon the opportunities and challenges of web tracking data and its suitability for empirical studies in the context of social and behavioral science.

NECESSARY PRIOR KNOWLEDGE OF PARTICIPANTS Basic knowledge on web-based surveys, including structured and unstructured datasets, is beneficial but not a pre-requisite.

LITERATURE THAT PARTICIPANTS NEED TO READ PRIOR TO PARTICIPATION none

RECOMMENDED ADDITIONAL LITERATURE Bosch, O. J., & Revilla, M. (2022). When survey science met web tracking: Presenting an error framework for metered data. *Journal of the Royal Statistical Society (Series A)*, 185, 408-436. <https://doi.org/10.1111/rssa.12956>

Bach, R. L., & Wenz, A. (2020). Studying health-related internet and mobile device use using web logs and smartphone records. *PLoS ONE*, 15(6), Article e0234663. <https://doi.org/10.1371/journal.pone.0234663>

Cernat, A., Keusch, F., Bach, R. L., & Pankowska [2024]. Estimating measurement quality in digital trace data and surveys using the MultiTrait MultiMethod Model. *Social Science Computer Review*. <https://doi.org/10.1177/08944393241254464>

Revilla, M. (2022). How to enhance web survey data using metered, geolocation, visual and voice data? *Survey Research Methods*, 16(1). <https://doi.org/10.18148/srm/2022.v16i1.8013>

INFORMATION ABOUT THE INSTRUCTORS

JAN KAREM HÖHNE is junior professor at Leibniz University Hannover in association with the German Center for Higher Education Research and Science Studies (DZHW). He is head of the CS3 Lab for Computational Survey and Social Science. His research focuses on new data forms and types for measuring political and social attitudes.

JOSHUA CLAASSEN is PhD student and research associate at Leibniz University Hannover in association with the German Center for Higher Education Research and Science Studies (DZHW). His research focuses on computational survey and social science with an emphasis on digital trace data.

MAXIMUM NUMBER OF PARTICIPANTS 25

WILL PARTICIPANTS NEED TO BRING THEIR OWN DEVICES IN ORDER TO BE ABLE TO ACCESS THE INTERNET? WILL THEY NEED TO BRING ANYTHING ELSE TO THE WORKSHOP? Yes, they should bring their devices (laptop and smartphone).

MONDAY, 31 MARCH 2025
10:00 - 01:00 PM
KONFERENZRAUM III

WORKSHOP 2

PRESENTATIONS

STRUCTURED INFORMATION EXTRACTION WITH LLMs

PAUL FERDINAND SIMMERING

Q Agentur für Forschung GmbH, Germany

DURATION OF THE WORKSHOP 2,5h

TARGET GROUPS Analysts and researchers working with text data, e.g. transcripts, news articles, social media posts or reviews

IS THE WORKSHOP GEARED AT AN EXCLUSIVELY GERMAN OR AN INTERNATIONAL AUDIENCE? International

WORKSHOP LANGUAGE English

DESCRIPTION OF THE CONTENT OF THE WORKSHOP This workshop is an introduction to the application of Large Language Models (LLMs) for structured information extraction in market research and social sciences. Participants will implement solutions to natural language processing tasks such as text classification, entity recognition, and sentiment analysis. The session includes hands-on exercises in Python using the library “instructor”.

Participants will learn about strategies for prompting, few-shot examples and fine-tuning. The approaches taught are compatible with a wide range of open source and commercial models. Discussion sections of the workshop will cover the methodological and technical possibilities and limitations of LLMs for information extraction.

GOALS OF THE WORKSHOP

- Get hands-on experience with structured information extraction.
- Get an overview of available models, tools and prompting tactics
- Learn about evaluation, efficiency and limitations
- Share experiences and use cases

NECESSARY PRIOR KNOWLEDGE OF PARTICIPANTS Basic knowledge of Python. R users can use the guide recommended literature to get up to speed quickly. The code examples in the workshop can be followed with minimal coding knowledge, extending them requires a bit more.

LITERATURE THAT PARTICIPANTS NEED TO READ PRIOR TO PARTICIPATION

Starter guide which will be sent before the workshop. It will contain instructions for using Google Colab and installing the required Python packages.

RECOMMENDED ADDITIONAL LITERATURE Primer on Python for R users: https://rstudio.github.io/reticulate/articles/python_primer.html

INFORMATION ABOUT THE INSTRUCTORS

PAUL SIMMERING is a data scientist at Q Agentur für Forschung where he works on social media and review analysis. He has presented research on sentiment analysis at GOR 23 and GOR 24

MAXIMUM NUMBER OF PARTICIPANTS 20

WILL PARTICIPANTS NEED TO BRING THEIR OWN DEVICES IN ORDER TO BE ABLE TO ACCESS THE INTERNET? WILL THEY NEED TO BRING ANYTHING ELSE TO THE WORKSHOP? Participants will need to bring a laptop. An OpenAI API key will be provided for use during the workshop. The recommended development environment for beginners is Google Colab, which is free and runs in the browser. A starter guide will be provided. Advanced users are welcome to use an IDE of their choice and are also welcome to use a different LLM platform than OpenAI that is compatible with instructor, such as Anthropic, Cohere, Gemini and local models using Ollama.

MONDAY, 31 MARCH 2025
01:30 - 04:30 PM
KONFERENZRAUM III

WORKSHOP 3

PRESENTATIONS

WORKING WITH RESEARCH PARTNERS – STRATEGIES FOR EFFECTIVELY COLLABORATING ON A RESEARCH PROJECT WITH EXTERNAL AGENCIES

MARIO CALLEGARO

Callegaro Research, United Kingdom

MONICA SANCHEZ

Untold Research, United States

DURATION OF THE WORKSHOP 2,5h

TARGET GROUPS Researchers who have worked or are looking into working with third parties companies for market/survey research projects and want to improve the collaboration and the quality of the project outcome

IS THE WORKSHOP GEARED AT AN EXCLUSIVELY GERMAN OR AN INTERNATIONAL AUDIENCE? International

WORKSHOP LANGUAGE English

DESCRIPTION OF THE CONTENT OF THE WORKSHOP Researchers in both commercial and academic sectors are increasingly relying on research partners for the entire process of market/survey research projects. The goal of this course is to describe good practices in working with external agencies and managing a survey research project with a third party vendor.

This first part of the course will cover a broad spectrum of research management topics, including: determining the project scope and objectives, proposal preparation including requests for proposals (RFP), statements of work (SOW) and how to evaluate and select a vendor, how to best communicate with vendors, setting up timelines and schedules, types of contracts and subcontracting, budgeting, monitoring and quality controls, data security, research ethics, and deliverables and reporting.

The second part of the course will discuss strategies of effective collaboration and how to best work together on a project such as assigning and managing tasks, managing time and managing documents/files and documenting the project. We will also discuss strategies for communicating with the research agency.

We will close the workshop looking at trends in the vendor industry and the use of AI from the vendor side.

GOALS OF THE WORKSHOP

- Dealing with the different stages of negotiations and documents needed when working with third parties. Streamline the partnership with third parties in order to obtain high quality data
- Effectively managing the workload and the roles of the internal team versus the research partner

NECESSARY PRIOR KNOWLEDGE OF PARTICIPANTS No necessary previous knowledge.

LITERATURE THAT PARTICIPANTS NEED TO READ PRIOR TO PARTICIPATION

There is basically no literature we are aware of at this time, we might uncover something during the preparation of the final version of the slides.

INFORMATION ABOUT THE INSTRUCTORS

MARIO CALLEGARO has 15 years of experience at Google having worked with large research international agencies on numerous million dollar projects. Before that, he worked on the agency side at Knowledge Panel (now Ipsos Knowledge Panel).

Mario holds a M.S. and a Ph.D. in Survey Research and Methodology from the University of Nebraska, Lincoln. Mario has published numerous papers on survey research methods, book chapters, and two books, one on online panels and one on web surveys.
<https://www.linkedin.com/in/mcallegaro/>

MONICA SANCHEZ brings over 15 years of experience in qualitative and quantitative research, leading national and global studies. From ideation and study design to data collection, analysis, and reporting, she ensures top-quality research through up-to-date methodologies and strict quality controls. Monica has presented findings across the U.S., Latin America, and Europe in both English and Spanish. Her expertise spans various industries, including consumer goods, technology, and non-profit/government sectors. She holds a Master's in Survey Research and Methodology and a Bachelor's in Journalism from the University of Nebraska - Lincoln.
<https://www.untoldresearch.com/>

MAXIMUM NUMBER OF PARTICIPANTS 35

WILL PARTICIPANTS NEED TO BRING THEIR OWN DEVICES IN ORDER TO BE ABLE TO ACCESS THE INTERNET? WILL THEY NEED TO BRING ANYTHING ELSE TO THE WORKSHOP? No need to bring a laptop

MONDAY, 31 MARCH 2025
01:30 - 04:30 PM
KONFERENZRAUM II

WORKSHOP 4

PRESENTATIONS

COLLECTING AND ANALYZING SMARTPHONE (SURVEY) DATA USING THE GESIS APPKIT

JULIAN KOHNE, MAREIKE WIELAND, LUKAS RAHNKE-OTTO, MAXIMILIAN HAAG

GESIS - Leibniz Institute for the Social Sciences, Germany [all]

DURATION OF THE WORKSHOP 2,5h

TARGET GROUPS Social Scientists including, but not limited to, Media & Communication Scientists, Psychologists, Quantitative Sociologists, and Political Scientists

IS THE WORKSHOP GEARED AT AN EXCLUSIVELY GERMAN OR AN INTERNATIONAL AUDIENCE? International

WORKSHOP LANGUAGE English

DESCRIPTION OF THE CONTENT OF THE WORKSHOP Smartphones are a ubiquitous technology that most people have in their pockets almost everywhere and anywhere. This makes them an excellent choice to collect surveys- and other data such as photos, screenshots, app usage, or sensor data from research participants in everyday life. They are especially suitable for intensive longitudinal designs where constructs are typically measured multiple times per day or per week to investigate fine-grained temporal patterns or changes in behavior, opinions, or attitudes. This workshop will introduce the design of intensive longitudinal studies both on a conceptual level and in practice by using the GESIS AppKit, our new infrastructure for easily creating smartphone data collections for Android and iOS. Participants will interactively create and try out their own data collection studies, and learn how to preprocess, analyze, and visualize the resulting data.

GOALS OF THE WORKSHOP

- 1.** Introduction to intensive longitudinal data collections, how to design these studies, and what differentiates them from “classic” longitudinal data collections with fewer repeated measures. [45 mins – Lukas Rahnke-Otto / Mareike Wieland]
- 2.** Create your own free GESIS AppKit account, learn how to create your own data collections using intensive longitudinal surveys, photo and image uploads, how to link your data to other survey tools, how to invite people to your study, and how to schedule intensive longitudinal data collections. [45 mins – Lukas Rahnke-Otto / Mareike Wieland]

- 3.** Learn how to export your data from the AppKit website and how to preprocess and visualize it using R. [45 mins – Julian Kohne / Maximilian Haag]

NECESSARY PRIOR KNOWLEDGE OF PARTICIPANTS Basic Knowledge of R and Rstudio is recommended [changing working directories, importing data, using ggplot2].

LITERATURE THAT PARTICIPANTS NEED TO READ PRIOR TO PARTICIPATION none

RECOMMENDED ADDITIONAL LITERATURE https://gesiscss.github.io/AppKit_Documentation/ [not required]

INFORMATION ABOUT THE INSTRUCTORS

<https://www.gesis.org/institut/ueber-uns/mitarbeitendenverzeichnis/person/JULIAN.KOHNE>

<https://www.gesis.org/institut/ueber-uns/mitarbeitendenverzeichnis/person/MAREIKE.WIELAND>

<https://www.gesis.org/institut/ueber-uns/mitarbeitendenverzeichnis/person/MAXIMILIAN.HAAG>

<https://www.gesis.org/institut/ueber-uns/mitarbeitendenverzeichnis/person/LUKAS.OTTO>

MAXIMUM NUMBER OF PARTICIPANTS 40

WILL PARTICIPANTS NEED TO BRING THEIR OWN DEVICES IN ORDER TO BE ABLE TO ACCESS THE INTERNET? WILL THEY NEED TO BRING ANYTHING ELSE TO THE WORKSHOP?

- 1.** Laptop with R and RStudio installed
- 2.** Smartphone running at least either Android 16 or iOS 11
- 3.** All participants will need access to the internet on both Laptop and Smartphone

MONDAY, 31 MARCH 2025
01:30 - 04:30 PM
ONLINE

WORKSHOP 5

PRESENTATIONS

MAKING INFERENCE FROM NONPROBABILITY ONLINE SURVEYS

CAMILLA SALVATORE

Utrecht University, Netherlands, The

DURATION OF THE WORKSHOP 2,5h

TARGET GROUPS This workshop is intended for students, researchers, and practitioners who work with nonprobability sample surveys and seek to assess data quality and explore methods for drawing inferences.

IS THE WORKSHOP GEARED AT AN EXCLUSIVELY GERMAN OR AN INTERNATIONAL AUDIENCE? International

WORKSHOP LANGUAGE English

DESCRIPTION OF THE CONTENT OF THE WORKSHOP Academic and market researchers often rely on opt-in (volunteer) online panels to collect survey data. These panels allow for the rapid collection of responses, resulting in large datasets at a relatively low cost. While this approach is convenient and cost-effective, it has a major limitation: making inferences about the target population is not straightforward. The issue is that the sample is non-probabilistic; the panel consists of volunteers who self-select themselves, introducing selection bias. Addressing this bias requires statistical adjustments under strong assumptions.

In this workshop, we will discuss methods for making inference using nonprobability samples, focusing on how to:

- Identify challenges related to data quality and errors
- Design research projects based on nonprobability sample surveys
- Select appropriate methodologies to address selection bias
- Acknowledge the limitations of these methodologies

The workshop will include case studies from existing literature and practical exercises in R. Participants do not need prior knowledge of R. An online interactive page will be provided with explanations, code and results.

GOALS OF THE WORKSHOP

- Identify challenges related to data quality and errors
- Design research projects based on nonprobability sample surveys
- Select appropriate methodologies to address selection bias
- Acknowledge the limitations of these methodologies

NECESSARY PRIOR KNOWLEDGE OF PARTICIPANTS Participants should have basic pre-existing knowledge about survey methodology and statistics.

LITERATURE THAT PARTICIPANTS NEED TO READ PRIOR TO PARTICIPATION Participants can watch this video by the Pew Reserch Center about nonprobability surveys: <https://www.pewresearch.org/short-reads/2018/08/06/what-are-nonprobability-surveys/>

RECOMMENDED ADDITIONAL LITERATURE

- Mercer, A., Lau, A., & Kennedy, C. (2018). For weighting online opt-in samples, what matters most? Pew Research Center report. <https://www.pewresearch.org/methods/2018/01/26/for-weighting-online-opt-in-samples-what-matters-most/>
 - Salvatore, C. (2023). Inference with non-probability samples and survey data integration: a science mapping study. *Metron*, 81(1), 83-107. <https://link.springer.com/article/10.1007/s40300-023-00243-6>
 - Mercer, A. W., Kreuter, F., Keeter, S., & Stuart, E. A. (2017). Theory and practice in nonprobability surveys: parallels between causal inference and survey inference. *Public Opinion Quarterly*, 81. <https://academic.oup.com/poq/article/81/S1/250/3749176>
- Additional references will be available in the slides

INFORMATION ABOUT THE INSTRUCTORS

DR. CAMILLA SALVATORE works as an assistant professor at the department of Methodology and Statistics at Utrecht University, where she specializes in survey research. Her interests include inference with nonprobability samples, survey weighting, nonresponse, the use of digital trace data and their integration with surveys. <https://www.uu.nl/medewerkers/CSalvatore>

WILL PARTICIPANTS NEED TO BRING THEIR OWN DEVICES IN ORDER TO BE ABLE TO ACCESS THE INTERNET? Online Workshop

DATE: TUESDAY, 01 APRIL 2025

LOCATION: MAX-KADE-AUDITORIUM

09:00 - 10:15 AM



Stephanie Eckman
Researcher and Data
Scientist at University of
Maryland, USA

SESSION 1:

HIGH QUALITY TRAINING DATA FOR AI MODELS: LESSONS FROM 20 YEARS IN SURVEYS

CV: Stephanie Eckman has a PhD in statistics and methodology from the University of Maryland. She has worked in government, non-profits, academia, and technology. She has conducted surveys and taught statistics around the world.

ABSTRACT:

Whether future AI models are fair, trustworthy, and aligned with the public's interests rests in part on our ability to collect accurate data about what we want the models to do. However, collecting high-quality data is difficult, and few AI/ML researchers are trained in data collection methods. The talk bridges artificial intelligence and survey methodology, demonstrating how techniques from survey research can improve training data quality and model performance. The talk concludes with practical recommendations for improving training data collection and ideas for joint research.

DATE: WEDNESDAY, 02 APRIL 2025

LOCATION: MAX-KADE-AUDITORIUM

10.30 - 11:15 AM



Prof. Brent Mittelstadt
Professor of Data Ethics and Policy at the Oxford Internet Institute (OII) at the University of Oxford, and the OII's Director of Research

SESSION 8:

KEYNOTE 2:

"DO LARGE LANGUAGE MODELS HAVE A DUTY TO TELL THE TRUTH?"

Professor Brent Mittelstadt is Professor of Data Ethics and Policy at the Oxford Internet Institute (OII) at the University of Oxford, and the OII's Director of Research. He also coordinates of the Governance of Emerging Technologies (GET) research programme which works across ethics, law, and emerging information technologies. He is a leading data ethicist and philosopher specializing in AI ethics, professional ethics, and technology law and policy.

In his current role he leads the Trustworthiness Auditing for AI project, a three-year multi-disciplinary project with University of Reading cutting across ethics, law, computer science, and psychology to determine how to use AI accountability tools most effectively to create and maintain trustworthy AI systems. He also co-leads the A Right to Reasonable Inferences in Advertising and Financial Services project which examines opportunities and challenges facing sectoral implementation of a right to reasonable inferences in advertising and financial services.

ABSTRACT:

UNIVERSITY OF OXFORD, UNITED KINGDOM

Careless speech is a new type of harm created by large language models (LLM) that poses cumulative, long-term risks to science, education, and the development of shared social truths in democratic societies. LLMs produce responses that are plausible, helpful, and confident but that contain factual inaccuracies, inaccurate summaries, misleading references, and biased information. These subtle mistruths are poised to cause a severe cumulative degradation and homogenisation of knowledge over time.

This talk examines the existence and feasibility of a legal duty for LLM providers to create models that "tell the truth." LLM providers should be required to mitigate careless speech and better align their models with truth through open, democratic processes. Careless speech is defined and contrasted with the simplified concept of "ground truth" in LLMs and prior discussion of related truth-related risks in LLMs including hallucinations, misinformation, and disinformation.

EU human rights law and liability frameworks contain some truth-related obligations for products and platforms, but they are relatively limited in scope and sectoral reach. The talk concludes by proposing a pathway to create a legal truth duty applicable to providers of both narrow- and general-purpose LLMs, and discusses "zero-shot translation" as a prompting method to constrain LLMs and better align their outputs with verified, truthful information.

JURY MEMBERS OF 2025



Holger Geißler
marktforschung.de /
succeet



Yannick Rieder
Johnson & Johnson,
DGOF Board Member,
Jury Chair

Sabine Menzel
L'Oréal



Tabea Tesch
Royal Canin,
Mars Inc.



TUESDAY, 01 APRIL 2025
03:45 - 04:45 PM
MAX-KADE-AUDITORIUM

I: IMPACT & INNOVATION AWARD I

SESSION CHAIR: YANNICK RIEDER
JANSSEN EMEA, GERMANY

PRESENTATIONS

OPTIMIZING LOYALTY PROGRAMS WITH AI-DRIVEN INSIGHTS: A CASE STUDY

KAMILA ZAHRADNICKOVA Lakmoos AI and Prague University of Business and Economics, Czech Republic

RELEVANCE & RESEARCH QUESTION Raiffeisenbank sought to optimize its loyalty programs to increase customer engagement and strengthen loyalty. The goal was to identify customer preferences and features that could drive higher program adoption and satisfaction.

METHODS & DATA To address this, Lakmoos AI designed a targeted survey combining synthetic data and customer segmentation analysis:

1. Targeted Survey Design: The survey explored customer opinions on current loyalty features, preferences for new incentives, and variations in preferences across demographic groups.

2. Synthetic Panel Data Simulation: Simulated customer profiles provided predictive insights into how different demographics might respond to features like cashback, reward points, and exclusive benefits.

3. Real-Time Data Validation: Survey findings were validated against real customer data to ensure insights were accurate and actionable.

4. Quantitative and Qualitative Integration: Alongside quantitative questions, open-ended prompts offered qualitative insights into customer priorities and suggestions for program enhancements.

RESULTS Key insights revealed:

- **High-Value Features:** Cashback and flexible point redemption options were highly appealing across demographics, reflecting a preference for tangible rewards.

- **Segmentation Preferences:** Younger customers preferred gamified elements [e.g., badges, status levels], while older customers favored straightforward, financial-based rewards.

- **Personalization Desire:** Many customers indicated greater engagement when programs offered personalized rewards based on spending habits and lifestyle.

ADDED VALUE The findings led to impactful changes for Raiffeisenbank:

1. Program Redesign: Enhanced loyalty programs now include more cashback options, personalized recommendations, and gamified features for younger users, boosting enrollment and engagement.

2. AI Integration: The successful pilot led to AI-driven research being integrated into Raiffeisenbank's standard processes, enabling faster and more precise customer insights.

3. Synthetic Panels in Design Sprints: The adoption of synthetic panels in research democratized insights, making design sprints more inclusive and efficient.

These measures strengthened Raiffeisenbank's ability to create customer-centric, innovative loyalty programs, enhancing their competitive position in the market.

WORDS TO NUMBERS: HOW TO QUANTITATIVELY SIZE AND PROFILE QUALITATIVE PERSONAS

NADJA BÖHME Factworks, Germany
MELISSA RAMONE Yahoo

RELEVANCE & RESEARCH QUESTION Yahoo News sought to validate and size user personas based on preexisting qualitative findings to optimize product and marketing strategies. With a strong user base and brand identity, Yahoo News aimed to drive revenue growth. However, the team lacked a unified understanding of core users, which slowed innovation and delivery. They initially developed five personas through qualitative interviews but needed quantitative validation to determine their accuracy and distribution. Factworks was engaged to refine and enrich these personas for strategic prioritization.

METHODS & DATA To quantify personas, Factworks translated qualitative insights into a structured survey administered to a representative Yahoo user sample. Instead of traditional segmentation techniques, the k-nearest neighbors algorithm was used to classify users based on similarity to preidentified personas. Respondents too dissimilar to any persona were excluded.

This approach allowed Yahoo to size each persona group and prioritize them strategically. Statistical testing helped identify key distinguishing traits, creating more distinct personas. Additionally, a Typing Tool was developed, enabling persona classification via a short questionnaire in Excel for future use in individual and batch scoring.

RESULTS

- Yahoo teams gained a clear, shared understanding of their audiences through the study's active socialization.
- Personas informed product, design, and editorial strategies across six News squads.
- They played a crucial role in cross-functional brainstorming sessions and roadmap planning.
- Product, design, and editorial roadmaps were prioritized based on persona insights.
- The study extended beyond News, influencing Yahoo's Search and Games teams.

ADDED VALUE Yahoo applied these insights to redesign and modernize its Homepage, tailoring updates to the three target personas. Results from iterative experiments included:

- 90% decrease in bounce rate
- 37% increase in classic page views
- 51% increase in Daily Active Users (DAUs)

Following the launch on June 13, 2024:

- Desktop experience saw a 30% increase across sessions, interactions, and page views.
- Mobile experience saw a 40% rise in user sessions and an 8% increase in total page views.

INTELLIGENT DOCUMENTS: EVOLVING EXISTING RESEARCH WORKFLOWS TO EASE AI ADOPTION

GEORG WITTENBURG Inspiritent
SOPHIA MCDONNELL Verian

RELEVANCE & RESEARCH QUESTION Our innovation, Intelligent Documents, redefines the delivery of research insights by transforming static reports into dynamic, interactive resources that allow end-users to engage directly with data. These documents are not just comprehensive reports—they are intelligent interfaces that respond to user questions, enabling follow-up inquiries and deeper exploration of insights. Researchers can navigate findings, request additional analyses, and validate each step of the analysis directly from familiar formats such as Microsoft Word, PowerPoint, or PDF.

Unlike traditional reports, which present insights in a fixed format, Intelligent Documents encourage real-time engagement. Every insight is backed by detailed validated data, which researchers can explore further by clicking embedded links that lead to the analytical foundation of each finding. For example, a marketing director or policy expert might ask, "What other factors are driving this trend?", "How does this vary across demographics?" or "How specifically were these results derived?" and receive immediate, data-driven answers within the same document. This interactivity is facilitated by utilizing Large Language Models (LLMs) to apply Retrieval Augmented Generating (RAG) to a selection of findings that are pre-generated for each analytical result, essentially as "speaker notes" for an LLM that guarantee verifiable answers free of hallucinations.

This interactivity elevates research from a static report to a dynamic platform for investigation, empowering end-users to make informed decisions on the spot. At the same time it combines "old" and "new" by embedding AI features into current document-centric workflows, thereby making the new capabilities of AI more accessibly and thus more relevant for day-to-day use.

TUESDAY, 01 APRIL 2025
05:00 - 06:00 PM
MAX-KADE-AUDITORIUM

II: IMPACT & INNOVATION AWARD II

SESSION CHAIR: YANNICK RIEDER
JANSSEN EMEA, GERMANY

PRESENTATIONS

NO BITTER NOTES: ENSURING DATA QUALITY IN STIEGL'S BRAND TRACKING STUDY WITH REDEM

JULIA MITTERMAYR Redem GmbH, Austria
CHRISTOPH AUBÖCK Media1

RELEVANCE & RESEARCH QUESTION In an increasingly complex global landscape, Germany's Federal Foreign Office sought to gain accurate insights into how Germany is perceived worldwide across various languages and cultures. Ensuring genuine, high-quality responses was paramount to answering the question, "How is Germany perceived?". However, the rising sophistication of survey fraud and AI-generated responses were identified as a significant potential threat to reliable insights. INFO GmbH, the agency leading the study, recognized that traditional quality checks are likely to be insufficient for a project of this scale and complexity. Hence, ReDem's AI-powered and human-refined tool was chosen to tackle the challenge of identifying and cleaning bad survey data, ensuring high data quality across diverse cultural and linguistic contexts.

METHODS & DATA INFO GmbH used ReDem's integration with the survey tool keyingress to apply multiple layers of real-time checks as the data was collected. Each response underwent a thorough, automated assessment: time checks ensured attentive engagement, grid question patterns were analyzed for unusual answer behaviour, and sophisticated algorithms reviewed open-ended answers in multiple languages. These reviews automatically flagged off-topic, duplicated, copy-pasted, or poorly worded responses, etc.

RESULTS By the end of the study, only high-quality data remained in the dataset for the Federal Foreign Office, empowering INFO GmbH to deliver precise and reliable insights with minimal manual intervention. ReDem's technology flagged and removed X% of responses in real-time due to low-quality indicators, catching fraudulent and inattentive data that traditional methods may likely have overlooked. INFO GmbH estimates that ReDem enabled them to capture an additional X% of poor-quality data, raising the overall quality of the dataset. Beyond ensuring data quality for the Federal Foreign Office, ReDem's

solution saved INFO GmbH approximately X hours of manual data review, allowing the agency to reallocate resources strategically and enhance the project's efficiency.

THE TEILHABE-COMMUNITY. AN INFRASTRUCTURE FOR RESEARCH PROJECTS INVOLVING INDIVIDUALS WITH DISABILITIES

LEA THÖNNES Aktion Mensch e.V., Germany
MATTHIES TOBIES Ipsos
ARMGARD ZINDLER Ipsos
ROBERT GRIMM Ipsos

RELEVANCE & RESEARCH QUESTION People with various impairments or disabilities are often not adequately represented in (population) surveys, and their specific subgroups are not identifiable in result analyses. Similarly, participatory product development with people with disabilities rarely takes place. This group wants a voice, a right protected by the UN Disability Rights Convention. Aktion Mensch develops innovations for an inclusive society, including the idea of a special panel. Ipsos has been brought on board for a collaboration to establish this panel. Ipsos and Aktion Mensch launched the "Teilhabe-Community" in early 2023. This panel of around 900 individuals with disabilities is available for diverse research about their life realities and everyday experiences. Due to barriers, these individuals often require specialized communication or customer engagement, necessitating an additional panel.

METHODS & DATA Ipsos hosts the "Teilhabe-Community" on its own online platform. Following a participatory approach, we work with people with disabilities together to ensure the community is as barrier-free as possible. We aim to capture a diverse range of impairments in our panel. In addition to online panels and social media, we leverage contacts through leading social welfare organizations (Aktion Mensch members). Workshops and events, involving advocates like facility directors and support staff, help convey our message. The introduced exchange feature on the platform, requested by individuals with disabilities, also serves the purpose of direct exchange with this group. Moderators assist inexperienced panelists in expressing and forming opinions.

RESULTS The "Teilhabe-Community" serves researchers and companies for varied inquiries, supporting both quantitative and qualitative methods, including offline activity recruitment. Applications span from usual social research to customer and user experience, including customer journey and more. The "Teilhabe-Community" enables companies to integrate diverse experiences of people with disabilities into product or service development, fostering an inclusive society. Aktion Mensch and others have already conducted surveys and website tests.

ADDED VALUE The unique feature of the "Teilhabe-Community" is its easy and quick access to this group. Panelists register in advance, sharing information on their impairments and demographics for targeted sampling for research purposes. The community also allows panelists to interact. Involving them in research significantly contributes to fostering inclusivity.

DEVELOPING A DETAILED UNDER- STANDING OF PATIENT JOURNEYS AND TREATMENT PATHWAYS THROUGH A CO-CREATED DIGITAL-FIRST APPROACH, TO RAPIDLY COLLECT REAL-WORLD PATIENT EXPERIENCE DATA

DANIEL SIMON CHAPPLE DontBePatient Intelligence, Germany
daniel.chapple@dontbepatient.com

RELEVANCE & RESEARCH QUESTION The pharmaceutical client entered the prostate cancer indication for the first time, introducing innovative treatments deviating from standard care pathways. To ensure patient access, diagnostic and therapeutic pathways needed assessment and potential reshaping. The client aimed to identify healthcare professionals (HCPs) involved in various treatment stages across the USA, UK, and Germany and sought strategies to reach and inform patients effectively. A key goal was to strengthen engagement with key opinion leaders (KOLs) and patient organizations, aiming to produce at least three publications for collaboration.

METHODS & DATA A detailed survey, developed with KOLs and patient organizations in the target countries, combined clinical questions with validated scales (BPI, FACT-P) and adaptive designs. This facilitated the comparison of survey results to established metrics. Recruitment employed a digital pathway with social media, search engine ads, patient organizations, and databases while upholding data privacy and participant anonymity. Surveys were conducted in English, German, and Spanish for respective populations. Data analysis applied descriptive statistics and machine learning to uncover clusters and develop predictive models for identifying patients needing enhanced support.

RESULTS Data collection from over 15,800 prostate cancer patients was rapid, taking less than 60 days, with participants showing a median age of 70 and a representative distribution of disease severity. Notably, most participants had no prior interaction with patient organizations but expressed interest in anonymous digital platforms for support and information. Urologists in Germany, clinics in the USA, and clinical oncologists in the UK were key influencers in patient decisions. Despite diverse healthcare journeys, patients shared similar attitudes and expectations across countries.

ADDED VALUE The project provided crucial data to refine communication strategies and define messaging channels for target segments. Findings facilitated physician engagement, reinforced early screening advocacy, and led to the creation of a digital prostate cancer community. Ongoing collaborations with KOLs and patient organizations yielded several research publications and informed treatment launch strategies, including expansion into Japan.

THE GOR IMPACT AND INNOVATION AWARD 2025 IS SPONSORED BY



JURY MEMBERS OF 2025



Dr. Eva Aizpurua
National Centre for
Social Research, UK



**Prof. Dr.
Nicola Doering**
Ilmenau University of
Technology, Germany

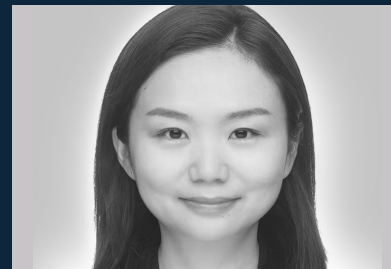
**Prof. Dr.
Florian Keusch**
University of Mann-
heim, Germany, Jury
Chair



Dr. Henning Silber
University of Michigan,
USA



**Prof. Dr.
Christoph Kern**
Ludwig-Maximilians-
University
of Munich, Germany



Chan Zhang
PhD Zhejiang
University, China

**Prof. Dr. habil.
Ramona Schödel**
Charlotte Fresenius
Hochschule, Germany



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 2023 (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 2025.

AWARD CEREMONY:
WEDNESDAY 02 APRIL 2025
11:15 - 11:45 AM

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). To be considered for the award, papers must have been published

WINNER:

KÖHNE, J. UND MONTAG, C.

**CHATDASHBOARD:
A FRAMEWORK TO COLLECT, LINK,
AND PROCESS DONATED WHATSAPP
CHAT LOG DATA**

BEHAV RES 56, 3658-3684 [2024]
[HTTPS://DOI.ORG/10.3758/S13428-023-02276-1](https://doi.org/10.3758/S13428-023-02276-1)

JURY OF THE GOR POSTER AWARD 2025



Dr. Tobias Rettig
Universität Mann-
heim, Jury Chair



Dr. Camilla Salvatori
Utrecht University

Lisa Dust
Kao GmbH , Germany



Leonard Mach
horizoom GmbH



TUESDAY, 01 APRIL 2025
02:30 - 03:30 PM
FOYER OG

4.1: POSTER SESSION

PRESENTATIONS

MAPPING THE WAY FORWARD: INTEGRATING APIS INTO TRAVEL SURVEYS

**EVA AIZPURUA, PETER CORNICK, SHANE HOWE, MARTA MEZZANZANI-
CA, NICK WAUGH** National Centre for Social Research, United Kingdom

RELEVANCE & RESEARCH QUESTION Technological advances are making it increasingly possible to collect information that augments traditional survey research. This study explores the performance of a survey integrated with an Application Programming Interface (API) that allows respondents to identify locations on a map. Specifically, we examine the accuracy and usability of this API integration, focusing on its potential to improve data quality in the context of travel diaries.

METHODS & DATA This research is based on a large-scale pilot conducted in May and June of 2024, where 7,500 addresses in Wales were issued. The web survey was programmed using Blaise 5 and integrated with the Ordnance Survey Places and Names databases, enabling respondents to locate places on a map. Data were collected from 1,008 individuals, resulting in 2,743 reported journeys.

RESULTS Most respondents confirmed that the location information was correct (92% on day 1 and 95% on day 2), indicating a high rate of accurate matches for searched locations. The most commonly reported errors were related to the start or end points of the journeys, followed by the inclusion of journeys that did not occur. In 278 instances (7.9% of all map entries), respondents selected the "I could not find my location" option and provided a free-text description instead. These descriptions varied widely in specificity, ranging from precise locations to vague place names that could correspond to multiple places. While the API integration performs well overall, some respondents encountered difficulties with location precision, suggesting that some degree of post-survey editing will be required.

ADDED VALUE This study highlights the potential of API integration to enhance survey data collection by capturing detailed and specific geographic information. Our findings suggest that while the approach is largely effective, there are areas for improvement to ensure that respondents can accurately and easily identify locations on a map. This research contributes to the growing field of technological augmentation in survey methodologies, offering insights into the practical challenges and opportunities of API-enabled geographic data collection.

UNCOVERING HOUSING DISCRIMINATION: LESSONS LEARNED FROM A LARGE-SCALE FIELD EXPERIMENT VIA IMMOSCOUT24

ZAZA ZINDEL 1,2, LONG NGUYEN 1,2, STEFANIE HECHLER 3, ELISABETH ZICK 11 German Centre for Integration and Migration Research (DeZIM); 2 Bielefeld University, Germany; 3 Freie Universität Berlin, Germany

RELEVANCE & RESEARCH QUESTION Discrimination in the housing market remains a critical societal issue with profound implications for social inclusion and equality. Field experiments offer a robust method to observe discrimination systematically in real-world contexts. This poster focuses on the technical processes involved in designing and executing a large-scale field experiment on housing discrimination in Germany. It highlights iterative developments, shares lessons learned, and offers practical recommendations for researchers.

METHODS & DATA The experiment involved sending over 2,000 standardized rental applications via the online portal Immoscout24 to landlords in 10 major German cities. These applications systematically varied applicant names to signal different ethnic backgrounds. The poster emphasizes key aspects of the technical implementation:

1. Factorial application form design: Creating realistic yet standardized application materials to ensure uniformity and reduce confounding variables.
2. Housing ad selection: Developing a script to identify and scrape rental ads systematically, ensuring representation across regions.
3. Automation of applications: Implementing tools to automate the application process while maintaining ethical and operational standards.
4. Response management protocols: Establishing automated systems for landlord interactions to uphold study integrity.
5. Response tracking and analysis: Constructing a database to monitor landlord responses, record callback rates, and facilitate streamlined analysis.

RESULTS Disparities in callback rates between applicants with German- and non-German-sounding names were observed. However, the poster primarily emphasizes how the technical framework enabled efficient data collection and processing. Key lessons learned in addressing challenges during each phase of the study are discussed.

ADDED VALUE This poster serves as a technical guide for researchers aiming to conduct large-scale field experiments using Immoscout24 or similar platforms. By focusing on practical implementation, it provides actionable insights into overcoming challenges. The framework is adaptable to diverse contexts and underlines the potential of online-based field experiments in studying discrimination.

STRAIGHTLINING IN CS VERSUS AD ITEMS, COMPLETED BY PHONE VERSUS PC

YFKE ONGENA, MARIEKE HAAN University of Groningen

RELEVANCE & RESEARCH QUESTION Straightlining has been shown to be more prevalent in surveys that are completed on a PC than on smartphones. However, on device use for self-completion, differential effects of different items on straightlining have not been researched extensively. Agree-disagree (AD) items are assumed to evoke more straightlining than construct specific (CS) items (i.e., items with a different response scale for each item, depending on the response dimension being evaluated). To fill this gap, we aim to answer the following research question: What are the combined effects of AD versus CS items and device use on response patterns that are assumed to be indicators of satisficing, such as straightlining?

METHODS & DATA Our survey was conducted in November 2024, with 3,500 flyers distributed across a neighborhood in a large Dutch city with subsequent face-to-face recruitment by students. The flyers included a QR code and URL for survey access. The survey was filled out by 556 individuals completing at least 50% of the questions (and 478 completing the full questionnaire), yielding a 13% response rate at household level. A smartphone was used by 85% of the participants, whereas 15% used a PC. Respondents were randomly assigned to four blocks of either five AD items or five CS items. Straightlining was defined in two ways; in a strict definition of providing exactly the same answer two five items asked in the same battery of items and by computing within respondent variance for each battery).

RESULTS Straightlining (both in the strict definition and in terms of lower variance) occurred more frequently in AD items than in CS items, with 18% of respondents in the AD condition showing straightlining, as opposed to 5% of respondents in the CS condition. PC respondents were more likely than smartphone respondents to straightline in battery items phrased as AD items, but this effect was not found when items were phrased as CS items.

ADDED VALUE Our study shows that using CS items might be more beneficial when the questionnaire is filled out on a computer than on a smartphone.

BAYESIAN INTEGRATION OF PROBABILITY AND NONPROBABILITY WEB SURVEY DATA

JOSEPH SAKSHAUG 1,2, CAMILLA SALVATORE 3, BELLA STRUMINSKAYA 3, ARKADIUSZ WISNIOWSKI 4 1 IAB, Germany; 2 LMU-Munich; 3 Utrecht University, the Netherlands; 4 University of Manchester

RELEVANCE & RESEARCH QUESTION The popularity of non-probability sample (NPS) web-surveys is increasing due to their convenience and relatively low costs. On the contrary, traditional probability-sample surveys (PS) are suffering from decreasing response rates, with a consequent increase in survey. Integrating the two types of samples in order to overcome their respective disadvantages is one of the current challenges. We propose an original methodology to combine probability and non-probability online samples to improve analytic inference on binary model parameters.

METHODS & DATA To combine the information coming from the two samples, we consider the Bayesian framework where inference is based on the PS and the available information from the NPS is supplied in a natural way through the prior. We focus on the logistic regression model, and conduct a simulation study to compare the performance of several priors in terms of mean-squared error (MSE) according to different selection scenarios, selection probabilities, and sample sizes. Finally, we present a real data analysis considering an actual probability-based survey and several parallel non-probability web surveys from different vendors which reflect different selection scenarios.

RESULTS The incorporation of biased non-probability data results in posterior estimates that have more bias, but generally, less variance than the non-informative priors. Overall, the informative priors reduce the MSE or, in the worst-case scenario, perform equivalently

to the non-informative ones. According to the different scenarios, we suggest in which case the proposed informative priors are beneficial. Finally, through a cost analysis, we show that such innovative Bayesian data-integration approach has practical implications in the reduction of survey costs.

ADDED VALUE The method provides a means of integrating probability and nonprobability web survey data to address important trade-offs between costs and quality/error. For survey practitioners, the method offers a systematic framework for leveraging information from non-probability data sources in a cost-efficient manner to potentially improve upon probability-based data collections. This can be particularly fruitful for studies with modest budgets or small sample sizes, where the greatest gains in efficiency can be achieved.

TUESDAY, 01 APRIL 2025
02:30 - 03:30 PM
FOYER OG

4.2: POSTER SESSION

PRESENTATIONS

EVERYBODY DOES IT SOMETIMES: REDUCING SOCIAL DESIRABILITY BIAS IN AN ONLINE SURVEY USING FACE- SAVING STRATEGIES

EMMA ZAAL, YFKE ONGENA, JOHN HOEKS University of Groningen, Netherlands, The
e.l.zaal@rug.nl

RELEVANCE & RESEARCH QUESTION Online surveys are indispensable for measuring human behaviors and cognitions. However, Social Desirability Bias (SDB) - the tendency to present oneself in a favorable light - poses challenges for survey research addressing sensitive topics. Attempting to reduce SDB, we conducted an online survey experiment in which we employed so-called face-saving strategies. Such strategies aim to ease the discomfort of admitting socially undesirable behaviors by reassuring respondents that deviating from normative behaviors can be unavoidable.

Still, it is unclear what mechanisms exactly drive the effectiveness of face-saving strategies and under what circumstances the method is most effective. This leads to our research question: "To what extent can SDB be reduced in an online survey integrating face-saving strategies using various operationalizations and question formats?"

METHODS & DATA This online experiment was conducted with participants recruited door-to-door using flyers with QR-codes, in two

neighborhoods in Groningen, the Netherlands (N=529). The survey included questions on volunteering, cohesion, disturbances, safety, and poverty. Our experimental conditions included face-saving formulations in questions and/or answers and were compared to control conditions with conventional formulations. For instance, we offered a face-saving preamble that employed downtowning (weakening a social norm, "It can sometimes be challenging to [...] It's also very normal to find this difficult."). We also investigated face-saving answer options by incorporating degrees of truth (reflecting partial agreement, like "occasionally") and credentialing (justifying norm-violations, like "no, but I don't have time").

RESULTS Compared to a yes/no format, face-saving answer options significantly increased the likelihood of socially undesirable responding, suggesting a reduction in SDB. A face-saving preamble did not seem to affect SDB. This poster presentation will offer results based on other conditions that were tested as well, providing insights into the effectiveness of face-saving in different question formats and formulations.

ADDED VALUE We examined novel operationalizations of face-saving strategies across various question formats, advancing our understanding of SDB-reduction and effective survey design. By refining and expanding the application of face-saving strategies, researchers and practitioners can ultimately enhance data quality significantly. This is in turn essential for informing evidence-based policies and interventions aimed at addressing societal issues.

DECODING STRAIGHTLINING: THE ROLE OF QUESTION CHARACTERISTICS IN SATISFICING RESPONSE BEHAVIOR

CAGLA EZGI YILDIZ, JESSICA DAIKELER, FABIENNE KRAEMER GESIS - Leibniz Institute for the Social Sciences, Germany
HENNING SILBER University of Michigan
EVGENIA KAPOUSOUZ NORC at the University Of Chicago

RELEVANCE & RESEARCH QUESTION Satisficing response behavior, including straightlining, can threaten the reliability and validity of survey data. Straightlining refers to selecting identical (or nearly identical) response options across multiple items within a question, potentially compromising data quality. While straightlining has often been interpreted as a sign of low-quality responses, there is a need to distinguish between plausible and implausible straightlining (see Schonlau and Toepoel, 2015; Reuning and Plutzer, 2020). With this research, we introduce a model that classifies straightlining into plausible and implausible patterns, offering a more nuanced understanding of the conditions under which straightlining likely indicates optimized response behavior (plausible straightlining) vs. satisficing response behavior (implausible straightlining).

For instance, straightlining behavior is plausible when answering attitudinal questions with items worded in the same direction, but it becomes implausible when items are reverse-worded. This study further examines how question characteristics – including grid size, design (e.g., matrix vs. single-item formats), straightlining plausibility – influence straightlining behavior.

METHODS & DATA For our analyses, we use the German GESIS Panel, a mixed-mode (mail and online), probability-based panel study, leveraging a change in the panel's layout strategy in 2017 that shifted multi-item questions from matrix to single-item designs, offering a unique quasi-experimental set-up. We conduct multilevel logistic regression analyses to assess the effect of question design and grid size on straightlining behavior. We further conduct difference-in-difference analyses to examine the format change's effect on plausible and implausible straightlining.

RESULTS Our initial multilevel regression analyses, using data from 3514 respondents and 18 grid questions from the wave before and after the design switch, show that matrix designs are associated with higher levels of straightlining compared to single-item designs. Our preliminary analyses, based on coding by five survey methodology experts, classify 22.2% of these questions as exhibiting plausible straightlining, with the remainder showing implausible patterns. Further analyses investigate how these classifications correspond to conditions under which straightlining reflects optimized versus satisficing response behavior, offering deeper insights into the role of question characteristics.

ADDED VALUE This research enhances questionnaire design and the accurate identification of low-quality responses, addressing gaps in linking question characteristics to straightlining plausibility.

LONG-TERM EFFECTS OF LATE RESPONDING IN PANEL SURVEYS

SILVIA SCHWANHÄUSER Institut für Arbeitsmarkt- und Berufsforschung (IAB), Germany

RELEVANCE & RESEARCH QUESTION Late responding survey participants – i.e., respondents who complete a survey only after a certain delay or after a higher recruitment effort – are often assumed to provide lower data quality. Several studies have therefore focused on the potential measurement error introduced by late respondents. On the other hand, studies have shown that the addition of late respondents can increase the accuracy of estimates by reducing sample selectivity and thus nonresponse error. Although this trade-off between measurement and nonresponse error is well documented, less is known about the potential long-term effects of late responding. This is especially true for web surveys, where self-selecting processes and measurement errors cannot be mitigated by interviewers. Therefore, this study investigates whether differences in measurement and nonresponse are limited to one wave of a panel survey or more persistent, also affecting follow-up waves. To answer this question, this study examines the long-term effects of late responding on 1) the selectivity within tranches, 2) panel attrition, and 3) several internal indicators of data quality such as response duration, non-differentiation, and item nonresponse.

METHODS & DATA This case study uses a unique experimental design, in which a large-scale panel survey was combined with a high-frequency web survey, in between two main waves. During the 26-week field period of the main surveys, respondents were divided into three disjunct tranches based on the timing of their participation in the main surveys. Thus, tranche 1 includes early respondents while tranche 3 includes late respondents. For the follow-up survey, each tranche was repeatedly invited to complete a short web survey, giv-

ing each respondent the opportunity to participate in three additional sub-waves. The timing of the invitation also depended on the tranche allocation, thus ensuring equal conditions for all tranches.

RESULTS First preliminary results indicate lower response rates and hence higher attrition in sub-waves for late respondents of the main survey, as well as faster response durations.

ADDED VALUE The study informs practitioners about the response quality and long-term effects of late respondents in web surveys. This information is important for managing attrition in panels and developing appropriate recruitment strategies.

POLITICAL COMMUNICATION ON SOCIAL MEDIA: STRATEGIC ANALYSIS OF THE BAVARIAN STATE ELECTION CAMPAIGN 2023

JAKOB BERG University of Regensburg, Germany

RELEVANCE & RESEARCH QUESTION This research explores political communication on Instagram during the 2023 Bavarian state election campaign, analyzing if and how candidates' strategies influence reach and engagement. A model categorizes key strategies, including self-presentation, party representation, thematic content, negative campaigning, and direct calls to vote. The study investigates which strategies drive Bavarian audience interaction and how Instagram serves as a platform for political mobilization.

METHODS & DATA The study is based on 46,897 Instagram posts from 619 candidates, analyzed using advanced machine learning techniques. OCR was applied to extract embedded text in images, while BERTopic identified thematic patterns. A prompt-guided classification with GPT-4o-mini enabled a nuanced categorization of posts according to predefined communication strategies. Evaluation metrics showed high accuracy, ensuring the scalability and reliability of the methodology.

RESULTS The research confirms Instagram's importance as a platform for political outreach also in Bavaria, with high interaction rates and active participation from candidates across relevant parties. While all identified communication motives were present, their frequency and effectiveness varied. Posts focusing on personal self-presentation and political content, such as policy issues and party-related topics, were dominant. In contrast, fundraising and internal communication were rare, reflecting their lower relevance in this electoral context. Regression analysis revealed that personal self-presentation significantly boosted engagement, aligning with findings that informal content fosters relatability and connection. Negative campaigning also had a slight positive effect, capturing audience interest in the Bavarian election context. Surprisingly, posts focused on key policy issues and thematic content, while frequent, had little impact on engagement rates. This supports findings in social media research indicating that personal and informal content resonates strongly with audiences, likely because it promotes a sense of connection and relatability.

ADDED VALUE The findings also suggest that the identified communication strategies alone are insufficient to explain the varying success

of political posts in terms of engagement rates and reach. It appears that other post characteristics - such as technical factors (e.g., video format or interactive elements), emotional triggers, or the offline popularity of candidates - play a more significant role in driving engagement on Instagram

TUESDAY, 01 APRIL 2025
02:30 - 3:30 PM
FOYER OG

4.3: POSTER SESSION

PRESENTATIONS

MODULAR SURVEY DESIGN: EXPERIMENTAL EVIDENCE FROM THE GERMAN INTERNET PANEL (GIP)

BENJAMIN GRÖBE, CAROLIN BAHM, ANNE BALZ, TOBIAS RETTIG

University of Mannheim, Germany

ALEKSA MÖNTMANN forsa Gesellschaft für Sozialforschung und statistische Analysen mbH

RELEVANCE & RESEARCH QUESTION A large body of interdisciplinary literature has clearly demonstrated the negative effects of lengthy surveys. A shorter questionnaire, by contrast, could reduce the burden on respondents and again contribute to better data quality. Starting from this assumption, we investigate how the use of a modular survey design, in which a self-administered survey is split up into several parts that can be completed at the respondent's convenience, affects participation and module completion rates, order of starting modules and respondents' evaluation of the questionnaire.

METHODS & DATA For this purpose, we embedded a randomized experiment over three survey waves in the German Internet Panel (GIP), a long-standing probability-based online panel of the German population. We split the questionnaires into three to four topic modules, each of which took between three and 15 minutes to complete and randomly divided respondents into five groups, each receiving different information about the topic, length and incentives of the modules. While the respondents in group 1 received the questionnaire as a single survey in the usual way, group 2 was additionally shown the topics of the individual survey modules. Participants in groups 3, 4, and 5 were given an overview of all survey topics and were free to choose the order and timing in which they completed the different survey modules, but were given different additional information about the length and incentives of the survey modules.

RESULTS Our results indicate that respondents did not complete the survey more often when they could choose the order in which they answer the modules, even when they knew the length and incentive

next to the topic of each module. Rather than determining the order in which modules should be answered, respondents showed a clear preference for answering all survey modules in the given order at one time.

ADDED VALUE Our aim is to find out how web-based surveys, especially long-term panel studies, can be optimized in order to increase data quality (i.e. higher completion rates, respondent satisfaction) and - in the long term - to retain respondents in the panel.

GXC - AN R PACKAGE FOR SPATIAL LINKING OF EARTH OBSERVATION DATA WITH SOCIAL INDICATORS

DENNIS ABEL, STEFAN JÜNGER GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION This poster will present the usability and benefits of a new R package for the flexible integration of georeferenced survey data or other social science datasets with customized Earth observation (EO) data derived from public sources like the Copernicus data services.

METHODS & DATA The unique feature of the tool is the possibility of carrying out both geographically and temporally medium- to high-resolution queries, which at the same time function efficiently on simple workstations. Our tested workflow development has identified five major levers: parameter type, indicator intensity, focal time period, baseline time period, and spatial buffer. Flexibility on these five attributes will be maximized for users. The tool also offers the functionality to automatically derive spatio-temporal links with other georeferenced data (e.g., surveys, digital behavioral data). Users benefit from the core variables integrated into the interface for social research. Examples include data on local air quality and pollutants, extreme weather events, or land use changes.

RESULTS The workflow will be exemplified on the basis of a research project which we currently conduct on the effects of flooding exposure on climate change opinion.

ADDED VALUE The project advances the automatization of these data integration processes between social science data and EO data based on an open-source, user-friendly tool.

UNDERSTANDING REDEMPTION PATTERNS: A STUDY OF POINTS-BASED INCENTIVE SCHEMES IN ONLINE PANEL SURVEYS

CHRISTINE DISTLER, MUSTAFA COBAN, WEIK JONAS IAB Nürnberg, Germany
christine.distler@iab.de

RESEARCH QUESTION Many online surveys offer incentives to enhance response rates, betting on stronger motivation for a response once respondents' participation costs are rewarded. Commonly, respondents receive incentives such as cash or vouchers. Additionally, online panel surveys may include a points-based incentive program allowing respondents to accumulate reward points throughout the study and redeem them anytime to get a shopping voucher. We would like to address the following research questions: What are the distinct patterns of reward point redemption among online survey participants, and how can these be categorized into behavioural clusters? How do demographic and socioeconomic factors influence reward point redemption behaviours?

METHODS In 2023, the Institute for Employment Research in Germany launched a new online panel survey of the German workforce (IAB-OPAL) using a push-to-web approach. The quarterly survey utilises a post-paid points-based incentive program, allowing respondents to earn reward points in their accounts after completing the survey. They can collect these points over time and redeem them for shopping vouchers from various providers at their convenience.

We comprehensively assess respondents' redemption behaviours of across five survey waves using individual tracking data on inflows and outflows of reward points of 13.513 panelists. First, we analyse recurring redemption patterns and identify distinct behavioural clusters by applying time series k-means. Second, we explore other dimensions of redemption behaviour, such as the timing of point redemption across different demographic groups and specific temporal trends. Lastly, we investigate the demographic and socioeconomic drivers of redemption behaviours, giving special attention to the respondents who collect reward points without redeeming them.

RESULTS The analysis reveals several key insights into reward point redemption behaviours within the IAB-OPAL panel survey. Respondents exhibited a wide range of behaviours, from frequent small redemptions to rare but large-point redemptions. Through clustering methods, distinct behavioural groups were identified.

ADDED VALUE Our findings shed light on the dynamics of reward point redemption in online panels and have practical implications. We provide valuable guidance for designing online panel surveys that may incorporate a points-based incentive program. Moreover, our results can assist survey practitioners in budget planning, decision-making, and fieldwork preparation.

TUESDAY, 01 APRIL 2025
02:30 - 03:30 PM
FOYER OG

4.4: POSTER SESSION

PRESENTATIONS

CHILDREN'S SOCIAL MEDIA BEHAVIOR AND SOCIODEMOGRAPHICS: A SEGMENTATION APPROACH

MICHAEL WÖLK, ELISABETH WOLFSTEINER, STEFAN EIBL, SONJA HAUER, VERENA TATZER-HANTEN FH Wiener Neustadt GmbH, Austria

RELEVANCE & RESEARCH QUESTION The growing influence of social media has sparked increasing concerns about its impact on digital well-being, particularly among young people. A survey from 2023 revealed that adults are highly concerned about the effects of social media on their children's mental health (Knight Foundation, 2023).

Our literature review highlighted key areas of existing research, such as the relationship between parents' social media sharing and children's privacy (Ong et al., 2022), the adverse effects of digital media on toddlers (e.g., Barr, 2022), emotional developmental delays, anxiety, and depression among children (O'Riley et al., 2018; Primack et al., 2017), and the short-term benefits of social media breaks (e.g., Brown & Kuss, 2020).

Building on these findings, the current research aims to (1) explore adolescents' (10-14 years) social media behavior and (2) identify different social-media-user segments based on their social media behavior and sociodemographic variables.

METHODS & DATA To answer the research question, a quantitative online survey was conducted. A convenience sample of 297 (Austrian) respondents (parents of at least one child aged 10-14) completed the questionnaire for small remuneration. To identify social-media-user segments, we collected information about adolescents' consumption behavior of social media (i.e. use of different platforms, duration of use, social-media-activities, active vs. passive social-media-usage, different social-media-rules set by parents) evaluated by their parents. Finally, we collected sociodemographic data (age, gender, type of school of adolescents; age, gender, highest education of parents). We then conducted a two-step cluster analysis (applicable for nominal and metric data) to derive segments based on social media behavior of adolescents.

RESULTS We are currently refining and deepening the results and implications of the cluster analysis which will – hopefully – be discussed at GOR 2025.

ADDED VALUE Exploring adolescents' social-media-behavior is critical from both a scientific and a practical perspective by providing insights into how different segments of adolescents engage with social media, enabling tailored interventions for digital well-being, education, and digital literacy. Additionally, it supports policymakers in creating ethical, segment-specific strategies for communication, product design (e.g. apps), and online safety regulations.

SMART SURVEY IMPLEMENTATION: EXPERIENCES FROM EXPERIMENTS IN THREE EUROPEAN COUNTRIES

MAREN FRITZ, FLORIAN KEUSCH University of Mannheim, Germany
NINA BERG Statistics Norway
PETER LUGTIG Utrecht University

RELEVANCE & RESEARCH QUESTION Smart surveys combine surveys with smart elements from sensors, for example, the use of the smartphone camera for receipt scanning in a household budget survey and the use of geolocation tracking to identify activities in time use surveys. At this point, relatively little is known on how to best implement smart surveys in the general population for official statistics, and what influence the different features of smart surveys have on participation behavior.

METHODS & DATA In 2024, fieldwork experiments were conducted in Norway, Belgium, and Germany to test various options in how to design and field smart surveys as part of national household budget surveys and time use surveys. The experiments in the three countries varied several design features to test their effect on recruitment and participation rates to smart surveys.

In Norway, the use of different platforms from which the data collection app could be accessed and the use of CATI interviewers for recruitment and follow-up was tested. In Belgium and Germany different recruitment protocols were tested, including the use of different appeals in the invitation letters focusing on features of the smart survey (e.g., use of the camera to scan receipts) and secondary data collection modes (e.g., paper questionnaires instead of app).

RESULTS We find that recruitment and participation rates vary across countries, and that the differences between within-country experimental conditions are relatively small. The poster will present results on differential recruitment and participation rates and nonparticipation bias in the three countries.

ADDED VALUE This research is part of the Smart Survey Implementation (SSI) project, funded by EUROSTAT, which aims to enhance data collection for official statistics across Europe through digital innovation. This experiment specifically addresses recruitment challenges in app-based surveys and evaluates the potential of mobile technology to streamline participation in official household budget and time use surveys.

SCENARIO-BASED MEASURES OF SMARTPHONE SKILLS IN ONLINE SURVEYS

WAI TAK TUNG, ALEXANDER WENZ University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION Digital skills have become important for navigating in today's information society. While prior digital inequality research has mostly focused on studying general internet uses and skills, research on smartphone-specific inequalities is still scarce. In addition, existing measurement instruments mostly rely on survey-based self-reports or small-scale laboratory-based performance tests that are susceptible to measurement and representation errors.

In this study, we examine the feasibility of using novel scenario-based measures to evaluate the level of smartphone skills in the general population. Scenario-based measures evaluate smartphone skills by assessing how well respondents perform a set of smartphone activities described in a hypothetical situation.

METHODS & DATA Data were collected in the German Internet Panel, a probability-based online panel of the general population aged 16-75 in Germany, in March 2022. Respondents were asked to answer three scenario-based questions and rate their general smartphone skills. The scenario-based questions asked respondents to correctly order a set of steps to carry out smartphone activities, such as buying a train ticket with an app that is not yet installed on their device. We examine response distributions and correlations between the scenario-based and self-reported measures. We also assess whether predictors of smartphone skills differ between the two measures.

RESULTS The scenario-based and self-reported measures are significantly positively correlated and measure the same underlying construct as determined by an exploratory factor analysis. Compared to self-reports, the scenario-based measures, however, have substantially greater rates of item-nonresponse. Older and less educated smartphone owners are significantly less likely to respond to the scenario-based questions.

The predictors of smartphone skills differ by respondents' sociodemographic characteristics across the two measures. Older, female, and more educated respondents are more likely to underreport their smartphone skills in the self-report compared to the scenario-based questions.

ADDED VALUE Methodologically, we demonstrate the feasibility of using scenario-based measures of smartphone skills in an online survey. Substantively, we contribute to the growing body of research on the second-level smartphone divide.

WHO DONATES THEIR GOOGLE SEARCH DATA? PARTICIPATION IN A DATA DONATION STUDY DURING THE 2025 GERMAN FEDERAL ELECTION

SINA CHEN, BARBARA BINDER GESIS, Germany

RELEVANCE & RESEARCH QUESTION Data donation is a relatively new, user-centered approach to collecting digital trace data, increasingly relevant in various research fields (Haim et al., 2023). It offers significant potential to understand online behavior, especially amid growing API restrictions, as it can help validate and enrich survey data. However, little is known about who participates in these studies (Keusch et al., 2024). Previous results hint that participation is non-random (Welbers et al., 2024). Our study aims to identify correlates of participation, moving beyond sociodemographics to consider political preferences and vote choice.

METHODS & DATA Wave 60 participants from the GLES Tracking pre-election survey ($n \approx 2,000$, CS, CAWI, online access panel) are invited after the 2025 German Federal Election to complete a follow-up survey on their vote choice, followed by a request to donate their Google search histories from six weeks before to one week after the election. Besides search terms, we collect URLs of clicked search results and the order of searches and clicks. We experimentally vary the invitation framing to evaluate how different descriptions of the collected data (“web services”, “data collected by Google”, “data collected by Google Search”) influence participation rates. The digital trace data will be linked to both survey datasets to analyze participation across sociodemographics, political attitudes, electoral preferences, and vote choice. Our design captures multiple participation stages, from initial opt-in to final data donation.

RESULTS Data is collected one week after the 2025 German Federal Election (February 23, 2025). At the GOR conference, we will present results on participation rates by invitation framing, sociodemographics, political preferences, and vote choice, differentiating between respondents who opted in, consented, processed, and ultimately donated their data.

ADDED VALUE This study provides insights into how sociodemographic characteristics relate to participation in data donation studies. Additionally, we examine political preferences and vote choice, as well as how the specificity of the data requested in the invitation affects willingness to participate. Since Google search data is highly private, our findings will inform optimal framing for requests involving sensitive data. Our study advances methodological approaches in digital behavioral data research and provides practical guidelines for designing effective data donation studies.

TUESDAY, 01 APRIL 2025
02:30 - 03:30 PM
FOYER OG

4.5: POSTER SESSION

PRESENTATIONS

PAY TO STAY? EXAMINING THE LONG-TERM IMPACT OF INITIAL RECRUITMENT INCENTIVES ON PANEL ATTRITION

ALMUTH LIETZ, JANNES JACOBSEN, JONAS KÖHLER, MADELEINE SIEGEL, JÖRG DOLLMANN, SABRINA J. MAYER German Center for Integration and Migration Research, Germany

RELEVANCE & RESEARCH QUESTION This study explores the long-term effects of incentives provided during the recruitment wave of a panel study, focusing on their impact on panel consent and subsequent dropout risks. Specifically, it examines how the amount and conditionality of incentives influence participation decisions and whether transitioning from a pre-paid incentive in the recruitment wave to a post-paid incentive in subsequent waves affects dropout rates.

These questions are critical for developing sustainable strategies for panel recruitment and retention, particularly in the context of probability-based online access panels.

METHODS & DATA The analysis is based on data from the DeZIM panel, a probability-based online access panel in Germany. Logistic regression models were used to examine panel consent and participation behaviours across different experimental conditions. The dataset includes responses from 9,168 participants in the recruitment wave and longitudinal data spanning 70,926 person-years across subsequent waves. The experimental design varied incentive types (pre-paid vs. postpaid) and amounts (€5 vs. €10), enabling a detailed assessment of their effects on respondent behaviour.

RESULTS The findings reveal that pre-paid incentives significantly reduce panel consent compared to post-paid incentives. However, the amount of the incentive (€5 vs. €10) does not significantly influence consent rates. Long-term analyses show no substantial effects of either the incentive type or amount on participation rates across subsequent panel waves.

Furthermore, switching from a prepaid incentive in the recruitment wave to a postpaid incentive in the first panel wave does not increase dropout risks in future waves. These results suggest that higher or unconditional incentives do not necessarily yield higher response rates or long-term participation in panel studies.

ADDED VALUE This study contributes to the literature on survey methodology by challenging the common assumption that higher or unconditional incentives automatically enhance response rates. The findings emphasize the importance of designing incentive strategies that balance short-term participation gains with the long-term sustainability of panel studies. By providing empirical evidence from a probability-based online panel in Germany, the study highlights the potential for more cost-effective and nuanced incentive policies in maintaining panel participation over time.

SOCIALLY DESIRABLE RESPONDING IN PANEL STUDIES – HOW DOES REPEATED INTERVIEWING AFFECT RESPONSES TO SENSITIVE QUESTIONS?

FABIENNE KRAEMER GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Social desirability (SD-) bias (the tendency to report socially desirable opinions and behaviors instead of revealing true ones) is a widely known threat to the validity of self-reports. Previous studies investigating socially desirable responding (SDR) in a longitudinal context provide mixed evidence on whether SD-bias increases or decreases with repeated interviewing and how these changes affect response quality in later waves. However, most studies were non-experimental and only suggestive of the mechanisms of changes in SD-bias over time.

METHODS & DATA This study investigates SDR in panel studies using a longitudinal survey experiment comprising six waves. The experiment manipulates the frequency of answering identical sensitive questions (target questions) and assigned respondents to one of three groups: One group received the target questions in each wave, the second group received the target questions in the last three waves, and the control group received the target questions only in the last wave of the study. The experiment was conducted within a German non-probability (n = 1,946) and a probability-based panel (n = 4,660). The analysis focusses on between- and within-group comparisons to investigate changes in answer refusal and responses to different sensitive measures.

To further examine the underlying mechanisms of change, I conduct moderator and mediator analyses on the effects of respondents' privacy perceptions and trust towards the survey (sponsor).

RESULTS First results show a decrease of answer refusal and SDR with repeated interviewing for most of the analyzed sensitive measures. However, these decreases were non-significant for both between-group comparisons and comparisons over time.

ADDED VALUE Altogether, this study provides experimental evidence on the impact of repeated interviewing on changes in SD-bias and contributes to a deeper understanding of the underlying mechanisms by examining topic-specific vs. general survey experience and incorporating measures on privacy perceptions and trust towards the survey (sponsor).

USING APPS TO MOTIVATE RESPONDENTS: EXPERIENCES FROM AN EXPERIMENT IN A REFUGEE PANEL STUDY

FLORIAN HEINRITZ, JUTTA VON MAURICE Leibniz Institute for Educational Trajectories, Germany

MICHAEL RULAND, THOM WEISS, KATHARINA SANDBRINK infas Institute for Applied Social Science, Germany

RELEVANCE & RESEARCH QUESTION Be it online panels or traditional panel studies, the commitment of participants to a longitudinal study is essential for the stability of a panel. In many panels, respondents are therefore contacted by letter at regular intervals with information about the study. With the growing importance of smartphones, tailored apps that use smartphones as a channel to communicate with respondents can help to keep in touch with respondents by sending them notifications and messages about the study. In this context the question arises, whether it makes sense to contact respondents more frequently or less frequently via app and how do respondents react to notifications?

METHODS & DATA In the study "Refugees in the German Educational System (ReGES)", an experiment was conducted with the "my infas" app that varied the frequency of contact via the app. One experimental group of respondents was contacted 6 times, while the rest were only contacted 12 times in the same period. The data from this experiment is used to cluster the participation behavior more precisely using a sequence analysis.

RESULTS Of the 2,740 respondents to whom we sent messages via the app as part of the experiment, only 271 people read at least one message. Due to this low number of participants, the potential for analysis is limited. Nevertheless, it is possible to draw some insights from this experiment into how respondents react to messages. Four clusters were identified: The One-Time Clickers, the Curious, the Awakened and the Interested. If we look not only at whether the message and the results were read, but also at when the messages were read, we see that many respondents read the messages very late (on average 28 days after receiving them).

ADDED VALUE As many people do not react to them at all or very late, the analyses suggest that using an app in panel studies is not a guaranteed success, but potential issues such as respondents turning off push notifications or uninstalling the app need to be considered as time-critical messages sent via apps may not reach all respondents.

THE GOR POSTER AWARD 2025 IS SPONSORED BY HORIZOOM



JURY OF GOR THESIS AWARD 2025



Dr. Olaf Wenzel
Wenzel Marktforschung,
Jury Chair



Dr. Carina Cornesse
GESIS – Leibniz-
Institut für Sozial-
wissenschaften

Dr. Frederik Funke
LimeSurvey &
datenmethoden.de



**Prof. Dr.
Bettina Langfeldt**
University of Kassel



Dr. Tabea Tesch
Royal Canin, Mars Inc.

SOLICHATUS ZAHROH 1, PETER LUGTIG 1, YVONNE GOOTZEN 2, JONAS KLINGWORT 2, BARRY SCHOUTEN 1,2 1Utrecht University, The Netherlands; 2 Statistics Netherlands, The Netherlands

RELEVANCE & RESEARCH QUESTION The general population travel survey is burdensome for the respondents as each respondent should document a full-day trip, including the precise start times, end times, and locations. Predicting travel purposes automatically would be more beneficial than traditional travel surveys to overcome the burdensome. This study aims to answer research question “How well can we predict the travel purpose using sensor data from a smartphone-based travel diary study?”

METHODS & DATA CBS collected the data from November 2022 to February 2023 using the ODin app and contains a total of 505 users. Administrative data, or demographic variables, from the CBS database is linked with the location data. Multiple bounding boxes with varying radiuses were determined from OSM for different tags associated with trip purposes. The data will be partitioned into training and testing sets with a ratio of 80:20 from the 4961 locations available.

This study used an Artificial Neural Network (ANN) and as comparison, the balanced accuracy of Random Forest (RF), Naive Bayes (NB), Support Vector Machine (SVM), and Extreme Gradient Boosting (XGB) were evaluated. Weather data that has a potential to assist trip-purpose prediction was also included.

RESULTS The initial analysis was performed without OSM data. Later, OSM was included and treated as count data and percentages. ANN model with OSM data from 1 radius as count data and percentage was the most optimal model. The model's accuracy increased to over 70% for the overall model. ANN and SVM showed similar results while RF and XGB performed almost perfect. Unfortunately, NB was not good enough. The confusion matrix indicates that despite the accurate classification of classes, the limited number of observations, notably in sports and education, results in low accuracy. Regrettably, the inclusion of weather data did not enhance the model.

**TUESDAY, 01 APRIL 2025
10:45 - 11:45 AM
HÖRSAAL D**

GOR THESIS AWARD I: BACHELOR/ MASTER

SESSION CHAIR: OLAF WENZEL
WENZEL MARKTFORSCHUNG, GERMANY

PRESENTATIONS

**PREDICTING TRAVEL PURPOSE IN A
SMARTPHONE-BASED TRAVEL SURVEY**

ADDED VALUE Sensor data from a smartphone-installed travel diary application successfully predicts the purpose of a trip by utilising spatial and temporal patterns. Prediction models give more importance to trends that occur over time in specific locations. Some types of stops would be more important for making accurate predictions than others.

SATISFICING IN A GERMAN SELF-ADMINISTERED PROBABILITY-BASED PANEL SURVEY

JULIA WITTON

Deutsches Institut für Wirtschaftsforschung, Germany

RELEVANCE & RESEARCH QUESTION When applied to the survey research context, satisficing theory describes a range of behavioral strategies survey participants may use to reduce the cognitive effort required to answer questions conscientiously and truthfully (i.e., optimally). The resulting response error undermines data quality, affecting both reliability and validity of survey results. While general recommendations for preventing satisficing exist, they do not fully account for the complexity and individuality of the behavior. Therefore, satisficing remains a significant challenge to survey researchers. If we could predict which respondent is at risk of a certain type of satisficing behavior, we may be able to prevent satisficing using targeted interventions, particularly in a panel survey context where we have more information on satisficing types and their correlates over time. This study aims to explore the potential for such targeted approaches by investigating the robustness and predictiveness of survey satisficing in self-administered mixed-mode panels.

The key research questions are:

1. Can distinct patterns of satisficing behavior be identified using latent class analysis (LCA)?
2. Do these patterns replicate across survey waves and modes?
3. Do respondent characteristics correlate with the identified satisficing pattern?
4. Can future satisficing behavior be predicted based on satisficing patterns exhibited in previous survey waves?

METHODS & DATA The analyses base on the first three waves of the German Social Cohesion Panel (SCP), which is conducted jointly by the German Institute for Economic Research (DIW Berlin) and the Research Institute Social Cohesion (RISC) recruited in 2021. The SCP is a mixed-mode panel survey with participants self-selecting into either paper-and-pencil (PAPI) or web (CAWI) mode. The sample consists of 17,029 individuals nested in 13,053 households.

I generated several indicators of satisficing behavior, including extreme and midpoint response selection, open and closed question nonresponse, speeding (in CAWI), and nondifferentiation (Nd) across item batteries. Nd is the tendency to select the same or similar response categories across a number of items, resulting in overall limited response variability. I measured nondifferentiation using the mean root of pairs method, distinguishing between unidimensional (weak Nd) and multidimensional or reverse-coded item batteries (strong Nd). To identify distinct patterns of satisficing behavior, I employed latent class analysis (LCA). To account for the hierarchical data structure, I used robust standard errors. For each combination of survey

wave (waves 1, 2, 3) and mode (PAPI, CAWI), I estimated six latent class models (assuming one to six latent classes). From the 36 models, I selected the six final models based on information criteria (BIC, CAIC, SABIC) and classification diagnostics (entropy, average posterior probabilities, odds of correct classification). Finally, I interpreted the six resulting latent class models regarding their homogeneity and distinctiveness.

Afterwards, I used multinomial logistic regression to investigate whether individual-level characteristics such as sociodemographic factors, predict class membership. Finally, I applied logistic regression to predict the most likely latent class membership in a wave by the estimated posterior class probability for the same latent class in a previous wave.

RESULTS Across the online survey mode, the LCA models identified three consistent latent classes that differ in their propensity to engage in certain satisficing strategies: The largest class are Optimizers. Counterintuitively, optimizers do not completely dispense with satisficing but exhibit comparably little and unspecific satisficing behavior. Optimizers may occasionally skip questions or apply nondifferentiation to reduce cognitive effort, sometimes speeding through questions.

Extremists are the second largest class. A typical Extremist is very likely to nonrespond to at least one closed as well as open question, reliably generating item missings. Extremist's responses will likely be the highest or lowest extreme values of the given response scales. By providing an approximate answer, they circumvent the cognitive engagement that nuanced and differentiated responses necessitated.

Indifferents are the smallest class. Typical Indifferents can be identified through nondifferentiation with a tendency to select the mid-points of response scales. Furthermore, they are at risk to speed through the survey.

However, in the paper mode, models demonstrate variability in their global-level structures across waves. A fourth class of "Missers," exclusively emerging in the second wave's paper mode, was identified as primarily engaging in item nonresponse. Besides, no class of optimizers was identified in the paper mode of the third wave. In the multinomial analyses, I found individual-level characteristics, such as education and income to be associated with class membership, with limited predictive power.

Regarding the robustness of class membership, the regression analyses revealed that the estimated posterior probabilities of belonging to a satisficing class in a previous wave were consistently significant predictors of belonging to that same class in the future, with odds ratios ranging from around 2 to 10. In some models, I found main effects of the survey mode as well as moderation effects of the survey mode on the effect of the past on the future satisficing strategy.

The effect sizes are consistently modest, with Nagelkerke R² values between 0.6 to 0.25.

ADDED VALUE This study provides several important insights for survey research. First, it demonstrates the feasibility of using LCA to identify distinct typical satisficing patterns in self-administered mixed-mode surveys. Second, the identification of distinct satisficing patterns across survey waves and modes suggests the potential for targeted interventions to mitigate satisficing. Third, the finding that

the global-level satisficing patterns replicate across survey waves for the online mode but not for the paper mode suggests that mode-specific approaches to targeted interventions may be necessary.

Fourth, the moderate predictiveness of past satisficing behavior on future satisficing as well as the limited predictive power of individual-level characteristics indicate that while individual-level characteristics play a role, situational factors also substantially influence satisficing. This implies that targeted interventions should not rely solely on past behavior but should be enhanced with real-time data and the application of learning algorithms. Fifth, the finding that even optimizers have non-negligible risks of engaging in undesirable response behavior suggests that innovative prevention methods should target not only the extreme cases but also individuals who are more under the radar but potentially more approachable.

TUESDAY, 01 APRIL 2025
12:00 - 01:15 PM
MAX-KADE-AUDITORIUM

GOR THESIS AWARD II: PHD

PRESENTATIONS

IDENTIFYING, CHARACTERIZING, AND MITIGATING ERRORS IN THE AUTOMATED MEASUREMENT OF SOCIAL CONSTRUCTS FROM TEXT DATA

INDIRA SEN University of Mannheim, Germany

In this abstract, I summarize the contents and contributions of my doctoral thesis “Identifying, characterizing, and mitigating errors in the automated measurement of social constructs from text data”. My doctoral research is highly interdisciplinary and sits at the intersection of Natural Language Processing (NLP) and Social Science.

RELEVANCE Computational Social Science (CSS) brings a transformative approach to social science, leveraging digital trace data—data generated through online platforms and digital activities [1]. This form of data is distinct from traditional social science data like surveys, providing vast, real-time information on social constructs, or human behavior and attitudes.

Given the large-scale nature of digital traces, new methods of analysis are needed to draw insights from them; particularly automated techniques using NLP and Machine Learning. Yet, despite their advantages, research with digital traces and automated methods also present unique challenges [2]. Digital trace data lacks well-established validity measures and faces biases from platform-specific artifacts,

incompleteness, and methodological gaps. Traditional social science emphasizes robust frameworks to minimize biases in survey data, but such practices are underdeveloped for digital trace data, making the rigorous study of its limitations critical. By systematically identifying and addressing these issues, we can advance the field and offer more reliable insights into complex social constructs like political attitudes, prejudicial attitudes, and health-related behaviors. This is what I set out to do in my doctoral dissertation, by exploring and answering the following research questions:

RESEARCH QUESTIONS

1. RQ1: What errors and biases arise when using digital trace data to measure social constructs?
2. RQ2: Can survey items enhance the validity of computationally measured social constructs from digital trace data?
3. RQ3: Can both manual and automated data augmentation techniques improve the robustness and generalizability of these computational measurements?

METHODS & DATA

DATA I use various online data sources in this dissertation and combine survey data and survey scale items. For RQ1, we conduct a scoping review of research that uses Twitter/X, Wikipedia, search engine data, and others. For RQ2, we utilize data from Twitter/X to study sexism, and Glassdoor, a platform where employees review their workplaces, to study workplace depression. To incorporate social theory into our computational models for RQ2, we also use survey items for sexism [3, inter alia] and workplace depression [4], as well as survey-based estimates of depression in the US for validation. For RQ3, we use data from Twitter/X, Reddit, and Gab to study sexist and hateful attitudes.

METHODS The study combines computational techniques, specifically NLP and Machine Learning (ML), with quantitative and qualitative approaches rooted in social science. For RQ1, we use a combination of literature review, case studies, and qualitative analysis to conceptualize an error framework inspired by traditional survey methodology, particularly the Total Survey Error (TSE) Framework [5]. Our framework, the “Total Error Framework for Digital Traces of Human Behavior on Online Platforms” (TED-On) adapts the TSE to address idiosyncrasies of digital traces, e.g., the effect of the online platform. Our framework helps in systematically identifying errors when using digital trace data for measuring social constructs.

For RQ2, we incorporate theory into computational NLP models in two ways — using survey items to guide the creation of labeled training data which is then used to train computational models, or to use them with sentence embeddings [6] for a semi-supervised approach. For RQ3, we use manual and automated data augmentation to create better NLP models. We use Large Language Models (LLM) like GPT3.5 for automated synthetic data. Finally, we also devise robust evaluation approaches that specifically account for the generalizability of computational methods.

RESULTS RQ1: Through the TED-On framework, two main error types are highlighted: (1) measurement errors due to content misalignment with constructs and (2) representation errors due to biased or incomplete data coverage.

RQ2: survey-inspired codebooks and models help structure the analysis of digital trace data around specific and holistic theoretical dimensions. Concretely, we find that models developed on our theory-driven

codebook outperform existing state-of-the-art models by 6% higher F1 scores. Our workplace depression classifier was also validated and found to correlate with state-level depression scores ($r = 0.4$).

RQ3: Manual and automated data augmentation techniques increase computational models' robustness, especially in cross-domain applications. For example, synthetic data created to detect sexism and hate speech resulted in models that generalize better across platforms, minimizing reliance on platform-specific artifacts. Both manual and LLM-generated augmented data improve the out-of-domain generalizability of computational models with improvements of 5-12% F1 scores for different domains, compared to scores of around 55% F1 from previous models.

ADDED VALUE This thesis makes significant contributions to three foundational aspects of CSS:

1. **Theory:** The research introduces measurement theory adapted for digital traces, creating a shared vocabulary for the interdisciplinary field. The TED-On framework aids in error documentation and model validation tailored to digital data challenges.
2. **Data:** This work develops several datasets: a Twitter/X dataset on sexism, synthetic training data for detecting sexism and hate speech, and a dataset detailing workplace depression rates across companies and states, offering valuable resources for CSS research.
3. **Methods:** The research introduces theory-driven and generalizable NLP models for identifying sexism and hate speech, and semi-supervised models for analyzing workplace depression.

These contributions provide a roadmap for future CSS studies which could further refine social construct measurement from new digital trace data sources

REFERENCES

1. Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A. L., Brewer, D., ... & Van Alstyne, M. (2009). Computational social science. *Science*
2. Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*
3. Glick, P., & Fiske, S. T. (2018). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. In *Social cognition*
4. Bianchi, R., & Schonfeld, I. S. (2021). The occupational depression inventory—a solution for estimating the prevalence of job-related distress. *Psychiatry Research*,
5. Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public opinion quarterly*
6. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Conference on Empirical Methods in Natural Language Processing*.

THE POWER OF LANGUAGE: THE USE, EFFECT, AND SPREAD OF GENDER-INCLUSIVE LANGUAGE

ANICA WALDENDORF Nuffield College, University of Oxford, United Kingdom

RELEVANCE & RESEARCH QUESTION This dissertation investigates a new social phenomenon: the use of gender-inclusive language (GIL). Leveraging language as a strategic research site, this paper-based dissertation contributes to understanding a current social phenomenon in Germany and advances the sociological understanding of how behavioural change occurs and is experienced. It creatively combines different online research methods (web scraping, text analysis, a digital field experiment, and qualitative online interviews). GIL refers to changing person nouns to be gender-inclusive, akin to the shift from policemen to police officers in English. In German, a researcher is a male researcher (Forscher) or a female researcher (Forscherin). A gender-inclusive alternative would be Forscher*in. The convention is to use the masculine form generically, which holds a strong and empirically documented male bias. In 2020, GIL was a relatively new and highly discussed topic in Germany. Everyday observations indicated an increase in GIL use (e.g. in iOS software and Spotify), yet academic sources viewed it as a marginal phenomenon. GIL is a potential equal opportunities tool owing to extensive research documenting the ability of GIL to mitigate the male bias in language. From a sociological perspective, it provides an opportunity to retrospectively study behavioural change using large-scale data without researcher interference.

This dissertation asks:

PAPER 1: Has GIL been as rapidly adopted as it seems? Under what conditions can difficult behavioural change take place?

PAPER 2: Does GIL increase the number of women who apply for a male-typed task? Do women perform better at a male-typed task when GIL is used?

PAPER 3: How has GIL use developed since its initial increase? How do individuals experience GIL?

METHODS & DATA Paper 1 uses web scraping to craft two unique datasets that were then analysed using quantitative text analysis, part-of-speech tagging, and manual annotation. Using a programme written in Python, I accessed the Deutscher Referenzkorpus (DeReKo) to collocate GIL occurrences in over 4 million newspaper articles published in five different media outlets between 2000 and 2021. This measured the frequency of GIL but not of the generic masculine. Therefore, I also web-scraped the five different media outlets to gather full-length newspaper articles, which I then annotated to measure the relative use of GIL. I conducted differential analyses by the political orientation of the media outlet (left, centre, right), the type of GIL (10 different types), and the author's gender (identified using part-of-speech tagging).

Paper 2 is co-authored with Klarita Gërkhani and Arnout van de Rijt. It is a digital field experiment that tests the effect of GIL, specifically in job applications where previous research has demonstrated an increase in girls' and women's attitudes and preferences towards stereotypically masculine jobs when GIL is used. We use Prolific, an online crowd-working platform, as an experimental labour market, where sign-up for our advertised task itself was an outcome variable.

We advertised participation in a stereotypically masculine task (solving maths problems) and varied the use of GIL in the advertisement (on Prolific) and the task description (on Qualtrics) in a 2x2 between design, which allowed the separation of a recruitment effect and a performance effect. We then measured how many women participated under each condition (two-tailed proportion test) and how well women performed when GIL was used (two-tailed independent t-tests). The experiment was pre-registered on OSF (<https://osf.io/x8ft4>) with a sample size of 2,000 based on power calculations. However, the participant pool was exhausted at 1,321 participants. It was fielded in Germany and Italy.

As GIL is a relatively new behaviour, Paper 3 revisits the DeReKo data for the years 2022/23 and combines it with data from Google Trends and the Dow Factiva database to see not only how GIL use develops after its initial increase but also whether and how it is talked about. It then builds on the quantitative findings with prospective longitudinal qualitative interviews in combination with ego-centric network data collection (21 cases). The qualitative interviews were conducted online using Microsoft Teams, an online tool that was essential as 1) it allowed access to a geographically diverse group of research participants and 2) for participants to engage in participatory research by using the whiteboard function to let them draw their network.

RESULTS

PAPER 1: In addition to observing an unexpectedly rapid increase in GIL (reaching 800 occurrences per million words, or 16.5% of potential use), two different trends are identified: whilst non-binary inclusive forms of GIL are increasingly used in the left-leaning newspaper, GIL that adheres to a binary notion of gender is favoured in the mainstream and right-leaning media. Three conditions for difficult behavioural change are identified: having a role model, the possibility of a low-threshold adoption, and incremental adoption.

PAPER 2: We find no effect of GIL on the share of women nor on the performance of women. In each condition, the share of women was 43%, and there was no statistically significant difference in women's mean performance (mean number of correct answers in math task was 3.3 with GIL, 3.5 without GIL). This may be because GIL only influences attitudes, not behaviours. It may also be an artefact of our research design, so we are planning a follow-up experiment.

PAPER 3: The data show that since 2021, the use of GIL has stalled, and the contestation of GIL has spiked. The current situation is what I refer to as incipient change: GIL has increased but has not replaced the generic masculine. Yet, it has also not waned. Turning to the interview data, I uncover a fragmented understanding of GIL (i.e., everyone knows GIL, but not everyone correctly understands what GIL is) alongside a surprisingly clear individual understanding of the circumstances in which GIL can or sometimes even should be used.

Rather than a general separation of GIL users and non-GIL users (polarisation), I argue that the use of GIL seems to be strongly tied to context, particularly the professional context, thus mapping onto a coordination-based micro-level theory.

ADDED VALUE This thesis combines multiple online research methods to study different facets of the same highly relevant social phenomenon: GIL. First, it shows the strength of using web scraping and computational tools to measure macro-level patterns and using digital tools to access the micro level, underscoring the importance of online research not just for quantitative research but also qualitative. Sec-

ond, it demonstrates how the shift of the labour market into the digital sphere enables new research designs and, with them, a new avenue of research possibilities.

WHO DONATES THEIR GOOGLE SEARCH DATA? PARTICIPATION IN A DATA DONATION STUDY DURING THE 2025 GERMAN FEDERAL ELECTION

SINA CHEN GESIS, Germany
BARBARA BINDER GESIS Germany Abstract

RELEVANCE & RESEARCH QUESTION Data donation is a relatively new, user-centered approach to collecting digital trace data, increasingly relevant in various research fields (Haim et al., 2023). It offers significant potential to understand online behavior, especially amid growing API restrictions, as it can help validate and enrich survey data. However, little is known about who participates in these studies (Keusch et al., 2024). Previous results hint that participation is non-random (Welbers et al., 2024). Our study aims to identify correlates of participation, moving beyond sociodemographics to consider political preferences and vote choice.

METHODS & DATA Wave 60 participants from the GLES Tracking pre-election survey ($n \approx 2,000$, CS, CAWI, online access panel) are invited after the 2025 German Federal Election to complete a follow-up survey on their vote choice, followed by a request to donate their Google search histories from six weeks before to one week after the election. Besides search terms, we collect URLs of clicked search results and the order of searches and clicks. We experimentally vary the invitation framing to evaluate how different descriptions of the collected data ("web services", "data collected by Google", "data collected by Google Search") influence participation rates. The digital trace data will be linked to both survey datasets to analyze participation across sociodemographics, political attitudes, electoral preferences, and vote choice. Our design captures multiple participation stages, from initial opt-in to final data donation.

RESULTS Data is collected one week after the 2025 German Federal Election (February 23, 2025). At the GOR conference, we will present results on participation rates by invitation framing, sociodemographics, political preferences, and vote choice, differentiating between respondents who opted in, consented, processed, and ultimately donated their data.

ADDED VALUE This study provides insights into how sociodemographic characteristics relate to participation in data donation studies. Additionally, we examine political preferences and vote choice, as well as how the specificity of the data requested in the invitation affects willingness to participate. Since Google search data is highly private, our findings will inform optimal framing for requests involving sensitive data. Our study advances methodological approaches in digital behavioral data research and provides practical guidelines for designing effective data donation studies.

ABSTRACTS

TUESDAY, 01.04.2025
10:45 - 11:45 AM
HÖRSAAL A

2.1: INNOVATIONS IN MARKET RESEARCH

SESSION CHAIR: GEORG WITTENBURG
INSPIRIENT, GERMANY

PRESENTATIONS

SHOW ME, HOW YOU TOUCH IT – THE EFFECTS OF VICARIOUS TOUCH IN ONLINE MARKETING VIDEOS

SEBASTIAN ZIAJA, CHRISTIAN BOSAU Rheinische Hochschule Köln
gGmbH, Germany

RELEVANCE & RESEARCH QUESTION Overcoming key purchase barriers and improving consumer experience is an ongoing trend in online shopping. Therefore, this research analyses the influence of vicarious touch in the digital context. It is assumed that the depiction of a hand touching the product can compensate for online shoppers' need for haptic information. Initial studies have already shown that this strategy leads to mental simulation of a product interaction and improved product evaluation [e.g. Luangrath et al., 2022; Liu et al., 2019], although no clear distinction has yet been made between product images and videos. Referring to the underlying theoretical framework of MNS activation (Adler & Gillmeister, 2019), it is assumed that the actual perception of movement is a prerequisite for such an effect, and therefore vicarious touch must always be presented in video format.

METHODS & DATA The experimental study uses quantitative data (N = 1000) from an online access panel, nationally representative in terms of age and gender. In a social commerce context, respondents reported their purchase intentions for four different products selected in a pre-study that differed in their haptic importance. The two between-factors touch (vicarious touch vs. no touch) and media format (image vs. video) were realized, resulting in four different product presentation types.

RESULTS Indeed, only for product videos did the use of vicarious touch significantly increase purchase intent compared to static images ($p < .001$, $d = .3$) and rotating product videos ($p < .05$, $d = .2$). There was no effect when vicarious touch was presented in static images. Interestingly, the video format itself was a significant driver of purchase intent, highlighting the overall benefit of dynamic product presentation ($p < .001$, $\eta p^2 = .012$). Furthermore, results did not differ on different

products, suggesting that online shopping generally creates a haptic information gap.

ADDED VALUE Using realistic e-commerce scenarios, this study shows that vicarious touch can significantly increase purchase intent, but only in product videos, not images. The findings highlight the need for dynamic, motion-rich content to compensate for the lack of haptic feedback online, making products more tangible and driving consumer engagement.

DATA-DRIVEN DECISION-MAKING IN REAL ESTATE WITH TENANTCM: UNLOCKING VALUE FROM TENANT SATISFACTION SURVEYS

**MATTHIAS DAVID KELLER, INNA BECHER, VERENA MACK, MARC
HERTER, JUSTUS RATHMANN, NICOLE HÄNZI** YouGov Schweiz AG

RELEVANCE & RESEARCH QUESTION Tenant satisfaction is a key performance indicator in the real estate industry, driving tenant retention, revenue stability, reputation, and operational efficiency. As tenant expectations evolve, it is crucial to efficiently capture feedback, address concerns, and benchmark performance against industry standards. Beyond benchmarking, the true value of tenant satisfaction surveys lies in generating actionable insights for decision-making at both the property and individual tenant levels. This raises the question: How can tenant satisfaction surveys be enhanced to provide detailed benchmarking opportunities and granular analysis for informed decision-making?

METHODS & DATA In collaboration with over 20 property owners across residential and commercial segments, we conducted tailored online tenant satisfaction surveys in Switzerland from 2019 to 2024 leading to nearly 75'000 completed surveys. Invitations were sent via email, postal mail, or tenant apps, with responses processed into a unified dashboard. This platform enables clients to benchmark performance against aggregated market results and allows property managers to compare performance across owners. Key challenges, such as identifying stakeholders and improving response rates, were addressed through targeted approaches. Actionable insights were ensured using multi-level evaluation, NLP for analysing open text comments, and structural equation modelling (SEM) to identify key drivers of satisfaction. Annual workshops with clients further refine survey design, dashboards, and workflows.

RESULTS Our methodology enabled dual-layer evaluations, offering detailed assessments at the tenant level and aggregated insights at property and portfolio levels. Property owners tracked satisfaction trends over time and benchmarked against market standards, while property managers addressed specific tenant concerns, expressed via numeric ratings and open comments, and compared performance across clients. Results are presented in an interactive dashboard with a measure management tool for tracking issues and enabling targeted improvements.

ADDED VALUE By advancing traditional survey methodologies, our approach places tenants at the centre of decision-making. The unified dashboard delivers tailored, comparable insights across residential and commercial segments. NLP and SEM efficiently identify key sat-

isfaction drivers, enabling timely interventions. Annual workshops ensure tools evolve with client needs, fostering tenant loyalty and supporting sustainable portfolio management.

BEYOND REPORTS 2.0: ENHANCING CUSTOMER SEGMENTATION WITH RETRIEVAL-AUGMENTED PERSONA BOTS

THEO GERSTENMAIER Factworks, Germany

RELEVANCE & RESEARCH QUESTION Segmentation research is invaluable for understanding diverse customer needs, but its impact depends on effective socialization within organizations. At last year's GDR conference, we introduced an AI-driven persona chatbot we developed using OpenAI's custom GPT functionality, aimed at making segmentation insights more interactive and accessible. However, inconsistent results and system issues limited its effectiveness. This time, we investigate whether a Retrieval-Augmented Generation (RAG) model can overcome these challenges and improve the accuracy, consistency, and practical application for segmentation insights. Our key question: Can RAG-based persona bots become a new way to communicate research insights, ultimately supporting strategic decision-making?

METHODS & DATA Our approach leverages a RAG model that combines a generative language model with a retrieval system. This approach limits the bot to only reference set data repositories (in our case segmentation research data) for more accurate factually-grounded responses. To evaluate its performance, we compare two data input approaches: an aggregated dataset representing segments' summarized characteristics versus individual respondent-level data. For each approach, we create a RAG-based chatbot trained on research data (survey data and persona descriptions) from a fictitious study on traveler segments.

RESULTS First testing suggests that the RAG model enhances the chatbot's ability to retrieve factual knowledge from the segmentation data with far fewer incorrect or fabricated responses than our previous model. Our hypothesis is that the aggregated data approach will offer more generalized insights into segment characteristics, while the respondent-level approach may yield more nuanced responses that reveal variations within each segment. This hypothesis and the general effectiveness of the RAG-based bot will be tested further, with results available for the conference.

ADDED VALUE This refined RAG-driven bot approach could represent a shift in how segmentation research is utilized across organizational functions. By providing reliable, interactive insights into customer segments—and allowing users to ask the bot questions beyond the original survey scope—this tool has the potential to increase engagement with and adoption of research findings. This innovation brings segmentation insights closer to teams across the organization, transforming them into an engaging, and strategic resource that extends well beyond traditional reporting.

TUESDAY, 01.04.2025
10:45 - 11:45 AM
HÖRSAAL B

2.2: WEB TRACKING

SESSION CHAIR: **DORIAN TSOLAK**
BIELEFELD UNIVERSITY, GERMANY

PRESENTATIONS

UNDERSTANDING PARTICIPATION IN WEB TRACKING STUDIES: A COMPARISON OF PROBABILISTIC AND NONPROBABILISTIC SAMPLING STRATEGIES

JOACHIM PIEPENBURG, BERND WEISS, SEBASTIAN STIER, FRANK MANGOLD, JUDITH GILSBACH, BARBARA BINDER GESIS, Germany

RELEVANCE & RESEARCH QUESTION Web tracking has become a critical method for studying digital behavior. However, most data collection efforts in this area face significant data quality issues, particularly due to limitations in sampling design that hinder representativeness and generalizability.

Additionally, factors influencing participation in web tracking are poorly understood, as is the precise definition of participation. This study addresses these gaps by (1) identifying key factors across three stages of web tracking participation – consent, tool installation, and attrition – and (2) examining whether these factors differ across probabilistic and nonprobabilistic sampling strategies.

METHODS & DATA This study draws on data from the GESIS Panel.dbd, which incorporates various panel recruitment strategies. Probabilistic recruitment includes sampling from the official population register and leveraging existing studies such as the 2023 German General Social Survey (ALLBUS). Nonprobabilistic recruitment involves multiple ad campaigns on META platforms (e.g., Facebook, Instagram) and open referrals. Following recruitment, participants were invited to a web tracking study, and we conducted multiple regression analyses to compare factors influencing participation across these sampling approaches.

RESULTS Our findings reveal that different factors influence each stage of web tracking participation, with distinct patterns emerging for consent, installation, and attrition. Additionally, recruitment strategy significantly impacts participation, with stable differences between probabilistic and nonprobabilistic samples, indicating persistent unobserved biases and underscoring the role of sampling design.

ADDED VALUE This study advances existing research by exploring web tracking participation factors in both probability and nonprobability samples. Our findings provide actionable insights for researchers and practitioners to enhance data quality and generalizability in web tracking studies, thereby strengthening the reliability and applicability of findings in digital behavior research.

SOCIOECONOMIC STATUS AND PATTERNS OF ONLINE BEHAVIOR IN GERMANY

BARBARA BINDER GESIS Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Do individuals from different socioeconomic status (SES) groups use the internet differently? The digital divide extends beyond disparities in access to adequate internet access. Being online offers potential benefits, but people may differ in their knowledge, opportunities, and capabilities to take full advantage of these benefits.

METHODS & DATA Using a linked dataset from the German General Social Survey (a large, probability-based survey) and respondents' web surfing behavior (GESIS Web Tracking), this study explores whether online behavior varies by SES background. Respondents participated in a web tracking study, which collected data on every individual website visit over two months following the installation of a browser plugin. This pilot study includes more than 4 million website visits from 500 participants, with the linked data providing around 340,000 website visits from 106 respondents. The websites visited were classified into content-based categories, such as "education," "job search," "healthy living," "personal finance," alongside categories like "shopping," "sports," and "video gaming" using two different third-party service providers. Through regression analysis, I examine whether SES is associated with particular types of website visits and whether this relationship is moderated by first- and second-level digital divide factors, such as access to fast internet connections and digital literacy.

RESULTS Preliminary results based on the webtracking data set show that characteristics of participants such as their education and political interest are associated with more frequent visits of particular website types such as news websites. Linking of the data sources will be possible next week. I will be able to show results on differences in online behavior by SES at the conference.

ADDED VALUE Differing online behavior may ultimately contribute to inequalities in education, the labor market, or financial well-being, potentially mitigating or reinforcing existing social inequalities. Thus, understanding these behavioral differences is crucial for reducing structures that exacerbate inequality. Results of this work may demonstrate the substantive value of linked survey and web tracking data.

BRIDGING GAPS OR DEEPENING DIVIDES? THE IMPACT OF ONLINE INTERMEDIARIES ON NEWS DIVERSITY

FELIX SCHMIDT Department of Computational Social Science, GESIS – Leibniz Institute for the Social Sciences

RELEVANCE & RESEARCH QUESTION Recent research demonstrated that intermediaries like Facebook, Twitter, search engines and news aggregators can broaden the diversity of news people consume. However, the personalized content users encounter on intermediaries remains a black box, as tracking tools have been limited in their ability to capture in-platform content. A central debate persists around whether algorithmically curated content diverges from preference-driven selective exposure -- where users actively choose to engage with specific news items based on their interests. Using a web tracking tool that captures the public content on Facebook and Twitter as well the content encountered on other websites, this study examines whether content exposure through intermediaries affects the diversity of news accessed by German internet users.

METHODS & DATA This study examines three months of web browsing histories, including survey responses, from a sample of German internet users (N = 739) to investigate how the use of intermediaries -- and the diversity of content encountered on these platforms -- affects direct visits to news websites and the diversity of news encountered. The analysis uses random-effects within-between (REWB) models, with data hierarchically structured at individual and daily levels.

RESULTS Preliminary findings align with existing research, demonstrating that engagement with intermediaries is positively associated with greater diversity in news exposure. Further, results show that individuals who engage with intermediaries tend to have richer and more diverse news diets within intermediary platforms compared to the diversity encountered through direct visits to news websites. However, this increased diversity comes with a caveat: users are also more likely to encounter hyperpartisan news through these platforms.

ADDED VALUE This study offers insights into the role of intermediaries in the dissemination of information and their potential impact for information diversity in digital environments. By providing a nuanced perspective on the mechanisms driving news diversity, it advances the field's understanding of the relationship between online intermediaries and diverse media diets.

TUESDAY, 01.04.2025
10:45 - 11:45 AM
HÖRSAAL C

2.3: VOTING BEHAVIOR AND INFORMATION SOURCES

PRESENTATIONS

THE IMPACT OF VOTING ADVICE APPLICATIONS ON VOTING BEHAVIOR: EVIDENCE FROM THE 2024 AUSTRIAN ELECTIONS

KATHARINA PFAFF, SYLVIA KRITZINGER University of Vienna, Austria

RELEVANCE & RESEARCH QUESTION Voting advice applications (VAAs) are digital tools providing personalized voting advice based on the match between party positions and the user's opinion. As such recommendations can influence decision-making processes and bear implications for electoral outcomes, understanding how VAAs influence voting behavior is essential for assessing the role of technology in strengthening or weakening democratic processes. This study addresses the following research question: "How do voting advice applications affect voting behavior?"

METHODS & DATA The quantitative analysis relies on novel data from a probability-based online panel in Austria, where national election was held on September 29, 2024. Using data from a post-election survey conducted between September 30 and October 21, this study examines individual factors influencing the use of various VAAs available during the election, voter motivations for using these tools, and their on vote choice.

RESULTS Data from this recent survey is currently being prepared for analysis. Preliminary analysis using data from up to n=1,687 panelists suggest that approximately one third of the respondents used at least one of the VAAs available for this election. A large share of respondents used VAAs to test whether the result corresponds to the party they intend to vote for. The authors will update more detailed results by January 2025.

ADDED VALUE Prior research suggests that VAAs facilitate decision-making by reducing costs of gathering political information. As a result, users are more likely to cast a ballot in elections as a result. The added value of this analysis lies in re-assessing this link using novel survey data in the context of the 2024 Austrian national elec-

tion. Unlike prior studies, it provides new insights into a setting, in which more than one VAA was available. This study thus offers a unique opportunity to assess the relationship between VAA usage and voter behavior in the most recent elections.

ASSESSING VOTER FATIGUE: MEDIA CONSUMPTION AND POLITICAL ENGAGEMENT ACROSS ISRAELI ELECTION CYCLES (2019-2022)

DANA WEIMANN SAKS, YARON ARIEL, VERED ELISHAR The Max Stern Yezreel Valley College, Israel

RELEVANCE & RESEARCH QUESTION This study examines voter behavior under conditions of frequent elections and political instability, focusing on Israel's unprecedented period of five national elections between 2019-2022. Analyzing this unique case of democratic stress, we investigate how recurring election cycles influence Israeli voters' media consumption and engagement patterns, and explore their relationship with voting intentions. The insights contribute to understanding the interplay between political volatility, media consumption, and democratic participation.

METHODS & DATA The study employed survey data from 2,000 Israeli participants recruited through the Midgam Project Web Panel, using stratified sampling aligned with Israeli Central Bureau of Statistics demographics. Data collection spanned four election cycles between 2019-2022, with surveys conducted prior to each election. Variables measured included traditional and digital media consumption patterns, social media engagement with political figures, and changes in voting intentions.

RESULTS Analysis revealed distinct trends across the 2019-2022 election cycles. Traditional media consumption peaked in 2019 but declined significantly in subsequent elections, reaching its lowest point in 2021 before a modest recovery in 2022. Digital media consumption showed steady growth throughout the period. Social media engagement with political figures exhibited a complex pattern: after decreasing in the second and third election rounds, it rebounded in 2022. Notably, logistic regression analysis indicated that while higher general social media consumption correlated with stable voting intentions, tracking politicians across multiple platforms significantly increased the likelihood of voting intention changes.

ADDED VALUE This research provides novel insights into voter behavior under conditions of repeated elections, challenging assumptions about voter fatigue in highly contested democratic environments. The findings demonstrate how different forms of media consumption influence political engagement and voting stability, particularly highlighting social media's nuanced role in shaping electoral behavior. These results have important implications for understanding democratic participation during periods of political instability and inform strategies for maintaining voter engagement in similar contexts.

**CLICK FOR CLARITY?
EXAMINING THE EFFECT OF OPTIONAL
INFORMATION ON PREDICTION
ACCURACY IN SWISS REFERENDA**

VERENA MACK, FABIAN BERGMANN YouGov Schweiz AG, Switzerland
SUSUMU SHIKANO, STEFFEN STELL Universität Konstanz, Germany

RELEVANCE & RESEARCH QUESTION Regular referenda are a central aspect of Swiss direct democracy. Due to the frequency of referenda and their diverse subject matters, public opinion on individual issues often solidifies only shortly before the voting date. Eligible voters commonly rely on official materials, such as booklets or the VotelInfo app, which provide condensed information to support their decision-making. Our study explores whether simulating this information process in a survey setting increases the predictive accuracy of survey data for referendum outcomes. Specifically, we investigate whether providing official referendum information through info buttons helps survey participants not only to become more decisive but also to shape or adjust their stated voting intentions.

METHODS & DATA We designed an experiment embedded in a survey on voting intentions in Swiss referenda. Around 5000 participants were randomly assigned to one of two groups: a control group that received no additional information, and a treatment group that could access concise official summaries, including voting recommendations from the Swiss Federal Council, Parliament, referendum committees, and party positions. This survey covers recent referenda in 2024 and the upcoming national referendum on 9 February 2025, addressing issues such as the Environmental Responsibility Initiative.

RESULTS Preliminary findings suggest that participants exposed to additional information were less likely to report uncertainty about their voting intentions. Significant differences emerged in the likelihood of intending to vote “yes” or “no” based on whether participants viewed the information. Moreover, participants who opted out of additional information displayed stronger political interest, greater confidence in casting a vote, and a lower perception of political system complexity. Our preliminary findings thus support the expectation that providing additional information helps voters with initially less definite voting intentions to demonstrate greater decisiveness and consistency, ultimately increasing the predictive accuracy of the survey data.

ADDED VALUE This study highlights how tailored, optional information fosters voting confidence and decision clarity, enhancing the reliability of pre-referendum survey predictions. By examining the integration of information aids into surveys, it provides methodological insights for improving survey-based forecasting and deepens the understanding of political engagement in direct democracy.

**TUESDAY, 01.04.2025
12:00 - 01:15 PM
HÖRSAAL A**

3.1: VIRTUAL INTERVIEWING

**SESSION CHAIR: TANJA KUNZ
GESIS – LEIBNIZ INSTITUTE FOR THE SOCIAL SCIENCES,
GERMANY**

PRESENTATIONS

**VIDEO-INTERVIEWING AS PART OF
A MULTI-MODE DESIGN IN PANEL
STUDIES: INSIGHTS FROM THE FIELD**

**JULIA WITTON 1, CARINA CORNESSE 2,
MARKUS GRABKA 1, SABINE ZINN 1,3** 1 Deutsches Institut für
Wirtschaftsforschung (DIW Berlin); 2 GESIS - Leibniz-Institut für
Sozialwissenschaften; 3 Humboldt-Universität zu Berlin

RELEVANCE & RESEARCH QUESTION To ensure continued survey participation and data quality among respondents, the panel survey landscape must adapt to the changing societal reality, especially in terms of mobility and digitalization. Computer-Assisted Web-Interviewing (CAWI) has been shown to be a useful self-administered and cost-effective survey mode. Nevertheless, in surveys with complex instruments, interviewer assistance may still be necessary. Given a growing prevalence of videoconferencing in many people’s lives, Computer-Assisted Live Video Interviewing (CALVI) has potential to be a solution.

METHODS & DATA To develop a high-usability CALVI methodology and examine fitness-for-purpose of this mode across societal subgroups, we have gathered data on panel respondents’ willingness to try CALVI in the Socio-Economic Panel (SOEP), developed and pretested an implementation of CALVI together with ifas Institute for Applied Social Science, and started fieldwork in an experiment comparing CALVI to online self-administered and offline face-to-face modes in the context of the SOEP Innovation Sample (SOEP-IS).

RESULTS To assess the potential of CALVI in household panel surveys in Germany, we introduced a hypothetical inquiry regarding respondents’ willingness to be surveyed via videocalls in the SOEP 2022 data collection wave. Of the 22,549 respondents who provided a valid answer, 39% expressed their willingness to try CALVI. Based on the results, we designed a CALVI implementation trial experiment for the SOEP Innovation Sample. In initial pretests, 600 participants of an ifas panel study were invited to try video interviews. In total, 73 of those made an appointment for the interview and 46 participated in the interview. On average, the interviews were 106 minutes long and were conducted by experienced and CALVI-trained ifas interviewers.

Our presentation will provide more in-depth results from the hypothetical scenario and the pretest as well as lessons learned from the experimental fieldwork.

ADDED VALUE The results of our feasibility project will inform future implementations of CALVI, especially with regard to usability for both respondents and interviewers as well as fitness-for purpose of the new mode for different population subgroups. Furthermore, our results will be used to develop targeted multi-mode strategies for data collection in SOEP-IS.

WILLINGNESS TO PARTICIPATE IN SURVEYS ADMINISTERED BY SMART SPEAKERS

CEYDA DEVECI, MAREK FUCHS, ANKE METZLER Technical University of Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION The Internet of Things (IoT) is a rapidly evolving technology, bridging the physical and virtual worlds [Zhang, 2021]. By integrating Internet-connected devices supported by hardware or software, IoT enables the emergence of smart devices with communication capabilities.

These capabilities create a new form of interaction between humans and computers, enabling new possibilities in survey research. Smart speakers could conduct interviews more cost-efficient than face-to-face interviews and potentially improve data quality compared to web surveys. However, it is unclear whether people are willing to participate in interviews administered by smart speakers. This study examines socio-demographic factors associated with the willingness to participate in such interviews.

METHODS & DATA Data was collected from two non-probability online access panels in the U.S. in Spring 2024 and a similar online panel in Germany in November 2024 with samples representing the general English-speaking U.S. population aged 18 and older and the respective German counterpart. Each country had around 2,000 respondents. The analysis included socio-demographic characteristics like gender, age, ethnicity, education, income, and metropolitan status. Additionally, participants' attitude on the importance of surveys were considered to explore factors influencing willingness to participate.

RESULTS In the U.S., 25% of respondents are "very likely" to participate in a survey conducted by a smart speaker. Willingness was significantly influenced by factors such as age, ethnicity, gender and attitudes towards survey importance. Younger respondents, males, African Americans and people who placed higher importance on surveys were significantly more likely to report willingness to participate. Analyses of the German data is ongoing to compare the willingness and the factors influencing the willingness between the two countries.

ADDED VALUE This study underscores the selective appeal of smart speaker surveys, suggesting that while not universally preferred, they may reach traditionally underrepresented groups. Smart speaker surveys can extend survey methodologies for technology-affine participants in mixed-mode designs. Although current acceptance varies across demographics, investment in this approach is advis-

able, as its applicability is expected to broaden with the advancing IoT technology.

DATA QUALITY INVESTIGATIONS OF ONLINE LIVE VIDEO INTERVIEWING: EMPIRICAL EVIDENCE FROM SEVERAL MAJOR UK SOCIAL SURVEYS

GABRIELE DURRANT University of Southampton, United Kingdom
SEBASTIAN KOCAR University of Queensland
TIM HANSON, MATT BROWN, CAROLE SANCHEZ UCL, University College London
MARTIN WOOD, KATE TAYLOR, MARIA TSANTANI NatCen
TOM HUSKINSON Ipsos

RELEVANCE & RESEARCH QUESTION Many surveys in the UK are transitioning to online data collection. However, long surveys or those involving complex elements, such as data linkage consent, cognitive assessments, and sensitive questions can be difficult to move to online self-administered data collection. As a result, in the UK several surveys explored online live video interviewing (VI), which represents a visual video-mediated, remotely interviewer-administered, and computerised survey mode.

This presentation investigates the use of VI, focusing on opportunities and barriers, and collating evidence from seven major social surveys in the UK, with an emphasis on longitudinal surveys and the collection of complex data. The research questions are:

1. What is the uptake of VI (response rates)?
2. What are the characteristics of those that respond via VI (representativeness)?
3. Are there advantages in using VI when collecting complex elements, in comparison to web only?

METHODS & DATA This study uses data from seven surveys conducted in the UK 2020-2023: National Child Development Study, British Cohort Study, Next Steps, Children of the 2020s, English Longitudinal Study of Ageing, European Social Survey and Health Survey for England pilot. We examine response rates, composition of VI samples and response to complex questionnaire items.

RESULTS One of the main findings is that VI in the UK is used either as the primary survey mode, or as a complementary mode in mixed-mode designs. VI as the primary data collection mode can lead to lower response rates and potentially also to an increase in representation bias. On the other hand, there are encouraging findings, including that this mode proves to be a very suitable approach for collecting complex elements. This is a key finding since previous research has identified limitations of other remote methods for collecting this kind of data, which is an important component of many studies, especially longitudinal studies.

ADDED VALUE This paper is the first to investigate the use of online VI in UK surveys. Particular benefits exist for longitudinal surveys. The collection of complex data is common practice in longitudinal studies and evidence presented here suggests VI performs well.

TUESDAY, 01.04.2025
12:00 - 01:15 PM
HÖRSAAL B

RESPONDENT ENGAGEMENT AND ATTRITION

SESSION CHAIR: ELLEN LAUPPER
 SWISS FEDERAL UNIVERSITY FOR VOCATIONAL EDUCATION
 AND TRAINING SFUVET, SWITZERLAND

PRESENTATIONS

ATTRITION PATTERNS AND WARNING SIGNS IN A LONG-TERM, HIGH FRE- QUENCY PROBABILITY ONLINE PANEL

TOBIAS RETTIG, ANNE BALZ University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION All longitudinal- and panel studies are confronted with the gradual loss of respondents, i.e., panel attrition. Over time, this leads to lower respondent numbers, loss of statistical power, and, if the attrition is nonrandom, systematic loss of specific respondents and thus biases in the remaining sample. Using data from a long term high-frequency panel, we investigate (1) which respondents are disproportionately lost over time and may be specifically targeted, oversampled, or the survey adjusted to facilitate participation, (2) what are warning signs of attrition, and (3) which survey features are associated with higher attrition and thus present opportunities for optimization.

METHODS & DATA Using data from a probability online panel of the German population, we analyze respondents' participation in over 70 panel waves spanning 12 years. This gives us the rare opportunity to analyze rich data from a high frequency and long running panel. Descriptively, we observe how the panel composition changes over time and which respondents are disproportionately lost. Using a survival model, we investigate risk factors for attrition on the respondent level, in their participation patterns, and in survey features.

RESULTS We observe high attrition over the first panel waves and slower but steady loss of respondents long term. Over time, the sample tends towards being higher educated, more likely to be married, and skews slightly more male. Attrition risk is lower for younger, higher educated, and full-time employed respondents. Higher attrition risk is associated with patterns of infrequent participation (breakoffs, missing panel waves, participating late during field time, item non-response), but not with interrupting the survey to continue later. Higher attrition risk is also associated with longer and poorly rated survey waves.

ADDED VALUE For panel practitioners, it is important to understand attrition patterns to accurately predict how many respondents to expect in future waves, when and how many respondents to recruit, and which groups should be specifically targeted or oversampled. We identify several groups that are at higher risk for attrition, early warning signs that may be used to counteract attrition with targeted interventions, and opportunities to optimize surveys for continued participation.

DO WE NEED TO INCLUDE OFFLINERS IN SELF-ADMINISTERED GENERAL POPULATION SURVEYS? AN ANALYSIS OF 150 SUBSTANTIVE VARIABLES IN A PROBABILITY-BASED MIXED-MODE PANEL SURVEY IN GERMANY

LENA REMBSER, TOBIAS GUMMER GESIS – Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Due to concerns about bias stemming from the undercoverage of non-internet users, most probability-based surveys try to include offliners (i.e., respondents not able or willing to participate online). Previous research shows that including offliners increases accuracy for some socio-demographic characteristics while not for others. These studies lack the inclusion of substantive variables. We aim to address this research gap by answering the following research question: Does the inclusion of offliners in a probability-based panel impact measures of substantive variables?

METHODS & DATA We use data from the GESIS Panel.pop, a probability-based self-administered mixed-mode panel of the German general population, surveyed via web and mail mode. We analyze around 150 substantive variables from six core studies collected since 2014, which we compare between the whole sample and a sample of only onliners (i.e., without offliners). To assess the impact of including offliners, we compute differences between both samples for each substantive variable and compute average absolute relative bias (AARB) for each variable and by (sub-)topic. In addition, we re-run these analyses for different definitions of onliners and offliners and for different recruitment cohorts.

RESULTS Data analysis is ongoing. Depending on the release of the GESIS Panel.pop data file version 54.0.0 in early 2025, which includes first Core Study data for the cohort recruited in 2023, we will have final results in early February 2025.

ADDED VALUE Our study contributes to the practical challenge of deciding whether to include the offline population in surveys by employing a costly and labor-intensive mail mode. We go beyond previous research by examining a wide range of substantive variables which will enable us to draw conclusions on topic areas in which including offliners is more warranted than in others. We expect our findings to be of relevance for survey practitioners and substantive researchers alike.

REDUCING POLITICAL SCIENCE SURVEYS' ATTRITION BIAS WITHOUT LIMITING SUBSTANTIVE RESEARCH: POTENTIALS OF ADAPTIVE SURVEY DESIGN AND MISSING DATA STRATEGIES

SASKIA BARTHOLOMÄUS GESIS – Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Adaptive Survey Designs that use respondents' topic interests can reduce the overrepresentation of politically engaged respondents in political science surveys. Politically disengaged respondents would receive a questionnaire combining political and non-political modules to boost participation, while politically engaged respondents receive a purely political questionnaire. However, assigning respondents based on political interest may distort research if variables of interest correlate with political interest. Instead, researchers could assign only a certain percentage of respondents with low political interest to tailored questionnaires and use missing data strategies based on main survey data to correct biased estimates. This paper aims to assess whether adaptive designs that rely on content variation bias substantive research and whether missing data procedures mitigate this bias.

METHODS & DATA Using the probability-based GESIS Panel, I simulate several datasets in which 50% to 100% of politically disengaged respondents' responses to various variables are set as missing. I then run several regression models using the original and simulated datasets to answer RQ1. To answer RQ2, I calculate the regression models using inverse probability weights and multiple imputation.

RESULTS Preliminary results suggest that adaptive survey designs that vary a questionnaires' content can bias substantive research if more than 50% of respondents are randomly assigned to question modules that are supposed to be more interesting. However, assigning only a certain percentage of respondents and applying missing data strategies instead reduces biased estimates and corrects confidence intervals.

ADDED VALUE While adaptive survey designs with content variation may reduce attrition bias in political science surveys, researchers must be cautious if variables of interest correlate with political interest as estimates may be biased. Assigning only a certain percentage of respondents to a question module that is supposed to be more interesting to them mitigates the risk of biased estimates when using inverse probability weights or multiple imputation.

QUANTIFYING QUESTIONNAIRE DESIGN: A HOLISTIC APPROACH TO DATA QUALITY AND USER ENGAGEMENT

EVA WITTMANN, CECILE CARRE Ipsos

RELEVANCE & RESEARCH QUESTION

The market research industry continues to face quality issues, often prioritizing the identification and removal of "problematic" respondents. However, this approach overlooks a crucial aspect of data quality: adapting questionnaires to align with respondents' lifestyles and enhancing their appeal to high-quality panel members. By focusing on survey design optimization, we can potentially improve data integrity at its source, rather than relying solely on cleansing methods.

This paper presents a quantitative perspective on enhancing data quality through optimized questionnaire design. Measuring questionnaire design effectiveness is complex due to the multitude of elements influencing respondent experience. While previous research has attempted to identify factors that make questionnaires difficult to answer within one research type and/or topic, we lack clear thresholds on what a good respondent is able to answer without risking diminishing response quality.

Our research addresses this gap by analyzing a comprehensive dataset from Ipsos' global operations over a multi-week period. We propose a questionnaire segmentation system, correlating design elements with Ipsos' quality indicators as a proxy for engagement. This approach allows us to identify measurable factors that minimize drops in respondent engagement and provide clearer recommendations for effective questionnaire structures. Importantly, we acknowledge the challenge of conducting research on research, given the significant influence of topic and interest levels on respondent behavior. Our methodology accounts for this variability, offering insights that are applicable across diverse research contexts.

METHODS & DATA Meta-study across Ipsos business

RESULTS Research to be finalised early 2025

ADDED VALUE By quantifying the impact of design elements on respondent engagement and data quality, with this study we hope to provide researchers with actionable strategies to create more respondent-centric surveys. Our findings contribute to a more holistic understanding of questionnaire design, moving beyond mere identification of problematic responses to proactively enhancing the overall survey experience.

TUESDAY, 01.04.2025
12:00 - 01:15 PM
HÖRSAAL C

3.3: DATA DONATION

SESSION CHAIR: JESSICA DAIKELER
GESIS, GERMANY

PRESENTATIONS

EXPLORING DATA DONATION THROUGH MOBILE COMPUTER DEVICES: A STUDY ON STUDENT USE OF AI TOOLS IN ACADEMIC CONTEXTS

DIMITRI PRANDNER Johannes Kepler Universität Linz
ANDREAS ZEMSAUER, DANIELA WETZELHÜTTER FH Oberösterreich
(Standort Linz)
THOMAS SCHÖFTNER Private Pädagogische Hochschule der Diözese
Linz

RELEVANCE & RESEARCH QUESTION In recent years, numerous computational tools, often based on artificial intelligence, have emerged. While several survey studies attempt to capture the facets of AI interaction, the prevalence of AI in society and the competencies needed to use it effectively for problem-solving remain largely unknown due to their complex nature. Data donations, which involve tracking AI interactions and augmenting them with other forms of user data, may offer a solution to this problem. Thus, we conducted a trial study in which mobile computing devices were provided to a set of students, enabling the collection of log files on digital trace data as they interacted with AI to solve problems. The study's objectives were to understand how students interact with AI and, at a methodological level, to what extent data donation through mobile computer devices enables a valid capture of AI usage behavior.

METHODS & DATA Using cluster sampling, we recruited 21 undergraduate and postgraduate students from three types of universities (general, applied, and educational) in Austria. The students were provided with mobile computer devices preloaded with various AI tools (e.g., ChatGPT) to use over approximately four weeks for course-specific tasks, such as creating a scientific poster, writing a report, or preparing for an exam. Usage data was collected via log files, permitting a comprehensive analysis of students' interactions with the tools.

RESULTS Preliminary findings indicate that data donation via log files from mobile computer devices offers a viable method for examining AI tool use. The collected data provided insights into students' approaches to prompt engineering, usage scenarios, and usage frequency. Study limitations included the small sample size due to resource constraints and the specific, education-related contexts of AI use.

ADDED VALUE Data donation via log files for collecting comprehensive, contextual data on digital behavior with AI tools offers a scalable design, particularly suitable for educational institutions or organizations seeking to understand digital behavior patterns. This method allows qualitative insights into AI interactions and quantitative analysis of usage frequency. Ethical considerations, such as data privacy and informed consent, were rigorously addressed throughout the study.

MEASURING THE ACCURACY OF SELF- REPORTED INSTAGRAM BEHAVIOR – A DATA DONATION APPROACH

FRIEDER RODEWALD, FLORIAN KEUSCH, DARIA SZAFRAN, RUBEN BACH University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION Current research on online behavior heavily relies on self-reported data, which, if flawed, can lead to inaccurate inference in subsequent analyses. Researchers examining online behavior require detailed measures beyond "time spent on a platform" to explore, for example, well-being, social media use, or online privacy, particularly to differentiate between active and passive social media use.

This study investigates the extent of misreporting in questions about fine-grained Instagram behavior by comparing them to objective measures collected via data donation. We also explore to what extent the accuracy of self-reports is dependent on the response format (rating scale vs. open text field) and the reference period ("last week" vs. "typical week").

METHODS & DATA We collected survey data from over 500 Instagram users in a German probability-based online panel regarding 25 distinct behaviors, including posting, liking, and commenting. Participants first complete survey questions on these behaviors. As part of the survey, we conduct a 2x2 experiment that randomly varies the reference period and, for a subset of behaviors, the response format. Respondents are then asked and, if they agree, instructed to download their Instagram usage data for the last three months and donate them to our research. We analyze correlation coefficients between behavioral self-reports and donated data to assess the accuracy of self-reports in general and for specific behaviors.

RESULTS Our study's data collection phase ended on 12 November, and we cannot present any results yet. We have successfully collected self-reported and donated behavior data from 122 respondents. We will update this abstract with the respective findings before 1 March 2025.

ADDED VALUE This study contributes in three ways. First, we inform the field of questionnaire design by offering insights into how to accurately inquire about specific online behaviors, which is particularly interesting for researchers who may not utilize data donation methods. Second, we examine the accuracy of self-reported data on individual Instagram behavior, helping researchers assess the validity of surveying self-reported online behaviors. Third, we illustrate the potential of data donation to gather detailed, fine-grained data on individual behaviors, which participants might be unable to report accurately.

TUESDAY, 01.04.2025
12:00 - 01:15 PM
HÖRSAAL D

3.4: DIGITAL BATTLE- GROUNDS

SESSION CHAIR: LONG NGUYEN
DEZIM – GERMAN CENTER FOR INTEGRATION AND
MIGRATION RESEARCH, GERMANY

PRESENTATIONS

WAR JOURNALISM ON TELEGRAM?

ETHICAL AND PROFESSIONAL

DILEMMAS OF TELEGRAM CONTENT

CREATORS DURING WAR

VERED ELISHAR, DANA WEIMANN SAKS, ARIEL YARON The Max Stern
Yezreel Valley College, Israel

RELEVANCE & RESEARCH QUESTION The Iron Swords War, which erupted in October 2023, has witnessed the emergence of Telegram as a significant news platform in Israel. The war was marked by the proliferation of alternative news channels and their exponential growth in followers. This study analyzes items from the nine most-followed independent Telegram news channels during the first year of the war. It examines how content creators address ethical and professional dilemmas traditionally associated with conventional journalism, as reflected in their posts. The main research questions are: What types of professional and ethical dilemmas emerge in these Telegram channels, and how do these align with or diverge from conventional journalistic dilemmas?

METHODS & DATA The study employs qualitative content analysis of posts from selected Telegram news channels, sampled between October 2023 and September 2024, beginning from the October 2023. A systematic random sampling yielded 300 posts (distributed equally across channels), which were thematically analyzed to identify explicit and implicit references to ethical and professional dilemmas. These were subsequently compared with the Israeli Journalists Association's ethical code of conduct to identify patterns of convergence and divergence.

RESULTS Content creators across all analyzed channels consistently grappled with professional and ethical dilemmas throughout the study period. The primary tension emerged between respecting victims' families' sensitivities and fulfilling the professional obligation to inform the public. A secondary dilemma involved balancing journalistic transparency against national security considerations. No-

tably, these channels demonstrated limited adherence to traditional journalistic practices regarding error correction and fact-verification protocols.

ADDED VALUE This research contributes to the emerging scholarship on Telegram as a news platform and presents the first systematic analysis of its role during active warfare. By examining content creators' navigation of professional and ethical challenges, it illuminates the evolving nature of wartime journalism in digital spaces. The findings enhance our understanding of both crisis reporting and digital alternative journalism while raising important questions about accountability in emerging news platforms.

EARLY WARNINGS:

FORECASTING EDITS AND DISPUTES ON WIKIPEDIA ARMED CONFLICT PAGES

MARIETH COETZER ¹, FABIAN BRAESEMANN ^{1,2} ¹ Oxford Internet
Institute, Oxford, United Kingdom; ² Einstein Center Digital Future,
Berlin, Germany

RELEVANCE & RESEARCH QUESTION In today's digital world, news websites, online encyclopedias and other knowledge platforms are crucial for many to gain access to reliable sources of information. On Wikipedia – one of the most frequented web pages globally – every edit to an article regarding an event must cite a reliable source. Therefore, we hypothesise that edits on Wikipedia pages can be predicted using related news articles. Understanding the process of updating Wikipedia pages via news articles gives insight into how information, including misinformation, integrates into a platform used by the public as a reliable source of information. We consider pages on armed conflicts as a case study, as these are often fast-developing events with much information need.

METHODS & DATA We examine the Russo-Ukrainian War, the Mali War, and the Sudanese Civil War over a 4-year time period. We use an API to collect online news articles as well as Wikipedia articles and edits. The Wikipedia pages and news articles are linked to each other using keywords which appear in both sources. We then use three machine learning models to predict the number of Wikipedia edits per day using the news sources. To examine how news articles relate to edits on a more granular level, we introduce a novel metric called the difference measure.

RESULTS We find that all the machine learning models perform well, producing an MSE score of less than one that is stable across different wars and pages. Of notable interest is the fact that the titles of BBC articles act as sufficient predictors. Additionally, the difference measure is able to supplement traditional controversy measures by identifying excessive edits in response to real-life events.

ADDED VALUE This framework is the first that forecasts when information on Wikipedia is likely to change on a granular scale using exogenous sources. Hence, it could be used as an early warning tool of a page's vulnerability on any online platform that is regularly updated by an active user community. This gives the moderators of the community more time to implement protection methods to preserve the quality of information.

IS WIKIPEDIA A BATTLEGROUND OF THE RUSSO-UKRAINIAN WAR?

LEOPOLD SVEN AUGUSTIN, FABIAN BRAESEMANN University of Oxford, United Kingdom

RELEVANCE & RESEARCH QUESTION

This article investigates the link between territorial conflict and digital disputes, focusing on Wikipedia during the Russo-Ukrainian war. Wikipedia, one of the most visited websites, is regarded as a “consensus truth” in Western societies and has increasing significance due to large language models that rely heavily on its content, making it a key data source for developing AI tools. Despite its role as a reliable information source, the platform’s open-access nature makes it vulnerable to manipulation and attempts to influence public opinion. Therefore, we hypothesize that the 2022 invasion of Ukraine led to heightened attention and disputes on Wikipedia, particularly in articles about contested Ukrainian regions.

METHODS & DATA To analyse the digital impact of territorial disputes, we use Wikipedia data and classify regions into three groups: disputed Ukrainian Oblasts, undisputed Ukrainian Oblasts, and Polish Voivodeships. Disputed Oblasts are defined as areas where Russian forces gained territory, or Russia annexed a region without Ukraine accepting the claim.

We develop a custom Natural Language Processing (NLP) method to identify dispute edits on the Wikipedia pages of disputed Ukrainian Oblasts.

In addition to using the number of dispute edits as a conflict proxy, we quantify the 2022 invasion’s impact with two further metrics: the number of revisions (proxying editor community attention) and the number of identity reverts (indicating generalized disputes). We use a difference-in-difference (DiD) regression setup to measure the invasion’s effect on all three dependent variables.

RESULTS We find a significant increase in attention and dispute on articles about disputed Ukrainian regions compared to those about undisputed regions. Furthermore, we find that dispute edits frequently involve debates over the Ukrainian versus Russian spelling of place names and discussions of national identity, reflecting the broader conflict.

ADDED VALUE The results presented here confirm our initial assumption that territorial conflicts are spreading into the digital realm in the case of the Russo-Ukrainian war. While research exists on other media, such as social media, newspapers, and mass media, our findings shed light on a previously overlooked digital conflict zone: attempts to change the identity of disputed regions on Wikipedia.

TUESDAY, 01.04.2025
03:45 - 04:45 PM
HÖRSAAL A

5.1: INNOVATIONS IN SAMPLING

SESSION CHAIR: SEBASTIAN LUNDMARK
UNIVERSITY OF GOTHENBURG, SWEDEN

PRESENTATIONS

HOROSCOPING AND SAMPLING: PREREGISTERED EXPLORATION OF THE IMPACT OF BIRTH MONTH ON RESEARCH OUTCOMES VIA THE ‘WHOSE BIRTHDAY IS NEXT’ SAMPLING STRATEGY

LYDIA REPKE GESIS - Leibniz Institute for the Social Sciences, Germany
JORIS MULDER Tilburg University, Netherlands
DANIEL OBERSKI 3 Utrecht University, Netherlands

RELEVANCE & RESEARCH QUESTION A large corpus of studies across various domains has demonstrated a birth month effect, wherein individuals born in specific months display distinct outcomes compared to those born in other months with respect to areas such as health, socioeconomic status, and behavior. In contrast, the common use of birthday-based sampling methods (e.g., selecting respondents in a household whose birthday is next or was last) in large-scale surveys assumes that birth month is uncorrelated with outcome variables. If a birth month effect exists, this assumption may introduce bias, particularly when comparing groups with systematic differences in household size, such as non-Western immigrants and majority populations in Western Europe.

METHODS & DATA We first develop a theoretical framework to explore the relationship between birth month effects and potential biases in birthday sampling designs. We conduct a preregistered empirical analysis using the LISS panel (Longitudinal Internet studies for the Social Sciences), a probability-based online panel of Dutch households. In the LISS panel, data is collected from all individuals within a household aged 16 and above. Through simulations across 12 different fieldwork periods (i.e., months of the year), we assess the extent of bias that might arise if the LISS panel employed the next-birthday sampling method instead.

We examine 35 variables, including personality traits, health outcomes, and socioeconomic status, to evaluate the potential impact on research outcomes.

RESULTS Our analysis does not reveal evidence for a strong birth month effect for the selected variables. The simulations show that, for the Dutch context, the next-birthday sampling method does not introduce substantial bias for the variables of interest.

ADDED VALUE Though a null finding, our study provides important insights for survey methodology. It suggests that, in the Netherlands, next-birthday sampling is unlikely to produce bias related to birth month effects, at least for the way how these variables are commonly measured in social science surveys. This contributes to the ongoing discussion on sampling methods and enhances the reliability of results in large-scale surveys.

SAMPLING REFUGEES IN COUNTRIES OF FIRST REFUGE – AN INTERNATIONAL SNOWBALL SAMPLING APPROACH WITH MULTIPLE TARGET POPULATIONS

MARVIN BÜRMAN, ARMIN KÜCHLER Bielefeld University, Germany

RELEVANCE & RESEARCH QUESTION As reported by the UNHCR for 2023, the majority of refugees are hosted in low- and middle-income countries (75%) and countries that neighbor their country of origin (69%). For refugees who are particularly vulnerable and unable to return to their country of origin, resettlement programmes aim to provide long-term prospects by resettling them to Western countries. And despite the fact that only a small proportion of those in need for resettlement are actually resettled (8% according to UNHCR), little is known about the living situation of those who are left behind. This study addresses this gap by conducting a web-survey targeting potential resettlement refugees using social contacts of already resettled refugees.

METHODS & DATA The study uses addresses of all refugees resettled to Germany since 2013 (i.e., approx. 17,000) from the German Central Register of Foreigners (AZR) to invite participants via postal mail to a web-survey. This survey marks the start of a snowballing approach, where refugees in Germany function as seeds and are asked to forward a survey-link (with a mobile “share” option) to up to three contacts, who still reside in countries of first refuge and may be eligible for resettlement (target population A). To reduce the risk of realizing too few cases – especially in the first step from Germany to abroad –, participants are also asked to share the survey to refugees who may not be eligible for resettlement (target population B). In the subsequent steps, this snowballing process continues in the countries abroad.

RESULTS Although data collection is ongoing at the time of the GOR conference, findings from cognitive pretests and from data already collected will be presented, offering insights into the potential but also challenges of surveying refugees via international snowball sampling.

ADDED VALUE This study contributes to the understanding of the living conditions of refugees while offering methodological insights into sampling hard-to-reach populations. By demonstrating how snowball sampling with multiple target populations can mitigate recruitment challenges, it provides valuable lessons for researchers focusing on vulnerable groups.

SOCIAL MEDIA SAMPLING FOR QUANTITATIVE SURVEYS IN HARD-TO-REACH COUNTRIES

ORKAN DOLAY, CLEMENS RATHE Bilendi & respondi, France

RELEVANCE & RESEARCH QUESTION Traditional online survey panels often lack coverage in smaller or less digitally integrated countries, limiting researchers’ ability to collect reliable data from these regions. Social media sampling presents a promising alternative for quantitative surveys in such contexts. This study investigates the feasibility, reliability, and potential biases of social media sampling as a data collection method. Using a multi-country survey spanning 18 nations—including Zimbabwe, Kazakhstan, Costa Rica, and Iceland—we address the question: Can social media sampling provide reliable insights for brand perception and societal measures, in countries where online panels are unavailable?

METHODS & DATA The study deployed a quantitative survey during June and July 2024 through targeted social media advertisements, optimized to recruit representative samples across diverse demographics. In total n9.000 Participants responded to a standardized questionnaire including brand image measures and selected questions from the World Happiness Report. Sampling quotas and algorithmic targeting ensured coverage of gender, age, regions, ethnicities and income levels each country. Reliability was assessed through comparative analysis against external data sources, where available, and consistency checks within the datasets.

We compared metrics like available audience in Meta, click-rates, completion rates, drop-outs per questions, differences in response behaviour between countries and between ethnic groups within countries.

RESULTS The results demonstrate that social media sampling can effectively generate diverse, balanced samples in countries lacking established online panels. Across the 18 countries, response rates and sample representativeness varied but were sufficient for robust analysis. Insights from brand image and happiness metrics revealed consistent trends across nations and offered valuable local context, while some limitations, such as underrepresentation of older rural populations, and low-income groups and some ethnicities in African countries were noted.

ADDED VALUE This research highlights the untapped potential of social media as a viable sampling solution for most of the investigated countries. By showcasing a rigorous approach to survey design, execution, and evaluation, this paper contributes a practical framework for using social media to extend the reach of quantitative research globally. The findings are particularly relevant for researchers, seeking solutions for data collection in emerging and underrepresented markets.

TUESDAY, 01.04.2025
03:45 - 04:45 PM
HÖRSAAL B

5.2: RESPONDENT NUDGING AND INCENTIVES

SESSION CHAIR: GEORG-CHRISTOPH HAAS
INSTITUTE FOR EMPLOYMENT RESEARCH, GERMANY

PRESENTATIONS

KNOCK-TO-NUDGE METHODS TO IMPROVE SURVEY PARTICIPATION IN THE UK

OLGA MASLOVSKAYA, CRISTIAN DOMARCHI, PETER WF SMITH
University of Southampton, United Kingdom

RELEVANCE & RESEARCH QUESTION The knock-to-nudge is an innovative method of household contact, first introduced during the COVID-19 pandemic when face-to-face interviewing was not possible. In this approach, interviewers visit households and encourage sampled units to participate in a survey through a remote survey mode (either web or telephone) at a later date. Interviewers may also collect contact information, such as telephone numbers or email addresses, or conduct within-household selection of individuals on the doorstep. This approach continued to be used post-pandemic, but there remains a knowledge gap regarding its advantages and disadvantages. It is still unclear whether knock-to-nudge approach leads to improvements in sample composition and data quality.

METHODS & DATA We analyse data from three UK surveys: the National Survey for Wales (NSW), the Transformed Labour Force Survey (TLFS), and the National Readership Survey (PAMCo), each of which employed different versions of the knock-to-nudge approach. Our goal is to assess whether this method improves sample composition across these three surveys. We begin with descriptive analysis and then apply logistic regression models to investigate the composition of subsamples that received different recruitment treatments. Finally, we compare the sample composition at various stages of the recruitment process.

RESULTS Our findings suggest that the knock-to-nudge approach is effective in improving sample composition, while also reducing costs compared to traditional face-to-face interviews.

ADDED VALUE This study contributes to the under-researched area of knock-to-nudge methods. The results indicate that, when carefully designed and implemented, this approach can enhance recruitment efforts and improve sample composition of the resulting samples in surveys.

RECRUITMENT INCENTIVE EXPERIMENT OF THE PROBABILITY-BASED PANEL HEALTH IN GERMANY: RESULTS ON OUT- COME RATES, NON-RESPONSE BIAS AND PANEL CASE COSTS

JOHANNES LEMCKE, STEFAN DAMEROW, ILTER ÖZTÜRK Robert Koch-Institut, Germany
NICOLAS FRENZEL BAUDISCH, THOMAS WEISS, JENNIFER ALLEN
infas Institut für angewandte Sozialwissenschaft GmbH

RELEVANCE & RESEARCH QUESTION The Robert Koch Institute (RKI) set up a probability-based panel infrastructure focused on public health research ('Health in Germany') (registered active panelists about 47.000). Due to declining response rates in recent decades, incentives have become increasingly important. Incentive experiments are therefore often carried out in order to achieve high data quality with a lower use of resources. Therefore, for the first recruitment study of the panel Health in Germany the RKI conducted an incentive experiment with a random sub-sample in order to test the effectiveness of different incentive schemes. The central questions of the incentive experiment are: Differentiated by incentive group (a) the response rate for panel registration, (b) the possible distortion due to non-response bias and (c) the panel case costs.

METHODS & DATA The study population comprises all persons aged 16 and over living in Germany. Around 170,000 addresses of the residents' registration offices were used as a random sample. The field period ran from January to May 2024. Some of the target persons were randomly selected for an incentive experiment with four groups of 1440 individuals each. The incentives were either paid unconditionally beforehand ('before') or were linked to registration for the panel ('after'). The incentive schemes of the groups were: (1) €5 before, €10 after, in cash; (2) €10 after, in cash; (3) €5 before, €10 after, as a voucher; (4) no incentive at all (control group).

RESULTS The incentive experiment replicates existing incentive experiment studies. The use of cash instead of vouchers and the unconditional payment ('in advance') of €5 substantially increases the willingness to participate (about 12 percentage points difference). Particularly in the hard-to-reach population group of people with a low level of education, a less biased sample composition can be observed in comparison to all other incentive groups.

ADDED VALUE The results of the incentive experiment provide insights into how high data quality can be achieved with fewer resources. In view of the large number of cases and probabilistic sampling, the findings can be transferred to similar epidemiological research projects.

TUESDAY, 01.04.2025
03:45 - 04:45 PM
HÖRSAAL C

5.3: LLMS AND SYNTHETIC SURVEY DATA

SESSION CHAIR: JOHANNA HÖLZL
UNIVERSITY OF MANNHEIM, GERMANY

PRESENTATIONS

SYNTHETIC RESPONDENTS, REAL BIAS? INVESTIGATING AI-GENERATED SURVEY RESPONSES

CHARLOTTE PAULINE MÜLLER Lund University, Sweden
BELLA STRUMINSKAYA, PETER LUGTIG Utrecht University, The Netherlands

RELEVANCE & RESEARCH QUESTION The idea to simulate survey respondents has lately been seen as a promising data collection tool in academia and market research. However, previous research has shown that LLMs are likely to reproduce human biases and stereotypes existent in their training data. Because of this, we further investigate the potential benefits and challenges of creating synthetic response datasets by following two major aims: 1. investigate whether AI tools can replace real survey respondents, and if yes, for which questions and topics, and 2. explore whether intentional prompts reveal underlying biases in AI prediction.

METHODS & DATA We compare already existing survey data from the German General Social Survey (Allbus) 2021, to AI-generated synthetic data with the OpenAI model GPT-4. For this, we took a random sample of 100 respondents from the Allbus dataset and created a so-called AI-Agent for each. Each Agent was calibrated based on general instructions and individual background information (14 variables). We chose to predict three different types of outcomes, including a numerical, binary, and open text/string format, each of them inheriting the potential to provoke certain biases, such as social desirability, gender, and age stereotypes. Furthermore, each item was tested across different contextual factors, such as AI model calibration and language settings.

RESULTS We found a deep lack of accuracy in the simulation of survey data for both numerical ($r = -0.07$, $p = 0.6$) as well as binary outcomes ($\chi^2(1) = 0.61$, $p = 0.43$, $V = 0.1$), while the explanatory power of the background variables for the predicted outcome, was high for both the

former ($r = 0.4$) and the latter ($r = 0.25$). Furthermore, we found no difference in the prediction accuracy between different input languages and AI model calibrations. While predicting open-text answers, individual background information was generally well considered by the AI tool. However, several potential biases became apparent, such as age, gender, and regional biases.

ADDED VALUE Our research contributes to a more ethically responsible application of AI tools in data simulation, highlighting an urgent need for more caution in the already-started utilization of AI-generated datasets.

TALK, TALK, TALK – UNLOCKING AI FOR CONVERSATIONAL RESEARCH

BARBARA VON CORVIN Human8 Europe, Belgium

RELEVANCE & RESEARCH QUESTION Conversational AI tools offer great advantages over human moderation when it comes to logistics: independence of time and moderator availability as well as multilingualism. One of the tools we tested even offers reporting.

At Human8, we were curious to see how good conversational AI tools are already, how research participants and consultants perceive them and where it makes sense to use these tools. We started experimenting with two conversational AI tools, both offering AI-moderated interviewing capabilities. We tested the conversational AI tools in different use cases to see how research participants react and where the tool can maybe even provide a benefit over human moderation apart from the logistic advantages.

METHODS & DATA Over the past months, Human8 has tried different AI moderation plug-ins. Currently, we're running an internal research project with $n=250$ in which we examine three use cases:

1. Consumer stories: consumers tell us about their lives in a more intimate channel
2. Agile qual research: AI replacing the moderator in a one-on-one interview situation
3. Augmented quant research: diving deep on a specific topic as part of a quantitative online survey

RESULTS Based on our research so far, but before final results will be available on February 21st (and will then be shared with the committee), we can already say that:

- People like the conversation with the AI bot.
- The reactions of the AI bot make the interviewee feel understood and motivate them to go further or even more into detail.
- The amount of info shared by the interviewee is higher than when typing text.
- The input we receive is more emotional and less over-rationalized than written text production, there is less filter.
- There is less interviewer bias.

ADDED VALUE Our results shed a light on conversational AI tools/ AI moderation plugins, examine different use cases and uncover what to consider when using these tools. We will share our learnings on where and in which situations AI moderation plug-ins can provide a relevant benefit with the audience and open the discussion.

TUESDAY, 01.04.2025
03:45 - 04:45 PM
HÖRSAAL D

5.4: SOCIAL MEDIA INFLUENCE IN WAR TIMES

SESSION CHAIR: ZAZA ZINDEL
BIELEFELD UNIVERSITY, GERMANY

PRESENTATIONS

RUMOR PROPAGATION ACROSS WAR STAGES: INFLUENCES OF PSYCHOLOGICAL, SOCIAL, AND MAIN- STREAM MEDIA FACTORS

YARON ARIEL, VERED ELISHAR, DANA WEIMANN SAKS The Max Stern
Yezreel Valley College, Israel

RELEVANCE & RESEARCH QUESTION This study investigates the psychological and media-related factors influencing rumor propagation during the Israel-Hamas 2023-2024 war, comparing early (first month) and later stages (ninth month) of the conflict. Using the unprecedented case of prolonged conflict in Israel, we examine how situational anxiety, trait anxiety, and psychological closeness interact with media consumption patterns to affect rumor dissemination. The research explores how these relationships evolve over time, contributing to our understanding of information behavior during extended crisis situations.

METHODS & DATA Data were collected through a longitudinal survey of 347 Jewish-Israeli participants recruited via the Midgam Panel, using stratified sampling aligned with Israeli Central Bureau of Statistics demographics. The initial sample of 500 participants achieved a 70% retention rate at the nine-month follow-up. Measurements included validated scales for state and trait anxiety (adapted from STAI), psychological closeness, media consumption across platforms, and rumor-spreading behavior. Analyses employed t-tests, correlations, and Hayes' PROCESS macro for mediation analysis.

RESULTS While rumor-spreading levels remained stable across time periods, significant decreases were observed in situational anxiety, psychological closeness, and both social and mainstream media consumption between early and late stages. Social media platforms consistently showed stronger associations with rumor spread than mainstream media, with correlation strengths increasing over time. In

the early stage, psychological closeness and trait anxiety significantly predicted rumor spread, mediated by media consumption. However, in the later stage, these factors directly influenced rumor spreading without significant mediation effects, indicating an evolution in the psychological mechanisms driving information sharing.

ADDED VALUE This study advances crisis communication theory by demonstrating how psychological and media factors evolve differently in their influence on rumor propagation during prolonged conflicts. The findings reveal that while overall rumor-spreading behavior remains stable, the underlying psychological and media-related mechanisms shift significantly. These insights contribute to developing more effective, stage-specific strategies for managing information flow during extended crises, particularly highlighting the distinct roles of social and mainstream media in different conflict phases.

SOCIAL MEDIA, ANXIETY, AND ETHNIC DISPARITIES – THE CASE OF THE JEWISH AND ARAB POPULATION GROUPS FOLLOWING OCTOBER 7 IN ISRAEL

VLAD VASILIU Max Stern Yezreel Valley College, Israel
HANEEN SHIBLI University of Washington, USA
GAL YAVETZ Bar Ilan University, Israel

RELEVANCE & RESEARCH QUESTION This study addresses the psychological impacts of crises and the role of social media in shaping those impacts, focusing on ethnic disparities between Jewish and Arab populations in Israel. The research investigates how social media consumption influences anxiety levels following the October 7 terror attack. Core research questions explore the relationship between social media use and Generalized Anxiety Disorder (GAD), comparing consumption patterns and psychological responses across these groups.

METHODS & DATA A cross-sectional study was conducted with 651 participants (505 Jewish, 146 Arab) using stratified sampling to ensure representativeness. Data was collected through online anonymous questionnaires available in Hebrew and Arabic during the fifth and sixth weeks post-attack. The survey assessed media consumption, perceived credibility of information sources, social media activity, and anxiety levels via the GAD-7 scale. Statistical analyses, including chi-square tests, Spearman correlations, Independent Samples Tests, and linear regression, were performed to analyze the data.

RESULTS Significant differences were observed in media consumption patterns and anxiety levels. Jewish participants reported higher social media usage and anxiety scores, while Arabs showed reduced usage and lower anxiety. Jewish participants trusted official sources like the IDF spokesman more, whereas Arabs relied on social media and journalists. A positive correlation between social media use and anxiety was evident among Jews, contrasting with Arabs, where higher GAD scores correlated with decreased social media usage. Linear regression revealed a link between anxiety and the adoption of new social media platforms during the crisis.

ADDED VALUE This research offers critical insights into how social media behavior during conflicts impacts mental health across different

ethnic groups. It emphasizes the need for tailored crisis communication strategies that account for diverse sociocultural contexts. The study highlights the dual role of social media as both a coping mechanism and a stress amplifier, providing a foundation for future policy development and intervention programs to mitigate psychological harm during crises.

WHAT MAKES MEDIA CONTENTS CREDIBLE? A SURVEY EXPERIMENT ON THE RELATIVE IMPORTANCE OF VISUAL LAYOUT, OBJECTIVE QUALITY AND CONFIRMATION BIAS FOR PUBLIC OPINION FORMATION

SANDRA WALZENBACH University of Konstanz, Germany

RELEVANCE & RESEARCH QUESTION The emergence of social media has transformed the way people consume and share information. As such platforms widely lack mechanisms to ensure content quality, their increasing popularity has raised concerns about the spread of fake news and conspiracy beliefs – with potentially harmful effects on public opinion and social cohesion.

Our research aims to understand the underlying mechanisms of media perception and sharing behaviour when people are confronted with factual vs conspiracy-based media contents. Under which circumstances do people believe in a media content? Do traditional indicators of quality matter? Are pre-existing views more important than quality (confirmation bias)? How is perceived credibility linked to sharing behaviour?

METHODS & DATA To empirically assess these questions, we administered a survey experiment to a general population sample in Germany via Bilendi in August 2023. As respondents with a general susceptibility to conspiracy beliefs are of major substantive interest, we made use of responses from a previous survey to oversample “conspiracy thinkers”.

Respondents were asked to evaluate the credibility of different media contents related to three vividly debated topics: vaccines against Covid-19, the climate crisis and the Ukraine war. We analyze these evaluations regarding the quality of the content (measured by author identity and data source), its visual layout (newspaper vs tweet), and previous respondent beliefs on the respective topic to measure confirmation bias.

RESULTS Our findings suggest that the inclination to confirm pre-existing beliefs is the most important predictor for believing a media content, irrespective of its quality. This general tendency applies to both, the mainstream society and “conspiracy thinkers”. However, according to self-reports the latter group is much more likely to share media contents they believe in.

ADDED VALUE Methodologically, we use an interesting survey experiment that allows us to vary opinion (in)consistency and objective quality of media contents simultaneously, meaning that we can estimate the relative effect of these features on the credibility of media

contents. We provide insights into the underlying mechanisms of the often debated spread of conspiracy beliefs through online platforms, with their practical implications for public opinion formation.

TUESDAY, 01.04.2025
05:00 - 06:00PM
HÖRSAAL A

6.1: INCREASING SURVEY RESPONSE

SESSION CHAIR: BARBARA FELDERER
GESIS, GERMANY

PRESENTATIONS

A SIMPLE INVITATION: A STUDY ON THE IMPACT OF SIMPLIFIED INVITATION LETTERS ON THE WILLINGNESS TO JOIN A PROBABILITY-BASED WEB PANEL IN SWEDEN

ALEXANDRA GARCIA NILSSON, FELIX CASSEL, BJÖRN RÖNNER-STRAND The SOM Institute, Sweden

RELEVANCE & RESEARCH QUESTION Reducing nonresponse bias in the recruitment to probability-based online access panels is essential for any panel striving to achieve accurate inferential statistics. However, certain groups of the population tend to be more difficult to recruit. For example, research have identified that foreign-born individuals and individuals with less educational attainment were more difficult to recruit.

Targeted efforts to enhance recruitment and reduce breakoffs in these groups might be especially efficient for reducing potential nonresponse bias. The present study assessed whether simplified invitation letters increased recruitment and reduced breakoffs in a probability-based web panel and whether a simplified language improved those rates even more so among hard-to-recruit subgroups of the population.

METHODS & DATA In a probability-based recruitment to the Swedish Citizen Panel (SCP) conducted in fall 2024, the sample was randomly assigned into one of two groups; One group (N = 9,000) were assigned to receive an invitation and reminder letter with the standard

language typically used in SCP recruitment. The experimental group (N = 9,000) received a revised version of the letter, written in simpler language that avoided academic jargon and words with many syllables.

RESULTS Data will be collected in November-December 2024 and will be analyzed and reported in an updated abstract in early January 2025. The effects on recruitment and breakoffs will be analyzed for the full sample, as well as for subgroups based on register information on sex, age, education and immigrant background.

ADDED VALUE The present experiment will demonstrate whether simplified invitation letters increase recruitment rates and decrease breakoffs in a probability-based online access panel, with a particular focus on the impact among hard-to-recruit populations. Simplified invitation letters may offer a cost-effective method to boost recruitment- and breakoff rates as well as reduce demographic skewness, and, in turn, nonresponse bias.

PICTURE THIS! THE INFLUENCE OF STRESSING THE CAMERA FEATURE IN THE MAIL INVITATION TO AN APP-BASED HOUSEHOLD BUDGET SURVEY ON PARTICIPATION BEHAVIOR

FLORIAN KEUSCH, MAREN FRITZ University of Mannheim, Germany
JOHANNES VOLK, LASSE HÄUFGLÖCKNER Destatis – Federal Statistical Office Germany, Germany

RELEVANCE & RESEARCH QUESTION Diary-based household budget surveys are often burdensome, as respondents must manually log details (e.g., product type, quantity, price) for each purchase over time. A smartphone app that allows participants to upload photos of shopping receipts could reduce this burden and increase survey participation. However, requiring an app that accesses the camera may raise privacy concerns, potentially affecting willingness to participate. This study investigates the influence of highlighting the app's camera feature in the invitation on both response rates and camera usage in the diary.

METHODS & DATA In November 2024, a sample of over 7,000 individuals in Germany whose addresses were drawn from residents' registration office lists were invited to participate in a smartphone-based household budget survey. Invitation letters included a QR code, URL, and personalized login credentials for the AusgabenAtlas app, compatible with Android and iOS devices. Participants were asked to log their daily spendings in a two-week diary, with the option to use the camera function for receipt uploads or input data manually.

Sample members were randomly assigned to one of three invitation treatments that varied in how the camera feature was framed: (1) no mention of the camera feature, (2) mention of the camera feature, and (3) mention of the camera feature with additional emphasis on the effort required.

RESULTS Data collection will conclude in December 2024, with preliminary findings available in January 2025. We will examine the effect of the invitation treatment on survey response rates and camera feature usage within the app.

ADDED VALUE This research is part of the Smart Survey Implementation (SSI) Project, funded by EUROSTAT, which aims to enhance data collection for official statistics across Europe through digital innovation. This experiment specifically addresses recruitment challenges in app-based surveys and evaluates the potential of mobile technology to streamline participation in official household budget surveys.

ASK ME NOW OR LOSE ME LATER – THE IMPACT OF IMMEDIATE FOLLOW-UP ON PARTICIPATION RATES, RETENTION AND DATA QUALITY IN WEB PANELS

TILDA EKSTRÖM, ALEXANDRA GARCIA NILSSON, ANDERS CARLANDER
 University of Gothenburg, Sweden

GOR 2025 Conference

Ask Me Now or Lose Me Later – The Impact of Immediate Follow-Up on Participation Rates, Retention and Data Quality in Web Panels

RELEVANCE & RESEARCH QUESTION Ensuring high participation rates and respondent retention is essential for sample and data quality in web panels. This study examined whether the time interval between recruitment and invitation to a respondent's first panel wave affected their likelihood to participate, of unsubscribing from the panel, and the quality of their survey responses. Drawing on Construal Level Theory (CLT), we posit that a longer interval may increase psychological distance and therefore decrease engagement in the web panel. By varying the time between recruitment and the first survey invitation, we aim to present optimal strategies to enhance participation and reduce attrition in web panels.

METHODS & DATA This study employed an experimental design using a newly recruited (spring 2024) non-probability sample of panelists (N = 3,140) from the Swedish Citizen Panel. Half of the sample (n = 1,570) were randomly assigned to receive their first panel wave invitation shortly after being recruited, whereas the other half (n = 1,570) were assigned to receive their first invitation six months after being recruited. Both groups responded to the same survey in December 2024.

RESULTS Data will be collected in December 2024 and will be analyzed and reported in an updated abstract in early January 2025. Added Value: Understanding the timing of follow-up surveys after recruitment may be essential for deciding whether panelists benefit from an immediate invitation to complete a panel wave shortly after being recruited or if delaying the invitation may be detrimental for engagement. Our findings provide insights for the management of web panels and strategies on newly recruited panelists to reduce dropout rates and improve data quality.

TUESDAY, 01.04.2025
05:00 - 06:00 PM
HÖRSAAL B

6.2: **MIXED MODE** **AND MODE** **TRANSITIONS**

SESSION CHAIR: BJÖRN ROHR
 GESIS - LEIBNITZ INSTITUTE FOR THE SOCIAL SCIENCES,
 GERMANY

PRESENTATIONS

LARGE-SCALE SOCIAL SURVEYS **WITHOUT FIELD INTERVIEWERS IN** **THE UK: AN EVIDENCE REVIEW**

CRISTIAN DOMARCHI, OLGA MASLOVSKAYA, University of Southampton, United Kingdom

PETER J. LYNN, University of Essex, United Kingdom

RORY FITZGERALD, NHLANHLA NDEBELE, RUXANDRA COMANARU
 City, University of London, United Kingdom

RELEVANCE & RESEARCH QUESTION Data collection organisations are shifting toward new approaches, with social surveys undergoing significant design and implementation changes. Since the COVID-19 pandemic, agencies have increasingly moved to online data collection due to dwindling response rates and rising fieldwork costs. A key challenge for self-completion general population surveys is the absence of field interviewers to facilitate recruitment and guiding respondents through the survey process. This research examines the UK survey landscape, aiming to identify recruitment methods for self-administered surveys, that can produce more representative samples of the general population.

METHODS & DATA We present findings from an information request sent to the UK's nine most important survey agencies. We collected information on surveys without field interviewers conducted between 2018 and early 2024, including publicly available technical and methodological reports and other survey materials, along with internal reports provided by the agencies. We processed and codified this information, building a spreadsheet containing 144 instances of 59 longitudinal and cross-sectional surveys, along with 227 communication materials.

RESULTS The responses for the surveys in our dataset use 57% online, 38% paper, and 5% telephone modes. Most surveys (84%) offer incentives to participants, with 92% being monetary and only 33% giv-

en unconditionally. Response rates vary widely – household-based cross-sectional surveys tend to have lower response rates (81% at 30% response or lower) than individual-based ones (47% at 30% or lower). Longitudinal surveys generally have the highest response rates. While only 35% of reports assess sample representativeness, the general trend confirms that mixed-mode surveys yield more representative samples than single-mode surveys.

ADDED VALUE To our knowledge, this review is the first coordinated effort to collate and summarise recruitment strategies for surveys without field interviewers in the UK. It covers sampling design, communication strategies and materials, incentivisation, fieldwork procedures, response rates, and report quality assessments. Our dataset provides insights into the current state of survey practice and helps identifying practices that might contribute towards higher response rates and better sample composition.

DOES WEB AS FIRST MODE IN A **MIXED-MODE ESTABLISHMENT** **SURVEY AFFECT THE DATA QUALITY?**

CORINNA KÖNIG, JOE SAKSHAUG Institute for Employment Research, Germany

RELEVANCE & RESEARCH QUESTION Due to declining response rates and higher survey costs, establishment surveys are (or have been) transitioning from traditional interviewer modes to online and mixed-mode data collection. Previous analyses have shown that mixed-mode designs maintain response rates at lower costs compared to face-to-face designs, but the question remains to what extent introducing the online mode affects measurement quality – this has rarely been addressed in the establishment survey literature.

METHODS & DATA The Establishment Panel of the Institute for Employment Research (IAB) was primarily a face-to-face survey until 2018. Since then, the IAB has experimented with administering a sequential web-first followed by face-to-face mixed-mode design versus the traditional face-to-face design. We address our research question by using this data and comparing the survey responses from the single- and mixed-mode experimental groups to corresponding administrative data from employer-level social security notifications. The accuracy of survey responses in both mode designs is assessed and measurement equivalence is evaluated.

Especially a lot of open-ended variables on numbers of employees with certain characteristics is used and additionally, we report on differences in accuracy between the individual web and face-to-face modes. Furthermore, we consider differences for several alternative data quality indicators, including item nonresponse, social desirability responding, and the use of filter questions. To account for selection and nonresponse bias weights are used throughout the analysis and as sensitivity checks, weights are estimated in different ways. In addition to propensity scores, random forest, and extreme gradient boosting were also applied.

RESULTS Preliminary results show that measurement error bias in online interviews is sometimes larger than in face-to-face interviews but compared to the mixed-mode design the difference is not significant anymore. Looking at sensitive questions it cannot be confirmed that

online respondents answer more socially desirable. Further findings indicate slightly larger item nonresponse in the online mode but considering the sequential mode design no difference can be found.

ADDED VALUE Thus, the study provides comprehensive insights into data quality for mixed-mode data collection in establishment surveys and informs survey practitioners about the implications of switching from single- to mixed-mode designs in large-scale establishment panels.

EXAMINING DIFFERENCES IN FACE-TO-FACE AND SELF-ADMINISTERED MIXED-MODE SURVEYS: INSIGHTS FROM A GENERAL SOCIAL SURVEY

ALEXANDRA ASIMOV GESIS, Germany

RELEVANCE & RESEARCH QUESTION General social surveys are traditionally conducted face-to-face, maintaining long time series for tracking public opinion trends. To ensure comparability, survey designs typically change minimally over time. However, face-to-face surveys have been experiencing declining response rates and higher costs. As a result, self-administered mixed-mode designs have gained popularity due to their ability to circumvent these challenges. Since switching modes is a major methodological change, investigating data comparability with the original mode is critical. Furthermore, self-administered mixed-mode designs can be implemented in two ways: concurrent or sequential. Each result in different proportions of web and mail survey responses. This raises the question: Does this difference in proportions affect comparability with face-to-face surveys?

METHODS & DATA This study uses the German General Survey 2023 (ALLBUS), which surveys the general population aged 18 and older and is traditionally conducted face-to-face. In 2023, ALLBUS included three randomized experimental groups: [1] face-to-face, [2] concurrent self-administered mixed-mode (mail and web), and [3] sequential self-administered mixed-mode (mail and web). This study examines data comparability by evaluating differences in nonresponse bias, sample composition, and measurement between face-to-face and the two mixed-mode designs.

RESULTS Overall, both self-administered mixed-mode designs produce similar results, with both showing slight strengths. The sequential design is slightly more similar to the face-to-face design in terms of nonresponse bias and sample composition. In contrast, the concurrent design achieves slightly smaller measurement differences compared to the face-to-face design.

ADDED VALUE This study offers valuable insights into the shift from face-to-face to self-administered mixed-mode designs by comparing concurrent and sequential approaches. It highlights their strengths in maintaining data comparability, with both designs producing similar results overall. This indicates that web and mail modes are comparable, as the proportion of these two modes varies between designs.

**TUESDAY, 01.04.2025
05:00 - 06:00 PM
HÖRSAAL C**

6.3: SURVEY INNOVATIONS

SESSION CHAIR: ALMUTH LIETZ
GERMAN CENTER FOR INTEGRATION AND MIGRATION RESEARCH, GERMANY

PRESENTATIONS

OPTIMISING ONLINE TIME USE SURVEYS: BALANCING QUALITY, EFFICIENCY, AND INCLUSIVITY

MARTA MEZZANANICA National Centre for Social Research, United Kingdom

RELEVANCE & RESEARCH QUESTION

Key words: time use survey, online diary data collection

Time-use surveys provide invaluable insights into how individuals allocate their time across various activities. Traditionally, paper-based diaries, often administered through face-to-face interviewers, have been the primary method for gathering such data. Advances in technology have enabled the development of online diary tools, but while digital tools offer advantages like reduced costs, cleaner data, and increased flexibility, they also raise concerns about data quality and inclusivity.

In recent years, two time use surveys using online diary tools have been developed and used in the UK, providing a unique opportunity to explore the benefits and challenges of this methodology.

This study assesses the impact of design choices on the quality and efficiency of online time-use surveys. By examining factors such as incentivisation, completion mode, use of support materials, and user interface, we identify strategies for minimising respondent burden while maximising data quality and response rates.

METHODS & DATA

Key words: time-use data, respondent burden, respondent journey, desk review, probability-based panel

This study analyses eight waves of time-use surveys conducted in the UK between 2020 and 2023, using sample from the NatCen's probability-based panel. A desk review of the online diary tools, participant materials, and analysis of available data (response rates, sample profiles, split experiments, paradata, and respondent feedback) was conducted to assess the performance of the fieldwork design and evaluate the respondent journey.

RESULTS

Key words: mobile-first design, incentives, response rates, representativeness

Several key factors influence the success of online time-use surveys. Invitations and reminders must be timely and effective. Online tools should be designed with a mobile-first approach. Printed support materials can guide participants and improve data quality. A telephone fieldwork option can significantly boost response rates and improve sample representativeness. While higher incentives can increase response rates, their effectiveness diminishes with increasing amounts. However, these benefits must be weighed against additional costs.

ADDED VALUE

Key words: practical considerations, time-use online tool, data quality

Practical considerations and recommendations for the design of online time-use studies, ensuring they meet the diverse needs and maximize data quality and sample representativeness.

WHAT DO PARTICIPANTS REFER TO WHEN ASKED ABOUT THEIR PLACE OF RESIDENCE?

MANUELA SCHMIDT University of Kaiserslautern-Landau, Germany
CLAUDIA SCHMIEDEBERG LMU Munich, Germany

RELEVANCE & RESEARCH QUESTION Social research is increasingly recognizing the relevance of spatial context to social phenomena. Thus, many surveys ask respondents about neighborhood characteristics. An open question, however, is the appropriate scale of the spatial context: Is it the immediate neighborhood, the municipality, or the county that matters? We investigate the consistency between respondents' subjective perceptions and objective characteristics across these spatial scales. Additionally, we examine how these perceptions differ between socio-demographic groups.

METHODS & DATA The analysis is based on data from the GLEN study, a nationally representative panel study started in 2024. In the self-administered push-to-web survey, respondents were asked to rate their place of residence on a number of characteristics, including green space, labor market, mobility, and health care accessibility. We test whether the respondents' subjective ratings are consistent with administrative data from official statistics at different spatial scales. Differential effects are tested for dimensions such as age, employment status, and urban vs. rural residence.

RESULTS [The survey is currently in the field. Results will be updated by end of February 2025]

ADDED VALUE We contribute to the understanding of spatial mechanisms by identifying the spatial scale that best aligns with respondents' subjective perceptions, and by highlighting how the area of reference may vary across demographic groups. Our results can give guidance for survey practitioners regarding the design of neighborhood questions as well as the selection of spatial context data as auxiliary variables.

WILL HARMONY LAST?

HARMONIZING TIME SERIES SURVEY

DATA WITH EQUATING UNDER

CHALLENGING PATTERNS OF DATA

AVAILABILITY

MATTHIAS ROTH, RANJIT KONRAD SINGH GESIS – Leibniz-Institut für Sozialwissenschaften, Germany

RELEVANCE & RESEARCH QUESTION Survey researchers often wish to combine survey data from multiple sources to create longer time series. One central problem when combining survey data is that variables of interest are measured with similar but not identical questions. Differences in the question text or the response options make data from questions incomparable, even though they measure the same construct. In this study, we evaluate the harmonization method Observed Score Equating in a Random-Groups Design (OSE-RG) for time series harmonization.

OSE-RG aligns differences between survey questions by transforming the response scale of one question into the format of another question requiring data from the same point in time. The challenge with time series is that we often do not have access to data from both questions at all points in time. To still use OSE-RG for time series harmonization, we need to re-use existing harmonizations at times where only data from one question is available. Thus, our research question is as follows: Is it possible to re-use existing OSE-RG harmonizations over time?

METHODS & DATA To explore the re-usability of OSE-RG harmonizations over time, we harmonize time series of ten pairs of survey questions from three German general population survey programs over a period of 14 years. The central idea is that we create a response scale transformation in one year and then re-use the response scale transformation to harmonize data from other years, tracking harmonization error in the process.

RESULTS We find that OSE-RG harmonizations are re-usable over time for some questions, for example questions measuring general health, but not for others, for example political interest questions. We conclude that OSE-RG is a viable choice for harmonizing time series survey data. However, researchers need to be aware of reduced re-usability of OSE-RG in time series harmonization as a potential source of bias.

ADDED VALUE The added value is threefold: First, we empirically demonstrate how time series survey data can be harmonized using OSE-RG. Second, we pinpoint the re-usability of harmonizations over time as a potential source of bias. Third, we discuss consequences for different patterns of data availability.

TUESDAY, 01.04.2025
05:00 - 06:00 PM
HÖRSAAL D

6.4: LLMS AS ANALYSIS TOOLS

SESSION CHAIR: JELDRIK BAKKER
STATISTICS NETHERLANDS, THE NETHERLANDS

PRESENTATIONS

GOING MULTIMODAL: CHALLENGES (AND OPPORTUNITIES) OF STREAMLINING DELIVERABLE PRODUCTION WITH AI

GEORG WITTENBURG Inspirient
SOPHIA MCDONNELL² Verian

RELEVANCE & RESEARCH QUESTION Among end clients and decision makers, each individual engages differently with the results of market research studies or opinion polls: Some read the headlines, some look at the charts, some read the entire report. The Artificial Intelligence (AI) community has made strides in automating text generation, but their promise of efficiency gains comes with the caveat of lacking trustworthiness.

Hence, for this contribution we ask three questions: How can we leverage AI to accurately describe our quantitative results? How can we tune this output so that it helps us produce reports more efficiently? How can we ensure AI-generated text can be trusted to be correct?

METHODS & DATA Verian Germany and Inspirient have worked together for the past three years to make Generative AI applicable to quantitative survey data through automated statistical reasoning. These prior results comprise both visual output of analyses (incl. charts) as well as corresponding formal chains of reasoning steps, which ensures that results can be linked back to source data and thus trusted. We now combine these into “speaker notes” for a Large Language Model (LLM) that we then utilize to generate descriptive textual output for each analytical result.

RESULTS Our system is able to generate descriptive text for typical charts that one may find in a survey report. We explain how to set up LLMs to accurately link their output back to their speaker notes, and how to control for this as part of post-processing. In our evaluation, we illustrate which AI speaker notes are required for which kind of output, which aspects can be controlled via prompting, and we discuss in how far client-ready output is achievable.

ADDED VALUE While our approach does not match the nuanced writing style and proficiency of an experienced human researcher, we can claim with confidence that the speed-up in getting to draft-level output is tremendous – in particular for lengthy reports. We thus envision a setup in which researchers merely need to fine-tune an AI-written draft report, incl. charts and accompanying text, while knowing that the factual statements in this deliverable can be trusted.

MEET YOUR NEW CLIENT: WRITING REPORTS FOR AIS

GEORG WITTENBURG Inspirient
PAUL SIMMERING, OLIVER TABINO Q Agentur für Forschung

RELEVANCE & RESEARCH QUESTION As organizations adopt Retrieval-Augmented Generation (RAG) for their Knowledge Management Systems (KMS), traditional market research deliverables face new functional demands. While PDFs of reports and presentation slides have effectively served human readers, they now are also “read” by AI systems to answer questions of human users—a trend that will only increase going forward. In order to future-proof the reports that are delivered today, this study evaluates information loss when transferring market research insights through different delivery formats into RAG systems. This open question emerged from a discussion at the DGOF KI Forum between market research buyers and suppliers.

METHODS & DATA We frame the transfer of information, incl. research insights, into clients’ KMS as a signal processing problem. The fidelity of the information transfer depends on the data format: Some formats, e.g., pictures of charts, incur an information loss while other formats, e.g., tables, do not. We model this loss using benchmarks for information extraction from different file formats and from graphs. Further, we assess the needs handled by current reporting formats and contrast them with new needs from RAG. This is done through expert interviews and an analysis of research reports from different institutes.

RESULTS Findings indicate that classic formats, while valuable for human interpretation, are not optimal for AI systems. Key limitations include difficulties in extracting information from graphs and styled slides, which lead to altered, de-contextualized, or lost information. Text-heavy reports offer greater compatibility, yet are not optimal either, e.g., when methodology is presented separately from results. Our study suggests that transitioning to complimentary special-purpose deliverables, designed explicitly for AI, enhances the retrieval accuracy of research insights within KMS, and thus for the client.

ADDED VALUE The choice of reporting format is critical for delivering insights to market research clients, especially now that these reports will also be consumed by AI. This study yields insights on new demands and improved formats for reports from suppliers. It also supports buyers of reports in their assessment of proposals and effective ingestion of results into their KMS for optimal information retrieval going forward.

WEDNESDAY, 02.04.2025
09:00 - 10:00 AM
HÖRSAAL A

7.1: SURVEY RECRUITMENT

SESSION CHAIR: CAMILLA SALVATORE
UTRECHT UNIVERSITY, THE NETHERLANDS

DESIGNING PASSWORDS FOR WEB SURVEY ACCESS: THE EFFECT OF PASSWORD LENGTH AND COMPLEXITY ON SURVEY AND PANEL RECRUITMENT

GEORG-CHRISTOPH HAAS, MARIEKE VOLKERT, STEFAN ZINS Institute for Employment Research, Germany

RELEVANCE & RESEARCH QUESTION For online probability surveys that recruit participants via postal invitation letters, passwords are used to manage access to the survey. These passwords serve several purposes, such as, blocking uninvited individuals, and preventing multiple submissions from the same individual. Research on web survey passwords has primarily focused on whether providing a password for survey access affects response rates. However, the chosen password strength, that is, length and complexity, may also affect response propensities. Password length refers to the number of characters in a password. Password complexity involves the set of characters from which the password can be derived (e.g., lowercase letters and numbers). Our research evaluates the effect of password length and complexity on survey access, response, panel registration and linkage consent rates.

METHODS & DATA We implemented a survey experiment by varying password length and complexity during the first wave of a general population online survey. For recruitment, every individual received a postal invitation letter with a web-link and QR-code directing to the survey, along with an individualized password. We conducted a 2×2 experiment that manipulated password length (five vs. eleven characters) and complexity (uppercase letters only vs. uppercase + lowercase letters + numbers). Additionally, we included a group that used the default length and complexity settings of the service hosting the survey (eight uppercase letters). Invited individuals were randomly assigned to one of these five groups across two different probability samples: employees (N=77,173) and welfare recipients (N=99,176).

RESULTS Results show that short as well as long passwords increase the access rate compared to the control group (16.7%, 19.2% vs. 14.9%). The positive effects of the password designs remain for response and panel registration rates. We also find that long passwords have a positive effect on the propensity to consent to linking survey with administrative data.

ADDED VALUE Our research sheds light on an often-overlooked aspect of postal survey invitations for web surveys: Passwords designs. Our talk shows how researchers can strategically design survey passwords, potentially influencing not only survey response rates but also other data quality indicators.

USING PANELISTS SELF-STATED MOTIVATIONS TO CRAFT EFFICIENT TARGETED EMAIL INVITATIONS TO AN ONLINE PROBABILITY PANEL

GEORG-CHRISTOPH HAAS¹, BENJAMIN BAISCH¹, MARK TRAP-PMANN^{1,2}, JONAS WEIK¹ ¹Institute for Employment Research, Germany; ²University of Bamberg

RELEVANCE & RESEARCH QUESTION In online panels, emails are a crucial element for recruiting respondents. Email invitations may substantially affect panelists' perception of the study's relevance, potentially influencing both response rates and sample composition. Previous research has examined the use of targeted appeals, where the wording in the invitation letter varies among pre-identified subgroups. In our study, we divide panelists into subgroups based on their self-stated motivations to participate. We then use these self-stated motivations to craft an appealing email invitation to invite panelists to a subsequent wave. Based on the Leverage-Saliency Theory, emphasizing the self-stated motivation in email invitations (Saliency) should have a positive effect on the panelists' response propensity, enhancing cooperation as well as reducing attrition within the panel. Our design enables us to answer the question: Do targeted invitations based on panelists self-stated motivations from a previous wave increase response rates in a subsequent wave?

METHODS & DATA We implemented a survey experiment in a German Online Probability Panel: IAB-OPAL. In wave 3, we asked 10,246 panelists to state their main motivation for participation, choosing among seven different motivations: topic, incentive, giving opinion, informing politics, curiosity, helping science, feeling obligated. In wave 4, we randomly assigned panelist either to the standard invitation or to an invitation that aligns with one of the self-stated motivations. Our treatment included a different subject line as well as a motivational email text.

RESULTS Results show that our treatment did not improve cooperation or reduce attrition within the panel. On the contrary, for the motivations "giving opinion" and "informing politics", results show that aligning the wording of the invitation email with panelists self-stated motivations from the previous wave reduces response rates compared to the standard invitation email.

ADDED VALUE Aligning panel communication documents with panelists' underlying motivations holds the promise to enhance cooperation and reduce attrition within the panel. At least for email invitations, we find that our results break with this promise, and lead to unexpected results regarding the underlying leverage-saliency-theory. Therefore, we find it noteworthy to share and discuss these results with experienced colleagues in the online survey community.

BACKING UP A PANEL WITH PIGGYBACKING – THE EFFECT OF PIGGYBACKING RECRUITMENT ON NONRESPONSE BIAS AND PANEL ATTRITION IN A MIXED MODE PANEL SURVEY

BJÖRN ROHR GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Sampling and recruiting respondents for (online) probability-based panels can be very expensive. One cost-intensive aspect of the process is drawing a separate sample and recruiting the respondents offline. To reduce the cost of panel recruitment, some mixed-mode or online panels (e.g., the GESIS Panel, the German Internet Panel, and the NatCen Panel) relied on piggybacking in some recruitments or refreshments. Piggybacking means that participants for the panel are recruited at the end of another probability survey so that no additional sample has to be drawn. Although this reduces the cost of panel recruitment, it might also introduce additional nonresponse. Whether or not the higher amount of nonresponse also translates to higher amounts of bias in practical applications of a piggybacking survey will be analyzed in my research. In addition to the bias for the initial recruited panelists, we will also investigate the effect piggybacking has on panel attrition.

METHODS & DATA To answer the research question, we use the GESIS Panel, a panel survey that was initially recruited in 2013 ($n = 4,961$) from a separate sample but later on refreshed three times with the help of piggybacking ($n = 1,710, 1,607, 764$). This setting allows us to compare the bias of both survey types against each other and disentangle the nonresponse bias introduced by piggybacking in contrast to regular nonresponse bias. To estimate the bias of the separate recruitment waves, we use the German Microcensus as a benchmark. The bias will be measured as a relative bias for demographic and job-related variables, as well as the difference in Pearson's r between benchmark and survey.

RESULTS Initial results show that piggybacking did significantly increase the number of nonrespondents, compared to a separate recruitment. As we are currently preparing the data for the bias analysis, actual results regarding the bias need to be added later.

ADDED VALUE Our work will give researchers a better understanding of the bias introduced through piggybacking and if this method is still a useful tool to reduce the cost of a probability (panel) survey without introducing high amounts of nonresponse bias.

WEDNESDAY, 02.04.2025
09:00 - 10:00 AM
HÖRSAAL B

7.2: DIGITAL BEHAVIOR AND DIGITAL TRACES

SESSION CHAIR: JULIAN KOHNE
GESIS - LEIBNIZ INSTITUTE FOR THE SOCIAL SCIENCES,
GERMAN

PRESENTATIONS

WHERE YOU ARE IS WHAT YOU GET? SAMPLE INCONSISTENCIES OF GOOGLE TRENDS DATA ACROSS DOWNLOAD LOCATIONS

JOHANNA HÖLZL, FLORIAN KEUSCH, JOHN COLLINS University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION Researchers increasingly use digital trace data sources such as Google Trends as an alternative or complement to survey data. However, besides technical limitations and issues of external and internal validity, several researchers have noticed issues with Google Trends' reliability. The data are based on an unknown sample of all Google searches. Downloading Google Trends data for the exact same parameters (i.e., search term, region, time) but at different points in time can therefore produce unreliable values on Google Trends' search index, especially for queries with low search volume. In this paper, we extend the research on Google Trends' reliability beyond the retrieval date by examining the effect of the download location on inconsistencies across samples: Do we get different values from Google Trends depending on where we download the data?

METHODS & DATA We retrieved Google Trends data for the same regions, time periods, and terms from four different countries on three continents (Austria, Germany, the U.S., and Australia). We then compared the search index values retrieved from each respective country to those downloaded in the other countries, keeping all parameters of the query constant.

RESULTS Our results show that values from Google Trends differ across download locations depending on the download day and the query's total search volume. Researchers can minimize these inconsistencies by averaging samples from several days for high search volume queries. Nevertheless, our results point to an additional limitation regarding the reliability and replicability of Google Trends data for its usage in social science research.

ADDED VALUE Our findings help researchers working with Google Trends data in making their research better replicable by averaging samples from several days for high search volume queries. Our results also serve as a tail of caution for research relying on APIs that provide samples of their digital trace data as the download location might impact the findings.

ONLINE LABOUR MARKETS IN THE CONTEXT OF HUMAN RIGHTS AND ENVIRONMENTAL DUE DILIGENCE

FABIAN BRAESEMANN 1,2, MORITZ MARPE 1 1 Datenwissenschaftliche Gesellschaft DWG Berlin, Germany; 2 Oxford Internet Institute

RELEVANCE & RESEARCH QUESTION Online labour markets (OLMs) reflect the globalisation of the past three decades, combined with accelerating digitisation, and are poised to reshape the future of work. For highly educated workers in developing and emerging economies, OLMs offer significant income opportunities. However, existing literature highlights issues such as insufficient regulation, lack of transparency, and inadequate policy focus. Recently, emerging frameworks like the German Act on Due Diligence in Supply Chains (Lieferkettengesetz, LkSG) have introduced legal mechanisms to address human rights violations in global value chains. These frameworks could also help regulate OLMs by requiring clients to exercise due diligence. This obligation, however, depends on the ability to identify clients and assign them corresponding responsibilities.

This study addresses two key research questions:

1. How can labour rights principles, such as the duty of care outlined in emerging supply chain regulations like the LkSG, be applied to OLMs?
2. How can digital trace data from OLMs be used to identify clients, assess their outsourcing behaviour, and evaluate compliance with these regulations?

METHODS & DATA We build on digital trace data from Braesemann et al. (2022) and the Online Labour Index (Stephany et al., 2021) who compile data on freelance projects from platforms like UpWork and Fiverr. Our analysis focuses on project histories to examine client outsourcing behaviour. Metrics such as wages, working hours, and gender imbalances are also assessed.

RESULTS Our study demonstrates the feasibility of identifying clients through project-ID matching algorithms, using a sample of 250 project IDs. Results show that small companies dominate outsourcing activities on OLMs. Wage distributions across case studies in Serbia, Egypt, and Bangladesh reveal that average freelance wages often exceed local minimum wages. However, significant variations exist across occupations and genders, underscoring the need for targeted policy interventions to ensure fair pay and gender equity.

ADDED VALUE This study highlights the potential of supply chain regulations to address regulatory gaps in OLMs by enforcing minimum wage standards and addressing gender disparities. It also advances methods for identifying and analysing clients on OLMs, providing actionable insights for policymakers and researchers.

WEDNESDAY, 02.04.2025
09:00 - 10:00 AM
HÖRSAAL C

7.3: QUESTIONNAIRE DESIGN

SESSION CHAIR: YFKE ONGENA
UNIVERSITY OF GRONINGEN, NETHERLANDS, THE

PRESENTATIONS

B:RADICAL – CO-DESIGNING ONLINE SURVEY QUESTIONS WITH CHILDREN AND YOUNG PEOPLE ON THEIR UNDER- STANDING AND THEIR EXPERIENCES OF RESPECT AND DISRESPECT

DIRK SCHUBOTZ Queen's University Belfast, United Kingdom

RELEVANCE & RESEARCH QUESTION In this presentation, I will showcase how children and young people can be involved as co-researchers and research advisors in designing online survey questions and analysing survey results collected in large-scale online social attitude surveys conducted among children and young people. There is a growing interest in participatory and collaborative research approaches and an increasing recognition that the involvement of co-researchers can enhance the relevance, validity and meaningfulness of research data. Traditionally, collaborative research approaches predominantly utilised qualitative research methods, but more recently survey researchers have also been more open to adopting co-production approaches. In this study I will showcase such a collaborative approach to online surveys involving children and young people.

METHODS & DATA This study is grounded in a rights-based approach to undertaking research with children and young people, using the Lundy Model with its four pillars (Voice, Space, Audience, Influence) to ensure that children and young people have a say, and their voices are listened to, when it comes to matters affecting their lives – here their experiences of respect and disrespect. Two Research Advisory Groups were formed and children and young people were trained as peer-researchers. Starting with explorative philosophical workshops, research questions were developed bottom-up, and were then included in two large-scale annual social attitude surveys undertaken in Northern Ireland – the YLT survey of 16 year olds and the KLT survey of 10-11 year olds.

RESULTS In this presentation, I will largely focus on the processes of designing online survey questions with children and young people as co-researchers, and I will reflect on the challenges of doing this. I may

be able to present some very preliminary survey results, as the survey data collections closes at the time of the GOR conference.

ADDED VALUE This presentation will be of value to those who are designing online surveys – in particular, but not exclusively, surveys for children and young people, and who are considering taking a collaborative co-production approach to data collection and analysis.

DOES SUCCEEDING ON ATTENTION CHECKS MODERATE TREATMENT EFFECTS?

SEBASTIAN LUNDMARK University of Gothenburg, Sweden
JON KROSNICK Stanford University, USA
LISANNE WICHGERS Lisanne Wichgers Consulting
MATTHEW BERENT Matt Berent Consulting

RELEVANCE & RESEARCH QUESTION Attention checks have become a common tool in questionnaires administered online. Confirming their popularity, 58 out of the 139 articles published in the Journal of Personality and Social Psychology in 2019 featured an online experiment where at least one attention check was used to exclude participants. If participants do not pay attention to survey questions and treatments, Type-II error may be inflated, increasing the likelihood of false negative results. A few studies have found that excluding participants who failed the attention check strengthened treatment effects, presumably because excluding them reduced noise and increased data quality. However, participants failing attention checks typically differ from passers in characteristics such as age, gender, and education, so excluding failers compromises sample representativeness and reduces sample size.

METHODS & DATA To assess the impact of excluding participants who fail attention checks on treatment effects and sample composition accuracy, data from sixty-six experiments were analyzed. In all experiments, online respondents were randomly assigned to one of two experimental conditions, with 750 people completing the questionnaire in each condition. This allowed for the assessment of whether the treatment effect became stronger when excluding attention check failers, the degree to which failing rates differed between the treatment and control conditions (which would compromise internal validity if failers are dropped), and the degree to which dropping failers compromised sample distribution of demographic characteristics.

RESULTS The results indicated that attention checks only weakly moderated treatment effects. Participants who failed the attention check showed statistically significant treatment effects despite, ostensibly, not paying attention. Including or excluding the failing participants did not alter any of the conclusions made about each of the sixty-six experimental treatment effects.

ADDED VALUE This meta-analytical study adds to the growing research investigating the appropriateness of attention checks in online questionnaire administration. The study allowed for differentiating whether certain types of attention checks were more efficient in detecting inattentive participants. Lastly, the study results add insights into how excluding data from participants failing attention checks affects a sample's resemblance to the general population in terms of several demographic characteristics.

BALANCING QUESTIONNAIRE LENGTH AND RESPONSE BURDEN: SHORT AND LONG-TERM EFFECTS IN THE IAB JOB VACANCY SURVEY

NICOLE GÜRTZEN, ALEX KÜBIS, ANDRÉ PIRRALHA IAB, Germany

RELEVANCE AND RESEARCH QUESTION Declining survey response rates are a global concern, also affecting official statistics and establishment surveys. This trend is evident in Germany, raising questions about what factors influence participation. Prior research indicates that longer questionnaires can deter respondents due to increased time and effort, leading to higher response burden and lower response rates. While this effect has been shown in household surveys, its impact on establishment surveys should also be researched. This study examines how questionnaire length affects response rates and response burden in the IAB Job Vacancy Survey, a relevant data source for understanding labor demand and recruitment in Germany. **Methods & Data:** In the 2023 wave of the IAB Job Vacancy Survey, around 2000 establishments were randomly assigned to receive either a concise two-page questionnaire or a detailed four-page version, differing in the number of questions. The survey employed a mixed-mode design, combining self-selected web or paper modes of data collection. This experimental design was implemented over three consecutive quarters, allowing the estimation of immediate and long-term effects throughout the year. By comparing the response behavior between the two groups over time, we assessed the influence of questionnaire length on participation and perceived burden.

RESULTS The findings show no significant differences in response rates between the two groups across the three quarters. However, establishments who received the longer questionnaire reported significantly higher levels of perceived response burden. This effect was particularly pronounced in the paper mode of data collection, where respondents expressed greater burden compared to those using the web mode. This increased burden did not translate into decrease short-term participation but may have implications for respondent satisfaction a data quality.

ADDED VALUE These results are significant for the design of establishment surveys and the production of official statistics. They suggest that reducing questionnaire length can lower response burden without significantly affecting response rates. This insight supports efforts to optimize the IAB Job Vacancy Survey's push-to-web design, aiming to enhance respondent experience and maintain high-quality data collection.

WEDNESDAY, 02.04.2025
12:00 - 01:15 PM
HÖRSAAL A

10.1: RELUCTANT RESPONDENTS AND ITEM NON- RESPONSE

SESSION CHAIR: DANIELLE REMMERSWAAL
UTRECHT UNIVERSITY, THE NETHERLANDS

PRESENTATIONS

ENCOURAGING REVISION OF 'DON'T KNOW' RESPONSES: COMPARING DELAYED AND DYNAMIC FEEDBACK IN WEB SURVEYS

ANKE METZLER, STELLA CZAK, HANNAH SCHWÄRZEL, MAREK FUCHS
Technical University of Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION In Web surveys, the absence of interviewers increases the risk of item missing. Therefore, general wisdom suggests to refrain from don't know (DK) options as they may encourage respondents to satisfice (Krosnick & Fabrigar, 1997). This, however, may lead to situations where respondents who cannot generate a valid answer randomly select one of the substantive response categories.

Previous studies indicate that interactive feedback can effectively improve response quality (Zhang, 2013; Al Baghal & Lynn, 2015). Interactive feedback can be provided either (1) after the questionnaire page is submitted (delayed feedback) or (2) immediately after respondents chose DK before they submit the page (dynamic feedback).

In this study, we apply interactive feedback in difficult single-choice questions that offer an explicit DK. If respondents select DK, we follow-up with either delayed or dynamic feedback to clarify question content. We assume that dynamic feedback is more effective in reducing DK since the feedback is provided while respondents are still engaged in the response process.

METHODS & DATA In a Web survey, conducted with a German online access panel in November 2024 (n=2,000), we implemented a be-

tween-subjects design experiment. In two single-choice questions the effectiveness of providing dynamic feedback (EG1) or delayed feedback (EG2) is evaluated compared to a control group providing no feedback (CG).

RESULTS Preliminary results indicate that both feedback types reduce the percentage of final DK responses. In the first experiment, positioned early in the questionnaire, delayed feedback appears to be more effective in reducing DK than dynamic feedback. In contrast, in the second experiment, placed later in the questionnaire, dynamic feedback exhibits stronger effects. T

These findings suggest that delayed feedback may be more effective when respondents are highly motivated, whereas dynamic feedback seems to reduce DK to a greater extent as respondent motivation decreases.

ADDED VALUE This study provides insights for survey researchers seeking to minimize DK answers and improve data quality in web surveys. By examining the distinct effects of dynamic versus delayed feedback on the revision of DK answers, this study helps understand how the timing of feedback influences respondent behavior.

ZOOMING IN: MEASURING RESPONDENT 'S REACTANCE AND RECEPTIVITY TO ASSESS THE EFFECTS OF ERROR-REDUCING STRATEGIES IN WEB SURVEYS

STELLA CZAK, HANNAH SCHWÄRZEL, ANKE METZLER, MAREK FUCHS
Technical University of Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION In Web surveys, respondents often exhibit signs of satisficing behavior, like speeding, non-differentiation, or item nonresponse. A common strategy to reduce such behavior uses prompts and interactive feedback to respondents (for example Baghal & Lynn, 2015; Kunz & Fuchs, 2019). However, the effectiveness of prompts is sometimes limited. Some respondents seem to react with an optimizing tendency while others tend to ignore such prompts. This raises concerns that prompts do not reach all respondents to the same extent.

In this study we assess the interaction of two personality traits with the effectiveness of prompts concerning non-differentiation, item nonresponse and other types of satisficing behavior. It is assumed that the respondents' level of reactance, which describes a person's inner resistance to restrictions on one's own freedom of action, prevents respondents from improving their answering behavior when exposed to a prompt.

By contrast, a person's receptivity describes the likelihood of a positive change in behavior due to interventions. It is assumed that respondents with higher levels of receptivity perceive prompts as a helpful resource and change their response behavior to the good.

METHODS & DATA In a web survey conducted in a general population sample drawn from a non-probability based online access panel (n=2,000) on "AI and digitalization" a series of experiments evaluat-

ing prompts on various types of satisficing behavior such as non-differentiation and item nonresponse have been implemented.

Also, two validated German psychometric scales were utilized to measure how responsive respondents are to interventions: One on reactance [Herzberg, 2002] and one on receptivity to instructional feedback [Bahr et al., 2024].

RESULTS Field work is still underway. In the analysis we aim to test, whether these personal traits are able to explain the differential effectiveness of the various established satisficing prompts.

ADDED VALUE The study contributes to a better understanding of the differential effectiveness of satisficing prompts. Based on the results we aim to tailor the frequency, the presentation and the wording of prompts to the respondents' psychometric profile. We assume that an improved respondent experience may foster effectiveness of interventions and ultimately data quality.

UNDERSTANDING ITEM-NONRESPONSE IN OPEN QUESTIONS WITH REQUESTS FOR VOICE RESPONSES

CAMILLA SALVATORE 1, JAN KAREM HÖHNE 2,3 1 Utrecht University, Netherlands, The; 2 German Center for Higher Education Research and Science Studies (DZHW); 3 Leibniz University Hannover

RELEVANCE & RESEARCH QUESTION Conducting smartphone surveys offers flexibility in collecting various types of responses. Among these various response modalities, voice responses stand out for their potential to facilitate deeper respondent engagement and expression. However, high item-nonresponse rates pose significant challenges to their large-scale use. Our research question is the following: which and to what extent socio-demographic characteristics, technological skills, and survey related aspects are associated with item-nonresponse?

METHODS & DATA In this web survey study, we use data that was collected in the Forsa Omninet Panel in Germany in November 2021. Forsa drew a cross-quota sample from their online panel based on age (young, middle, and old) and gender (female and male). In addition, they drew quotas on education (low, medium, and high). The quotas were calculated based on the German Microcensus, which served as a population benchmark. For recording respondents' voice responses, we implemented the open-source "SurveyVoice" (SVoice) tool that was implemented in the Forsa web survey system. The sample includes data from 501 respondents. Overall, we investigate data from eight open questions with requests for voice responses.

RESULTS Two main respondent groups emerged from the data: voice skippers (about 30% of respondents who skipped all voice questions but answered closed questions), and voice engagers (those who responded to at least one open voice question). Male respondents were more likely to be voice skippers. In contrast, respondents who expressed interest in the survey topic were less likely to be voice skippers. Among voice engagers, response rates were higher among those with good smartphone skills and positive survey perceptions (e.g. evaluating the survey as easy, not long, not intimate, and interesting). Response rates and response lengths also varied by question topic.

ADDED VALUE The study presents key insights into the characteristics associated with item-nonresponse in open voice questions. It provides an important starting point for future studies that aim at enriching web survey data collection through rich respondent narrations across various respondent groups.

WEDNESDAY, 02.04.2025

12:00 - 01:15PM

HÖRSAAL B

10.2: AI AND AUTOMATION IN (SURVEY)

SESSION CHAIR: DANIELLE REMMERSWAAL
UTRECHT UNIVERSITY, THE NETHERLANDS

PRESENTATIONS

BOTS IN WEB SURVEYS: PREDICTING ROBOTIC LANGUAGE IN OPEN NARRATIVE ANSWERS

JOSHUA CLAASSEN, JAN KAREM HÖHNE DZHW; Leibniz University Hannover

RUBEN BACH University of Mannheim

RELEVANCE & RESEARCH QUESTION Web survey data is key for social and political decision-making, including official statistics. Respondents are frequently recruited through online access panels or social media platforms, making it difficult to verify that answers come from humans. As a consequence, bots – programs that autonomously interact with systems – may shift web survey outcomes and social and political decisions. Bot and human answers often differ regarding word choice and lexical structure. This may allow researchers to identify bots by predicting robotic language in open narrative answers. In this study, we therefore investigate the following research question: Can we predict robotic language in open narrative answers?

METHODS & DATA We conducted a web survey on equal gender partnerships, including three open narrative questions. We recruited 1,512 respondents through Facebook ads. We also programmed two AI-based bots that each ran through our web survey 100 times: The first bot is linked to the LLM Gemini Pro, and the second bot additionally includes a memory feature and adopts personas, such as age and gender. Using a transformer model (BERT) we attempt to predict robotic language in the open narrative answers.

RESULTS Each open narrative answer is labeled based on whether it was generated by our bots (robotic language = "yes") or the respon-

dents recruited through Facebook ads (robotic language = “unclear”). Using this dichotomous label as ground truth, we will train a series of prediction models relying on the BERT language model. We will present various performance metrics to evaluate how accurately we can predict robotic language, and thereby identify bots in our web survey. In addition, we compare these results to students’ predictions of robotic language to study whether our BERT models outperform human judgement.

ADDED VALUE Our study contributes to the ongoing discussion on bot activities in web surveys. By investigating AI-based bots with different levels of sophistication that are linked to a LLM, our study stands out from previous research that mostly looked at less sophisticated rule-based bots. Finally, it extends the methodological toolkit of social research when it comes to identifying bots in web surveys.

ADDRESSING BIASES OF SENSOR DATA IN SOCIAL SCIENCE RESEARCH: A DATA QUALITY PERSPECTIVE

VANESSA LUX 1, LUKAS BIRKENMAIER 1, JOHANNES BREUER 1, JESSICA DAIKELER 1, FIONA DRAXLER 2, JUDITH GILSBACH 1,3, JULIAN KOHNE 1,4, FRANK MANGOLD 1, INDIRA SEN 2,5, HENNING SILBER 6, KATRIN WELLER 1,7, MAREIKE WIELAND 1

1 GESIS - Leibniz Institute for the Social Sciences, Germany; 2 University of Mannheim, Germany; 3 University of Konstanz, Germany; 4 Ulm University, Germany; 5 RWTH Aachen, Germany; 6 University of Michigan, USA; 7 University of Düsseldorf, Germany

RELEVANCE & RESEARCH QUESTION

Sensor data – social sciences – error sources – error framework

The everyday availability of sensors has opened new research avenues for the social sciences, including their combination with traditional data types, such as survey data. However, as sensors become more prevalent for the collection of digital behavioral information, concerns regarding the accuracy and reliability of the obtained sensor data have emerged. Error sources and biases of sensor data are very sensor-specific, which poses a challenge to social science researchers as the necessary technical expertise is often lacking. The paper gives an overview of these concerns and proposes a general error framework for the data quality assessment of sensor data in social science research, contributing conceptually and methodologically to enhance the assessment and reporting of sensor data quality.

METHODS & DATA Systematic review – thematically focused content analysis – expert group

Sensor error framework dimensions were extracted based on the results of a thematically focused systematic review (see preregistration here: <https://osf.io/vkxbt>) using qualitative content analysis and evaluated within an expert group.

RESULTS Data quality – error framework – technical and human error – measurement error – representation bias

The proposed error framework outlines error sources and potential biases for measurement and representation along the full research cycle (planning, data collection, data analysis, archiving and shar-

ing). We addressed the intricate relationship between general data quality dimensions and sensor-specific error sources by incorporating the multilayered character of sensor data arising from technical affordances and device effects. In addition, we identified three principles structuring error sources and biases for specific sensors: The interplay between researcher, study participant, and device, the spatial mobility of the sensor, and the continuous character of the error sources. The adoption of the framework is illustrated with sensor-specific examples.

ADDED VALUE Data quality assessment – reporting standards – replicability – interpretability

The proposed general error framework for sensor data bears the potential to enhance the assessment and reporting of sensor data quality in the social sciences. It provides guidance to researchers and facilitates better replicability and interpretability of sensor data.

IMPROVING THE MEASUREMENT OF SOLIDARITY IN THE EUROPEAN CONTEXT: RESULTS FROM A WEB PROBING IN FOUR COUNTRIES

VERA LOMAZZI, MARGHERITA PELLEGRINO University of Bergamo

RELEVANCE & RESEARCH QUESTION This research addresses how cultural biases affect cross-national comparability in attitudes toward solidarity. Recent studies highlight concerns about the comparability of solidarity measurements between countries. By implementing international web-probing, we aim to uncover these biases and improve the clarity of questions in future rounds of the EVS questionnaire to ensure reliable cross-country comparisons.

Cross-country comparability, solidarity, European Values Study.

METHODS & DATA We conducted web probing in – Italy, Portugal, Hungary, and the Czechia – utilizing nine solidarity-related items from the EVS 2017 questionnaire. The method involved inserting probes following closed-ended questions to explore respondents’ interpretations. A sample of 600 participants was surveyed, with responses analyzed qualitatively to identify variations in how terms like “Europeans,” “immigrants,” and “concern” are understood, and explaining why they chose to feel a certain level of concern toward Europeans. This data was translated and categorized using thematic coding across languages.

Web probing, open-ended questions, codification.

RESULTS For each response, we identified multiple categories, demonstrating the diverse interpretations of the same word within and especially across countries, despite the accurate translation of the EVS questionnaire.

For instance, in Italy and Hungary, the word “concern” was primarily interpreted as “cherish” or “care,” while in Portugal, it was more commonly identified with being “worried.” Similarly, the term “immigrants” was understood in multiple ways. In Hungary and Portugal, it was often perceived as referring to generic migrants, whereas in the Czech

Republic and Italy, respondents specified particular categories. For instance, 24 respondents from the Czech Republic described migrants specifically as Ukrainians, highlighting how contextual and temporary these interpretations can be.

Cultural variations, language, context, interpretation.

ADDED VALUE This study demonstrates the value of web probing as a tool for identifying and addressing cultural biases in international surveys. The insights gained provide a basis for refining survey instruments, ensuring that data on solidarity reflects a more accurate and culturally sensitive understanding across European countries. Cultural biases, comparative implications, data quality improvement.

parative analysis of creator performance stratified by gender. Using a unique data set of 36,166 videos that have been sampled using a rigorous hourly sampling approach to access content performance over a 41 day period.

ADDED VALUE My methodology demonstrates the feasibility of video content analysis on TikTok, contributing to both the theoretical understanding of digital creator economies and the methodological toolkit for social media research. It also demonstrates how embedding models can be leveraged for social scientific studies of visual data, which is still scarce in the field.

RESULTS I find differences in the distribution of content types across genders. The choice of content type is also related to the success of generating user engagement (views / likes). However, individual level intercepts also strongly predict performance of shared content.

WEDNESDAY, 02.04.2025
12:00 - 01:15PM
HÖRSAAL C

10.3: EXPLORING REPRESENTATION IN SOCIAL MEDIA

SESSION CHAIR: JESSICA DONZOWA
MAX PLANCK INSTITUTE FÜR DEMOGRAPHISCHE
FORSCHUNG, GERMANY

PRESENTATIONS

DANCING WITH DATA: UNDERSTANDING GENDER REPRESENTATION AMONG VIRAL TIKTOK CONTENT

DORIAN TSOLAK Bielefeld University, Germany

RELEVANCE & RESEARCH QUESTION Social media platforms have evolved into significant ecosystems, where content creators derive income from their digital presence. Despite its prominence as one of the biggest social media platforms, TikTok remains understudied in social science literature, primarily due to the methodological challenges in analyzing video content.

METHODS & DATA This paper presents a novel computational approach to evaluate content creators' performance through systematic video data analysis. By embedding the videos via VideoMAE, a state of the art model for embedding of visual data, I examine the performance variations across different content categories and conduct a com-

BEYOND BINARY BYTES: MAPPING THE EVOLUTION OF GENDER INCLUSIVE LANGUAGE ON TWITTER

**DORIAN TSOLAK 1,2, SIMON KÜHNE 1, STEFAN KNAUFF 1,2,
H. LONG NGUYEN 2,3, DOMINIK HANSEN 1**

1 Bielefeld University, Germany; 2 Bielefeld Graduate School
in History and Sociology; 3 DeZIM Institute

RELEVANCE & RESEARCH QUESTION Languages worldwide differ significantly in how they incorporate gender into grammar and phonetics. In the German language, the generic masculine form (e.g., saying "Lehrer" [teacher, male, sing.]) is used to refer to a group of people with unknown (or non-male) sex and has been criticized for rendering women and non-binary people invisible in language, thereby reinforcing gender biases and unequal power dynamics. Gender-inclusive language (GIL) has been proposed as an alternative to the generic masculine and involves various subtypes. Our study investigates the development of GIL on Twitter between 2018 and 2023. In addition, we study individual (gender) and contextual (regional) effects on the use of GIL.

METHODS & DATA We rely on a unique dataset of over 1 Billion German language Tweets. We present a pipeline to detect three types of GIL, namely binary feminization, non-gendered GIL and non-binary inclusive language. We do this through a combination of using a fine-tuned German BERT model, regular expressions, and a corpus of German gender-inclusive language words. User names are analyzed based on lists of male, female and unisex names. By inferring the place of residence for the users of more than 300 million Tweets, we shed light on the correlations of socio-structural variables and use of gender-inclusive language across Germany.

RESULTS We find that GIL adoption increases slightly over the studied 5 year period and we identify different trends among GIL types in this adoption. Furthermore, profiles with female usernames use GIL more often than those with masculine or unisex usernames. In addition, we find regional patterns with more use of GIL in urban regions and regions with a higher share of users with young population.

ADDED VALUE Our study makes several novel contributions to the understanding of gender-inclusive language adoption and digital so-

cio-linguistics. First, it provides insights into the real-world uptake of gender-inclusive language through the largest-scale analysis of German social media communication to date. Second, by linking language use to regional socio-structural variables, we offer the first comprehensive geographic analysis of gender-inclusive language adoption patterns in Germany.

ROMANCE DAWN: INVESTIGATING THE DYNAMICS OF COLLABORATION IN A CULTURAL PRODUCER COMMUNITY ON YOUTUBE

PAVEL DIMITROV CHACHEV Social Monitor, Romania
MARCEL KAPPES University of Mannheim

RELEVANCE & RESEARCH QUESTION Online communities are widely acknowledged to provide new opportunities for meaningful interaction between individuals with similar tastes in cultural consumption. However, the flip side of this coin – that online communities likewise provide new social connection opportunities for producers of cultural content – has received much less attention so far.

This project, therefore, starts with the premise that cultural producers foster “competitive co-operation” among themselves for mutual benefits through collaboration. We attempt to explore the specific dynamics of how such collaborations fosters growth and success of individual producers as well as of their community as a whole.

METHODS & DATA To investigate these and related claims empirically, we trace the development of an online community of cultural producers creating video content on a popular Japanese manga on YouTube. We develop a new interactive interface allowing for the systematic coding and preprocessing of Youtube data. The coding interface allows for the coding of collaboration cues: words and phrases indicating collaborations (e.g. “feat”, “w”, “@”). With this we construct a longitudinal collaboration network where cultural producers of this specific community share a directed tie if they have jointly published a video on YouTube, with the host as a source and the guest as destination.

RESULTS The results of negative binomial regression estimating view count indicate great benefits of being invited by other youtubers. A tie to a new host, i.e. an increase in indegree, heightens the yearly viewer yield of a youtuber by around 20%. Inviting new guests, i.e. an increase in outdegree, also increases the yearly viewer yield of a youtuber, however only by around 7%.

ADDED VALUE This study contributes to the study of online producers in two way. First we show a method how to code and preprocess youtube data. Second, we explore specific effects taking place in a cultural producer community pointing out very different effects of in- and outdegree. Both give insight into the dynamics behind emergence of communities online.

WHERE IS EVERYBODY? MEASURING SEMANTIC SOURCE POSITION AND CREATING ONLINE DISCOURSE TYPOLOGIES FROM CO-OCCURRENCE NETWORKS

DAN SULTANESCU, PAVEL DIMITROV CHACHEV, DANA C. SULTANESCU
Social Monitor, Romania

RELEVANCE & RESEARCH QUESTION Tremendous advances have been made in measuring the meaning of text data allowing us to quantify entire discourses within unified data structures such as a vector spaces or networks. What is often lost in the application of such methods is the perspectives that different types sources or actors may contribute to a discourse and how such perspectives may come to shape it. We demonstrate a method for evaluating the different positions text sources can have in co-occurrence networks, what we call the sources semantic position. We then create a typology of the networks that emerge based on the semantic position of the actors.

METHODS & DATA We use online data from a sample of leading US sources collected with the NewsVibe platform, which allows access to web news articles and facebook posts based on search queries. We sample articles around the 2024 US election. The co-occurrence network for each source is computed first, and then all keyword networks are added to a global co-occurrence network. This allows us to determine the position of each source within the discourse according to the keyword. Finally we evaluate the different types of networks that emerge from such an analysis.

RESULTS We show that depending on the search query, networks as well as source contribution can be highly contextual. Discourses around “economics”, but also discourses around candidates like Trump and Kamala are highly polarized. The sources in each of these cases split into almost separate parts of the network displaying what one could call semantic polarization. However networks around topics such as NATO are governed by strong influence by a few actors and very little participation by other sources.

ADDED VALUE We further the study of semantic networks by demonstrating how to measure the individual sources’ position to a discourse. By keeping the computations rather simple, non-expert third parties can also evaluate discourses by themselves, improving transparency and access to complex interpretation of data around meaningful online issues and actors.

WEDNESDAY, 02.04.2025
02:30 - 03:45 PM
HÖRSAAL A

11.1: NONRESPONSE BIAS AND CORRECTION

SESSION CHAIR: ANKE METZLER
TECHNICAL UNIVERSITY OF DARMSTADT, GERMANY

PRESENTATIONS

DO MACHINE LEARNING TECHNIQUES IMPROVE NONRESPONSE WEIGHTING?

BARBARA FELDERER, BJÖRN ROHR, CHRISTIAN BRUCH GESIS,
Germany

RELEVANCE & RESEARCH QUESTION Nonresponse weighting is an important tool for improving the representativeness of surveys, e.g. weighting respondents according to their inverse propensity to respond (IPW). IPW estimates a person's propensity to respond based on characteristics that are available for respondents and non-respondents. While logistic regression is typically used to estimate the response propensity, machine learning methods offer several advantages: they allow for very flexible estimation of relationships and the inclusion of a large number of potentially correlated predictor variables. ML methods are known to predict values very accurately. However, it is also known that the estimation of relationships of the weighting variables and the response propensity suffers from regularization bias. With regard to weighting, it is unclear which of these properties is more relevant and has a greater influence on the quality of the weighted estimate. In this study, we address the question of whether machine learning methods outperform logistic regression in performing IPW.

METHODS & DATA In a simulation study that mimics the three nonresponse models (separate cause model, common cause model, survey variable cause model) and varies the number of features that affect nonresponse, we apply IPW weighting using five different prediction models: Regression Trees (CART), Random Forest, Boosting, Lasso, and Logistic Regression. We conclude the analysis with an application to voting decisions collected in the German Internet Panel.

RESULTS Machine learning methods perform similarly well to logit regression and lead to a lower variance in the estimates than logit regression. Overall, the advantage of an excellent prediction seems to outweigh the disadvantages of regularization bias.

ADDED VALUE The presentation provides guidance on how to improve the weighting of surveys, which is a crucial task when drawing conclusions about the general population from a survey.

COMPANY NONRESPONSE IN GENDER INEQUALITY SURVEYS: CHALLENGES IN PARTICIPATION AND IMPLICATIONS FOR DATA QUALITY

ELENA FERRARI, FLAVIA PESCE ARS - Associazione per la ricerca sociale
MARGHERITA PELLEGRINO, VERA LOMAZZI University of Bergamo

RELEVANCE & RESEARCH QUESTION Gender inequality in the labor market is still nowadays a significant issue. However, gathering data from a local and representative sample to understand the barriers faced by employed women in terms of career progression, pay equity, and work-life balance remains challenging. This research aims to examine the difficulties of engaging companies to participate in surveys on gender inequality and the subsequent implications for data quality. Understanding these challenges is crucial for advancing research methods and, in this specific case, for better understanding social inequality and improving future survey methodologies in sensitive areas.

Sampling challenges, company engagement, labor market.

METHODS & DATA A representative sample of 1,400 companies, each employing over 50 workers, from three Italian provinces of (Milan, Bergamo, and Brescia) was contacted to participate in the survey, with the aim of exploring company policies on gender inclusion, work-life balance, career development, and diversity. Challenges related to corporate nonresponse, such as issues with contact information accuracy, email deliverability (e.g., emails marked as spam), and hesitancy to participate, were systematically documented and analyzed. To increase the number of responses several strategies were applied. Our research offers valuable insights into how to increase companies' responses.

Representative sample, survey strategies.

RESULTS We will present our findings on how different strategies for involving companies can impact response rates. We will begin by discussing the initial recruitment attempt and the subsequent strategies employed to increase participation. Finally, we will examine the advantages of this type of research, the complexities inherent in this methodology, and strategies to overcome these challenges. Response rates, methodology.

ADDED VALUE This study provides critical perspectives into the complexities of corporate engagement in gender inequality research. It highlights the importance of overcoming participation barriers to obtain a representative sample, by identifying strategies to improve company engagement, this research can inform future studies and help develop more effective methods to engage local samples and obtain representative results.

Engagement optimization, gender inequality research, research impact.

REASONS FOR PARTICIPATING IN (NON) PROBABILITY ONLINE PANELS AND HOW TO ASK ABOUT IT

TANJA KUNZ, IRINA BAUER GESIS – Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION Knowing why respondents join online panels helps researchers design more engaging recruitment strategies, especially for hard-to-reach populations. Understanding motivations also supports retention strategies by aligning the panel experience with participants' expectations. This study aims to explore the reasons why individuals join online panels and beyond, to assess how different question designs—specifically closed versus open-ended formats asking for single versus multiple responses—affect the measurement of participation motives and whether they differ between nonprobability and probability-based panels.

METHODS & DATA The study was conducted using data collected from two types of online panels: a probability panel recruited via random sampling and a nonprobability panel relying on self-selection. Respondents (Nprob = 19,876 and Nnonprob = 3,256) were randomly assigned to one of four experimental groups: closed question asking for a single response, closed question asking for multiple responses, open-ended question asking for a single reason, and open-ended question asking for multiple reasons. Responses were analyzed to assess differences in the diversity, ranking, and completeness of reasons reported across different formats and panel types.

RESULTS Preliminary findings suggest that the most important reason for participating and the prioritization of influencing factors differ between probability and nonprobability panelists. In addition, open-ended questions lead to slightly more diverse responses, but at the cost of drastically higher nonresponse rates. The different question formats also lead to sometimes striking differences in the ranking of reasons.

ADDED VALUE This study sheds light on the reasons for participation in online panels and highlights the methodological implications of question design.

DEEPER LOOK INTO EDUCATION BIAS IN WEB SURVEYS

MUSTAFA COBAN Institute for Employment Research (IAB), Germany
CHRISTINE DISTLER Institute for Employment Research (IAB), Germany

MARK TRAPPMANN Institute for Employment Research (IAB), Germany

RELEVANCE & RESEARCH QUESTION The COVID-19 pandemic has accelerated a trend in survey research to use online data collection for general population samples. High-quality web surveys recently achieved response rates comparable to or even exceeding those of telephone surveys. However, selection bias concerning education is often more pronounced. To address this issue, we analyze complete employment biographies of both respondents and non-respondents and focus on three main research questions: [1] How do the different stages of the recruitment process for an online panel contribute to education bias?

[2] Are there specific subgroups within the low-educated population who are even less likely to participate? [3] Are there interaction effects between education and other predictors of nonresponse?

METHODS & DATA In 2023, the Institute for Employment Research in Germany launched a new online panel survey of the German workforce (IAB-OPAL) using a push-to-web approach. Addresses were sampled from a powerful database comprising compulsory social insurance notifications by employers as well as unemployment insurance and welfare benefit records. We utilize this unique opportunity of a sampling frame containing detailed individual-level information on complete employment biographies.

RESULTS Our findings indicate that educational bias accumulates at every stage of the recruitment process. We observed that unit nonresponse is more pronounced among individuals with lower education levels, particularly for respondents aged 50 and older and foreign nationals. Additionally, nationality appears to have a greater impact on highly educated individuals, and women are less likely to participate unless they hold advanced degrees.

ADDED VALUE Using a detailed sampling frame that includes individual-level information from complete employment histories enables us to evaluate how educational bias emerges throughout the recruitment process. It also allows us to determine if response tendencies within different educational strata vary based on typically unobserved factors, such as experience with benefit receipt, occupations, or wages.

**WEDNESDAY, 02.04.2025
02:30 - 03:45 PM
HÖRSAAL B**

11.2: BOTS, AVATARS, AND ONLINE LABS

SESSION CHAIR: JOACHIM PIEPENBURG
GESIS, GERMANY

PRESENTATIONS

BOTS IN WEB SURVEY INTERVIEWS: A SHOWCASE

JAN KAREM HÖHNE, JOSHUA CLAASSEN DZHW, Leibniz University Hannover, Germany

SAIJAL SHAHANIA, DAVID BRONESKE DZHW, University of Magdeburg

RELEVANCE & RESEARCH QUESTION Cost- and time-efficient web surveys have replaced other survey modes. These efficiencies can potentially cover the increasing demand for survey data. However,

er, since web surveys suffer from low response rates, researchers start considering social media platforms as sources for respondent recruitment. Although these platforms provide advertisement and targeting systems, the data quality and integrity of web surveys recruited through social media might be threatened by bots. Bots have the potential to shift survey outcomes and thus political and social decisions. This is alarming since there is almost no research on bots and their completion characteristics in web surveys. Importantly, existing research does not distinguish between conventional (rule-based) and sophisticated (AI-based) bots. We therefore address the following research question: Do bots varying in sophistication show different web survey completion characteristics?

METHODS & DATA We programmed four bots with different levels of sophistication: two conventional (rule-based) and two sophisticated (AI-based) bots. To this end, we conducted a literature search to compile a list of capabilities that are key for bots to successfully complete web surveys. The two AI-based bots were additionally linked to the Large Language Model “Gemini Pro” (Google). We ran each bot $N = 100$ times through a web survey on equal gender partnerships and tested several bot prevention and detection measures, such as CAPTCHAs and invisible honey pot questions.

RESULTS The results indicate that both rule-based and AI-based bots come with impressive completion rates (up to 100%). In addition, we can prove conventional wisdom about bots in web surveys wrong: CAPTCHAs and honey pot questions pose no challenges. However, there are clear differences between rule-based and AI-based bots when it comes to web survey completion. For example, conventional bots are faster and select more options in check-all-that-apply questions than their AI-based counterparts.

ADDED VALUE By distinguishing rule-based and AI-based bots, this study stands out of previous studies and shows that both types of bots have different completion characteristics. In addition, it provides a useful framework for future research on the prevention and detection of bots in web surveys.

BRINGING THE LAB ONLINE: DEVICE EFFECTS IN PSYCHOLOGICAL BIAS TESTING IN ONLINE SURVEYS

ELLI ZEY 1, INIOBONG ESSIEN 2, STEFANIE HECHLER 1,3, SUSANNE VEIT 1 1DeZIM Institut, Germany; 2Leuphana University Lüneburg, Germany; 3FU Berlin, Germany

RELEVANCE AND RESEARCH QUESTION Self-report measures for sensitive topics, such as stereotypes and prejudices, are often compromised by social desirability bias. Indirect psychological bias tests offer a promising alternative by measuring implicit biases through reaction times, decision errors under time pressure, priming effects, and memory performance. Traditionally, these tests are conducted in controlled lab environments, which limits the sample size and diversity due to logistical constraints. However, when conducted online, researchers cannot control participants’ choice of device - but device type is associated with systematic differences (e.g. screen size, input method, and test environment) that may influence results. We developed a tool for integrating indirect psychological bias tests into online surveys: MIND.set. This study used MIND.set to investigate two key

questions: 1) How reliably can implicit biases be detected in online survey contexts? and 2) How does the type of device used (e.g., mobile vs. desktop) impact test outcomes?

METHODS AND DATA In 2023, we conducted an online survey with 2,707 participants from the general population, using quotas for gender, age, and education. Participants were randomly assigned to one of five indirect bias tests implemented via the MIND.set platform: Implicit Association Test (IAT), Affect Misattribution Procedure (AMP), Shooter Task (ST), Avoidance Task (AT), and Source Monitoring Paradigm (SMP). All tests focused on stereotypes of Arab-looking men versus White men, specifically regarding perceived threat. Participants self-selected their devices (mobile or desktop), and our preregistered hypotheses (OSF) examined the influence of device type on bias detection and bias scores.

RESULTS The analyses confirmed implicit biases on at least one bias indicator across all five tests. Crucially, bias scores were largely unaffected by device type. While minor variations were observed, these did not significantly undermine the reliability of results across different devices.

ADDED VALUE The MIND.set platform enhances the accessibility of indirect bias testing by offering a robust infrastructure for online research. This study is the first to systematically investigate device effects across multiple indirect bias tests, providing critical insights for researchers seeking to incorporate such tests into online surveys.

ENHANCING OPEN-ANSWER CODING IN QUANTITATIVE SURVEYS USING OPTIMIZED LLM TOOLS

ORKAN DOLAY, DENSI BONNAY Bilendi & respondi, France

RELEVANCE & RESEARCH QUESTION Accurate coding of open-ended responses in quantitative surveys is critical for generating insights. However, traditional manual coding methods are time-consuming and costly. LLMs present an opportunity to revolutionize this process. The research question explored in this study: How can LLM-based coding be optimized to outperform both human coders and baseline LLM implementations in terms of accuracy?

The goal of this research is to better understand how to improve automated coding via foundation models, and to assess the impact on coding quality of various strategies aimed at improving on vanilla LLM use. In particular, we considered the effect of:

1. few shot learning: how helpful is it to provide general or case-based examples?
2. prompt optimization: what is the best way to ask the LLM to perform the seemingly easy task of applying labels to verbatims?
3. input optimization: is there a way to format input labels so as to make it easier for the LLM to correctly apply them?
4. model choice: do newer generation LLMs fare better than older ones? are quicker/lighter models good enough?

METHODS & DATA This research employed comparative tests across 4 coding methods tested on multiple datasets from real-world surveys: (1) Human manual coding by client (=benchmark), (2) Human cod-

ing by external suppliers, (3) Initial implementation of an LLM-based coding tool (BARI V1), and (4) An optimized version of the LLM tool (BARI V2) enhanced through iterative improvements in prompt engineering, training data alignment and feedback loops. Key performance metrics being always coding accuracy (measured against benchmark 'Client').

RESULTS Our results suggest that:

1. general few shot learning is not particularly helpful
2. some prompting strategies do fare better, specially on trickier inputs
3. input format actually crucial the most crucial factor
4. smaller models aren't bad compared to bigger ones

ADDED VALUE This study showcases how optimizations of LLMs can bridge the gap between AI and human in coding open-ended survey responses. The findings provide insights into leveraging AI for more efficient and accurate data analysis, highlighting a transformative approach for researchers, practitioners, and industry stakeholders.

WEDNESDAY, 02.04.2025
02:30 - 3:45 PM
HÖRSAAL C

11.3: **SOCIAL MEDIA** **SURVEYS AND** **RECRUITMENT**

SESSION CHAIR: JOACHIM PIEPENBURG
GESIS, GERMANY

PRESENTATIONS

THE COVID-19 HEALTH BEHAVIOUR SURVEY: A CROSS-NATIONAL SURVEY CONDUCTED VIA FACEBOOK

DANIELA PERROTTA Max Planck Institute for Demographic Research, Germany

RELEVANCE & RESEARCH QUESTION The COVID-19 pandemic affected daily life in unprecedented ways, posing serious challenges for governments and societies. Nonpharmaceutical interventions (NPIs), such as stay-at-home orders, physical distancing measures, and mask mandates, were pivotal in reducing transmission, particularly during the early stages when vaccines were unavailable.

Understanding how populations responded to these interventions was crucial for developing effective communication strategies and policies.

However, the lack of comprehensive data on behaviors and perceptions during the pandemic posed a significant challenge. This study sought to fill this gap by investigating behavioral responses to COVID-19 across diverse demographic groups and countries, examining the interplay between threat perception, preventive behaviors, and compliance with NPIs.

METHODS & DATA To explore these dynamics, we conducted the COVID-19 Health Behavior Survey, a large-scale, cross-national online survey administered across eight countries: Belgium, France, Germany, Italy, the Netherlands, Spain, the United Kingdom, and the United States. Data collection relied on targeted Facebook advertisements, enabling rapid recruitment of participants during the pandemic's initial wave.

The survey, conducted between March 13 and August 12, 2020, yielded over 140,000 responses. It captured detailed information on participants' health status, behaviors, social contacts, and attitudes toward COVID-19. Statistical techniques were employed to address potential sampling biases and ensure robust insights.

RESULTS The results highlighted significant demographic and national differences in pandemic responses. Women and older individuals perceived COVID-19 as a greater threat than men and younger groups, leading to higher adoption rates of preventive measures such as mask-wearing and physical distancing. Threat perception was particularly influential among vulnerable populations, including the elderly and those with preexisting conditions. Social contact patterns also changed markedly, with physical distancing guidelines leading to a 48%-85% reduction in social contacts compared to pre-pandemic levels across surveyed countries, often exceeding the impact of lockdown measures.

ADDED VALUE This study provides valuable cross-national insights into behavioral responses during a global health crisis. By leveraging innovative survey methods and timely data collection, it underscores the importance of understanding population behavior to inform public health strategies and enhance preparedness for future pandemics. The findings offer actionable guidance for evidence-based policy-making and effective risk communication.

ESTIMATING FERTILITY INDICATORS IN LOW- AND MIDDLE-INCOME COUNTRIES: EVIDENCE FROM A NETWORK REPORT- ING ONLINE SURVEY IN SENEGAL

JESSICA DONZOWA 1,2, DANIELA PERROTTA 1, DENNIS FEEHAN 3, EMILIO ZAGHENI 1 1 Max Planck Institute für demographische Forschung, Germany; 2 Bielefeld University, Germany; 3 University of California, Berkeley, USA

RELEVANCE & RESEARCH QUESTION Data availability is often limited in developing countries, with timely administrative or survey data especially lacking. To address this, we propose a novel survey recruitment and estimation approach. First, we recruit survey participants

through Facebook advertisements. While social media surveys are common in high-income countries, they are less frequently used in contexts like Sub-Saharan Africa, where internet and Facebook penetration are low. We aim to assess the potential of this approach in such settings. Additionally, we explore the feasibility of a network reporting approach to estimate fertility rates.

METHODS & DATA We used Meta's advertising platform to recruit survey respondents, targeting Facebook users aged 18 and over in all 14 regions of Senegal. Data collection occurred over one week in October 2024. Our survey included a network reporting component, where respondents provided information about themselves and three people from their regular social network. This approach captures unique data typically inaccessible in standard Facebook surveys, including socio-demographic information such as age, gender, education level, and number and age of children. Our analysis aims to estimate birth rates, using data from the Demographic Health Survey (DHS) as a benchmark for sample composition and fertility rate accuracy.

RESULTS Our sample includes 350 respondents, with 44% women. About 24% live in Dakar, Senegal's capital, and 37% live in rural areas. The average respondent age is 33. On average, respondents reported contact with 10 people the previous day and provided detailed information for up to three. The network sample (n=567) is gender-balanced (50% women) with an average age of 30. About 23% of network members reportedly do not use Facebook. Further analysis will focus on fertility rate estimation, comparing our findings with DHS data to assess the reliability of our approach.

ADDED VALUE Our study addresses data gaps in African fertility estimates and introduces a new data collection method using social media. By comparing our results with DHS data, we aim to evaluate the potential of this approach for providing timely fertility estimates in African contexts, thereby enhancing understanding of population trends in the region.

SHOULD WE BE WORRIED? THE IMPACT OF PROBLEMATIC RE- SPONSES ON SOCIAL MEDIA SURVEYS

ZAZA ZINDEL German Centre for Integration and Migration Research (DeZIM); Bielefeld University, Germany

RELEVANCE & RESEARCH QUESTION The digital age has transformed survey recruitment, with social media ads enabling cost-effective and rapid access to diverse and hard-to-reach populations. Despite this promise, this method raises critical challenges - one of these is the tendency for measurement errors in form of problematic response behaviors. These behaviors—including satisficing, low-effort responses, and fraudulent participation—threaten data quality. While some studies have shortly mentioned these issues, most social media-recruited surveys do not address them systematically. This raises the question: Should we worry about problematic responses in social media-recruited surveys? This study examines whether problematic respondents in social media-recruited surveys systematically differ from others and assesses their impact on data quality and model estimates. The study addresses two core questions: [1] Are problematic respondents systematically different in socio-demographics and substantive answers? [2] Do they bias multivariate model estimates?

METHODS & DATA The study analyzes data from a web survey on labor market discrimination against women with headscarves in Germany, conducted in 2021/2024. Recruitment via targeted Facebook ads yielded 3,021 completed interviews at an average cost of €1.41 per respondent. Response quality was evaluated using indicators such as item non-response, straight-lining, speeding, and identity misrepresentation. Statistical tests, including chi-square, Fisher's exact, and Mann-Whitney U-tests, were employed to identify significant differences. Multivariate regression models assessed the impact of problematic behaviors on key outcomes, such as perceived anti-Muslim discrimination.

RESULTS Preliminary findings reveal that problematic respondents differ significantly in socio-demographic composition and substantive answers. For example, behaviors like straight-lining and speeding are more frequent among younger and less-educated respondents. Multivariate analyses show that problematic responses distort key estimates, particularly on discrimination experiences. Cleaning the dataset improves model fit and the reliability of results, emphasizing the value of robust quality control measures.

ADDED VALUE This study provides a systematic investigation of problematic response behaviors in social media-recruited surveys, shedding light on their prevalence, predictors, and implications for data quality. It underscores the necessity of incorporating quality control measures in future surveys, offering practical recommendations for researchers leveraging social media recruitment strategies.

OPTIMIZING SOCIAL MEDIA RE- CRUITMENT: BALANCING COSTS AND SAMPLE QUALITY IN NON- PROBABILISTIC PANELS

JESSICA DAIKELER, JOACHIM PIEPENBURG, BERND WEISS GESIS, Germany

RELEVANCE & RESEARCH QUESTION With social media platforms increasingly utilized for participant recruitment, it's critical to assess their effectiveness in building balanced non-probabilistic panels. This study investigates whether investing in targeted social media ads on platforms like Meta (Facebook and Instagram) can effectively balance recruitment costs and sample composition quality. We explore how different targeting criteria (such as age and gender) impact sample composition and evaluate the accuracy of platform-provided socio-demographic estimates. The study aims to understand the trade-offs between advertising budget allocation and sample representativeness, addressing the overarching question: do high-cost ad strategies improve recruitment outcomes?

METHODS & DATA Our data stems from the GESIS Panel Digital Behavioral Data Sample (GESIS Panel.dbd) recruitment campaign, conducted on Meta platforms in May 2023. We employed a structured ad campaign with four distinct targeting groups based on gender and age, allowing for comparison of recruitment efficacy across various demographics. The recruitment process included a click-through ad link leading to a welcome survey, followed by a registration step to join the panel. Incentives were provided to participants who completed registration. Data analyses focus on the socio-demographic composition,

recruitment costs, and response rates across ad groups, as well as the accuracy of Meta's demographic targeting information.

RESULTS Our findings reveal that targeted social media recruitment can enhance demographic balance in certain respects but does not fully eliminate sampling biases. For instance, Meta's age and gender targeting improved representation among older individuals but showed limitations for younger demographics. Additionally, while the socio-demographic estimates provided by Meta are generally reliable, slight misclassifications (around 5%) were observed. Cost analysis revealed that lower recruitment budgets yielded the most cost-effective samples, contradicting the notion that higher spending guarantees improved sample composition. Higher ad expenditures increased reach but also raised cost per participant, suggesting a strategic budget allocation is essential for optimal sample utilization.

Added Value allocation is essential for optimal sample utilization.

ADDED VALUE This study offers practical insights for researchers leveraging social media platforms in recruitment, particularly in non-probabilistic settings. By illuminating the cost-structure dynamics and demographic accuracy of Meta's ad-targeting tools, we provide guidelines for optimizing recruitment budgets without compromising sample quality.

WEDNESDAY, 02.04.2025
04:00 - 05:00 PM
HÖRSAAL A

12.1:
REDUCING
NONRESPONSE

SESSION CHAIR: ALEXANDRA ASIMOV
 GESIS, GERMANY

PRESENTATIONS

CAPI, OR NOT CAPI – THAT IS THE QUESTION: USING ADMINISTRATIVE DATA TO ASSIGN THE OPTIMAL MODE FOR MAXIMIZING RESPONSE-RATES IN A HOUSEHOLD PANEL

PATRICK LAZAREVIC, MARC PLATE Statistik Austria, Austria

RELEVANCE & RESEARCH QUESTION Selecting the appropriate mode(s) of data collection is a major consideration for every survey. Personal interviews (CAPI) – typically recognized as the gold standard of data

collection in surveys – have been increasingly called into question as the be-all and end-all of data collection methods, with a growing shift towards self-administered modes, particularly web surveys (CAWI). Reasons for this range from representation concerns due to shifting mode preferences and the flexibility these modes offer to busy respondents, all the way to practical constraints like health and safety concerns, interviewer availability, and budgetary restrictions. Yet, recruitment using CAWI alone might result in biases due to, e.g., systematic differences in digital skills. Thus, many surveys employ mixed-mode designs, raising the question of how to determine which mode should be offered to whom.

METHODS & DATA The Austrian Socio-Economic Panel (ASEP), a household panel of the Austrian population, experimentally tested a tailored mode-design using administrative data to assign half of the sample's households to their presumably preferred mode (CAPI/CAWI) while also offering the other mode after persistent non-response. The other households were randomly assigned to one of the mode-designs (CAPI-First/CAWI-First) as control groups. To evaluate the utility of the tailored mode-design concept, we employed a multi-faceted analytical approach by comparing a variety of indicators, such as the rate of proxy-interviews, the number of requested mode-changes, or the overall response-rates and resulting nonresponse bias, between the different mode-designs.

RESULTS Results were promising regarding the tailored mode-design. For example, response-rates were consistently higher compared to the other mode-designs, while proxy-rates and the nonresponse bias were considerably lower.

ADDED VALUE With the tailored mode-design, we present an interesting and promising alternative to already established single- or mixed-mode designs. This novel approach could prove helpful to decrease nonresponse bias and survey costs while maintaining data quality.

THE FRAMEWORK OF SURVEY BEHAVIOUR: AN EXTENSION OF THE FRAMEWORK FOR WEB SURVEY PARTICIPATION

JELDRIK BAKKER, JONAS KLINGWORT, VERA TOEPOEL Statistics Netherlands

RELEVANCE & RESEARCH QUESTION Why do people behave the way they do in surveys? The answer to this fundamental question in survey research can help increase survey participation, decrease break-off and improve data quality. Underneath this seemingly simple question is a complex interplay of factors influencing survey behaviour (i.e., the behaviour of [potential] respondents). While current frameworks, theories and models provide valuable insights into this behaviour, they all have limitations in understanding survey behaviour as a whole. Furthermore, none are generically applicable across survey behaviours for all modes, devices, and target populations (i.e., person, household, establishment).

METHODS & DATA We conducted an extensive literature review of both generic behavioural, and survey-specific frameworks, theories and models. Using the Framework for Web Survey Participation (Peytchev,

2009) as a starting point, we extended this framework into our generic Framework of Survey Behaviour.

RESULTS The resulting framework provides a holistic view of the factors affecting the key survey decisions and the underlying behaviours that shape those decisions. The key survey decisions reflect the three main goals in survey research: getting people to start the survey, complete the survey, and provide high-quality responses. These decisions are affected by five groups of factors: environmental factors, respondent factors, interviewer factors, survey design factors, and questionnaire factors. The underlying survey behaviours that shape those decisions are diverse and range from (proxy) responding, sacrificing, breaking off, and straightlining, to speeding.

ADDED VALUE By centralising behaviour in the framework, we offer a comprehensive approach that considers all human, organisational, and environmental elements involved in the survey process. The framework guides researchers in designing surveys and collecting high-quality data across diverse contexts. Understanding and being able to influence survey behaviour for the better is key in order to improve respondent engagement and data quality. Practical recommendations are provided, and future research areas are identified.

REFERENCES Peytchev, A. (2009). Survey breakoff. *Public Opinion Quarterly*, 73(1):74–97.

SURVEY DESIGN FEATURES THAT MATTER: A META-ANALYSIS USING OFFICIAL STATISTICS SURVEYS OF THE NETHERLANDS

JELDRIK BAKKER, JONAS KLINGWORT, VERA TOEPOEL Statistics Netherlands

RELEVANCE & RESEARCH QUESTION More and more surveys are being conducted, but response rates are declining. The solution to avoid the missing data problem is to avoid having any. Understanding which factors influence response rates is crucial for improving survey participation and reducing nonresponse bias. This study investigates which features significantly affect response rates and how they can be optimized to improve survey participation?

METHODS & DATA We conducted a multilevel meta-analysis using Statistics Netherlands' data from 38 person population surveys with over 1200 samples. The surveys were fielded over a seven-year period (2018–2024) and had a total sample size of over 7 million people. These surveys range from one-time to recurring studies with frequencies from weekly to biennially. We used 72 factors such as respondent factors (e.g., age, gender, nationality, device use), survey design factors (e.g., year, month, mode, incentive, fieldwork period, number of contacts, topic) and page & question factors (e.g., duration, number of block, pages & questions, number of question per question type, number of introduction texts) that potentially have an effect on the response rates.

RESULTS Preliminary findings suggest that the data collection mode, type of incentives, survey topic, the number of and types of questions, the device used by the respondent, and age and gender have

significant effects on the response rate. Interestingly, the length of the fieldwork, the number of reminders, and the periodicity of the survey show non-significant effects.

ADDED VALUE This study offers comprehensive insights into improving response rates for official statistics surveys, highlighting effective data collection strategies and identifying survey designs to avoid. Utilizing a wide range of different design features, the study serves as a practical toolbox for national statistical agencies, survey agencies, and survey researchers alike. Notably, it represents the first practical application of the Framework of Survey Behavior, which has also been submitted to this conference.

WEDNESDAY, 02.04.2025

04:00 - 05:00PM

HÖRSAAL B

12.2: APP-BASED DIARY STUDIES

SESSION CHAIR: FLORIAN HEINRITZ
LEIBNIZ INSTITUTE FOR EDUCATIONAL TRAJECTORIES,
GERMANY

PRESENTATIONS

THE EFFECT OF PERSONALIZED FEEDBACK ON PARTICIPANT BEHAVIOR IN DATA COLLECTION: USING PARADATA TO UNDERSTAND PARTICIPATION RATES AND PARTICIPANT ENGAGEMENT IN APP-BASED DATA COLLECTION

THIJS CORNELIS CARRIÈRE 1, BELLA STRUMINSKAYA 1, BARRY SCHOUTEN 1,2 1 Utrecht University, The Netherlands; 2 Statistics Netherlands

RELEVANCE & RESEARCH QUESTION Data collection with smartphone sensors and apps is increasingly used in substantive and market research. However, app studies suffer from low response rates and adherence. To increase participant engagement, some app-based studies provide personalized feedback or tailored insights that are generated based on the data participants provide and shown back to them. To date, studies on the effectiveness of personalized feedback on participation and adherence have been inconclusive. Given that implementing such feedback is costly, better insights are needed. We use app process data (paradata) to answer the research question:

‘What is the effect of personalized feedback on participation rates and participant engagement in an app-based general population study?’.

METHODS & DATA We use data collected using the Household Budget Survey app of Statistics Netherlands. A probability-based sample of 1441 households in the Netherlands was invited to download the app and report all their expenditures for two weeks. The app lets participants fill in their expenses manually or scan receipts.

Random half of the households received personalized feedback: graphs and figures summarizing all their expenses. We use paradata of the participants’ actions in the app to study whether participants that received personalized feedback differ in their app usage and study adherence compared to participants that did not receive feedback.

RESULTS From 1441 households, 290 participated in the study. We find no effect on participation rates and study engagement when promising feedback in the invitation stage and in providing participants with personalized feedback. We will present detailed findings of our analysis of paradata to demonstrate how participants interacted with the personalized feedback [e.g., number and duration of engagements, changes to reporting of expenses in the feedback and no-feedback groups], as well as and whether the feedback had an effect on specific subgroups.

ADDED VALUE The study advances our understanding of personalized feedback since paradata can show us how participants interact with personalized feedback. In light of our findings, researchers and practitioners can use personalized feedback as a tool to increase study engagement and/or make informed decisions on whether they should spend resources on research app feature of providing feedback.

A RECIPE TO HANDLE RECEIPTS? USABILITY-TESTING THE RECEIPT SCANNING FUNCTION OF AN APP-BASED HOUSEHOLD BUDGET DIARY

LASSE HÄUFLÖCKNER, JOHANNES VOLK Federal Statistical Office Germany (Destatis), Germany

RELEVANCE & RESEARCH QUESTION As part of the EU project Smart Survey Implementation (SSI), the Federal Statistical Office of Germany (Destatis) is participating in the development of Smart Surveys. Smart Surveys combine traditional question-based surveys and smart features, that collect data by accessing device sensor data. One smart feature is a receipt scanner, making use of the smartphone camera, allowing participants to upload pictures of shopping receipts in a survey app. The aim is to reduce respondent burden in diary-based household budget surveys. However, smart features can only help to simplify surveys if they can be easily used. Therefore, the usability of the receipt scanning function was tested with potential respondents.

METHODS & DATA Destatis conducted qualitative usability tests with a diary-based household budget app, featuring a prototype of a receipt scanner. 19 participants used the app, while their interaction with the app was observed, followed by an interview on their experiences.

RESULTS Given the choice between manual input and using the scanner, respondents prefer the scan function to record purchases. Participants appreciate the fast and easy way to record receipts, compared to the manual input of purchases.

Making use of the scan function does not raise privacy concerns – at least not as long as an advantage is clearly given and as the publisher of the app is known and trusted.

All participants were able to use the scan function, although user friendliness of the current state of development proved to be insufficient. As a prototype was tested, the scan results were prone to errors, which had to be corrected. Respondents do not accept having to correct data, as the effort involved is perceived as too high and results are expected to be accurate.

ADDED VALUE The study shows that a receipt scanning function per se is highly appreciated. However, in order to be used by the respondents, it is imperative that the function works perfectly, that its operation is easy to understand and involves little effort. Concerning the further development of this smart feature, the results confirm us in our approach but also show, where improvements are needed.

HOW MANY DIARY DAYS? SMART SURVEYS CAN HELP TO REDUCE BURDEN AND COSTS OF DATA COLLECTION FOR BEHAVIOURAL STATISTICS

**DANIELLE REMMERSWAAL 1,2, BARRY SCHOUTEN 1,2,
BELLA STRUMINSKAYA 1** 1 Utrecht University, Netherlands, The; 2 Statistics Netherlands

RELEVANCE & RESEARCH QUESTION Diary studies are used to capture detailed respondent behaviour, but they are highly burdensome for respondents, resulting in nonresponse and measurement-errors. Smart surveys (i.e using a smartphone or activity tracker) can make use of sensors that can replace questions, reducing the response burden. Researchers can then collect diary data for a prolonged time period.

We aim to answer the question “How many days should we collect data for a smart diary study in behavioural statistics?”, considering the response burden and the inter- and intra-personal variability of statistics.

METHODS & DATA We use data from four one-week smart diary studies: a physical activity study (2021, N = 414), a travel study (2018 and 2022, N = 185), and a budget study (2022, N = 10,421). All studies used probability-based samples, were conducted by Statistics Netherlands, and have different levels of response burden.

For each data source we calculate two statistics based on the frequency and duration of the studied behaviour. For each statistic we separate the inter-individual and intra-individual variance. We calculate the intraclass-correlation-coefficient (ICC), the proportion of intra-individual variance out of the total variance. We then calculate the reliability of a study period of 7 or fewer days, based on these variance components.

RESULTS The intra-individual variance for 7 days is high for each of the studies, meaning that not every day looks the same for studied individuals. The ICC is approximately 50% for travel statistics, 55% for physical activity statistics, and 90% for expenditure statistics. The reliability for the first two studies is around 0.80 at 7 days, meaning that 7 days provide enough information. While for expenditure, reliability is around 0.50 at 7 days, implying more days are useful.

ADDED VALUE Our results can guide the decision on the number of diary days by taking into consideration burden and data quality. The results can help practitioners efficiently allocate resources (e.g., lower sample size, shorter data collection), reducing the costs. Our conclusions and recommendations have relevance across a wide range of research areas concerning studies on (time-use) behaviour, and studies using (smartphone) sensors or apps.

of social identity by using phrasing such as “We need more responses from people like you”), (3) descriptive norm intervention (inducing a norm activation with a text like “Most people have already responded but we still need your answer”), and (4) a combined social identity intervention (2) and descriptive norm intervention (3).

RESULTS Data collection will occur in December 2024. The data will be analyzed by comparing response rates between the experiment groups, and non-response bias by analyzing age, gender and level of education of the respondents. The abstract will be updated with results early January 2025.

ADDED VALUE This study adds to the literature of reminders as a tool to increase response rates, while also not adding any cost. Further it adds to the research of theories of social identity and social persuasion by exploring if these psychological mechanisms can be utilized in the context of survey reminders.

WEDNESDAY, 02.04.2025
04:00 - 05:00 PM
HÖRSAAL C

12.3: REMINDERS AND TARGETING

PRESENTATIONS

I WANT YOU FOR THIS SURVEY!
TEXT INTERVENTIONS IN EMAIL
REMINDERS TO INCREASE RESPONSE
RATES IN WEB SURVEYS

FREJA WESSMAN, ANDERS CARLANDER University of Gothenburg, Sweden

RELEVANCE & RESEARCH QUESTION Declining response rates (Groves, 2006) is a potential threat for survey validity, e.g., non-response bias (Groves & Peytcheva, 2008). While reminders have been reported to increase the response-rate in web surveys (Sammot et al., 2021), few studies have investigated the effects of number of reminders, and even fewer have examined specific content (Pedersen & Nielsen, 2016). This study applies psychological theories of social identity (Tajfel & Turner, 1979) and social influence through descriptive norms (Goldstein et al., 2008), to test if text interventions in reminders can increase response rates and decrease non-response bias.

METHODS & DATA This is a 2x2 factorial experiment using the Swedish Citizen Panel, a non-commercial non-incentive web panel. Each group received a different intervention: (1) control group/no intervention (a general reminder), (2) social identity intervention (eliciting a sense

THE EFFECTS OF USING DIGITAL MAILBOX REMINDERS

**JULIA BERGQUIST, SEBASTIAN LUNDMARK, CORNELIA ANDERSSON,
KLARA PERSSON** The SOM Institute, University of Gothenburg, Sweden

RELEVANCE & RESEARCH QUESTION Text messages have often been used as a cost-effective way to remind sample persons to complete mailed paper-and-pencil and online questionnaires. However, in some countries, the likelihood of identifying phone numbers of sample persons have decreased rapidly.

Fortuitously, digital mailboxes have been implemented in Sweden where government authorities and companies can send mail digitally instead of physically. According to the Agency of Digital Government (2024), more than 6 million Swedes used these digital mailboxes in 2023. The present study assessed the effect that reminding potential respondents by digital mailboxes instead of through text messages.

METHODS & DATA The preregistered experiment was implemented in a self-administered mixed-mode survey (paper-and-pencil and web questionnaire) administered to a sample of 26,250 randomly selected individuals. Prior to being invited to complete the questionnaire, all sample persons were randomly assigned to one of two groups.

One group of sample persons received reminders via four regular mailings and four texts, whereas the other respondents received reminders through four regular mailings, two through texts, and two through their digital mailbox. Per the preregistration, the study will assess the two reminder strategies in terms of response rates, non-response bias, and data quality.

RESULTS The data collection period started in August 2024 and will finalize at the beginning of January 2025. Therefore, empirical results of the experiment will not be available until January 2025.
Added Value

CAN TARGETED APPEALS WIN OVER LATE RESPONDENTS IN A BUSINESS WEB SURVEYS?

ELLEN LAUPPER 1,2, ULF-DIETRICH REIPS 2 1Swiss Federal University for Vocational Education and Training SFUVET, Switzerland; 2 Konstanz University, Germany

RELEVANCE & RESEARCH QUESTION Nonresponse is a particularly salient issue in the context of business web surveys, as response rates for such surveys are on average lower than for personal web-surveys. Previous research in the field of business surveys has mainly focused on testing the utility of different contact and survey modes to increase response rates. The following study uses an experimental approach to test the implementation of an adapted and expanded design feature to increase motivation to complete the survey, originally proposed by Lewis et al. (2019). This is done by implementing a more interactive and targeted communication with late respondents, i.e. potential nonrespondents.

METHODS & DATA Therefore, a randomly selected sample of late respondents (N=3'000 in total) to a well-established Swiss business web survey was selected and asked to indicate their primary reason for nonresponse at various stages in the survey design. Subsequently, they were randomly presented with one of three appeals: a targeted appeal addressing their previously stated reason for nonresponse, a generic general appeal addressing some of the most prevalent reasons for nonresponse, or a main appeal providing arguments for the nonresponse reason of "not time". After reading the appeal respondents had to indicate whether they wanted to proceed immediately to the main survey or postpone the decision to participate.

RESULTS As hypothesized, a targeted appeal demonstrated a more pronounced motivational effect compared to the two alternative appeal types in persuading respondents to proceed to the main survey. However, the motivational effect was limited insofar as it did not extend to a higher motivation to provide more complete data. Furthermore, more detailed analyses revealed that the conversion rates are neither independent of the type of nonresponse reason nor timing factors.

ADDED VALUE The study demonstrates that the implementation of such an interactive feature is feasible and potentially worthwhile, even in the context of a challenging business survey. It shows the potential to motivate target respondents to participate, particularly at a later stage in the fieldwork period.

INVITED SESSIONS

TUESDAY, 01.04.2025
10:45 - 11:45 PM
MAX-KADE-AUDITORIUM

INVITED SESSION I:

SESSION BEI MARKT- FORSCHUNG.DE

SESSION CHAIR: HOLGER GEISSLER,
 MARKTFORSCHUNG.DE, GERMANY

PODIUMDISKUSSION: KI IN DER MARKT- UND MEINUNGSFORSCHUNG – WO SOLL DAS EIGENTLICH ALLES HINFÜHREN?

HOLGER GEISSLER Moderation, Geschäftsführer marktforschung.de & succet GmbH

DR. STEFFEN SCHMIDT Experte für Markenmanagement, Neuromarketing und Künstliche Intelligenz

HARTMUT SCHEFFLER Berater Marktforschung und Markenführung, vormals Geschäftsführer Kantar Deutschland

CHARLOTTE WEBER COO von Civey

RUTH WAKENHUT Head of Business Development KERNWERT GmbH

PODIUMSDISKUSSION Die aktuelle Diskussion um den Einsatz von KI in der Marktforschung ist auch zwei Jahre nach dem Launch von ChatGPT vor allem geprägt von der Frage, an welchen Stellen KI eingesetzt werden kann, um Prozesse zu vereinfachen oder zu automatisieren oder mittels KI neue Tools und Methoden zu entwickeln.

Was bislang kaum diskutiert wird, ist die Frage, welche Auswirkungen der verstärkte Einsatz von KI auf die Branche der Markt-, Sozial- und Meinungsforschung an sich haben könnte. Welche Veränderungen sind zu erwarten? Wie wird sich die Branche unter dem Einfluss von KI verändern?

In der Diskussionsrunde sollen folgende Aspekte thematisiert werden:

- Datenerhebung und Qualität
- Datenanalyse und Aufbereitung der Daten
- Markt (Reputation der Branche, Umsätze, Rentabilität, Bedeutung von Instituten)
- Mitarbeitende

PODIUMSTEILNEHMER

- Dr. Steffen Schmidt, Experte für Markenmanagement, Neuromarketing und Künstliche Intelligenz

- Hartmut Scheffler, Berater Marktforschung und Markenführung, vormals Geschäftsführer Kantar Deutschland
- Charlotte Weber, COO von Civey
- Ruth Wakenhut, Head of Business Development KERNWERT GmbH
- Prof. Dr. Dirk Lippold, Dozent, Autor, vormals Geschäftsführer Capgemini Deutschland GmbH

MODERATION Holger Geißler, Geschäftsführer marktforschung.de & succet GmbH

WEDNESDAY, 02.04.2025
09:00 - 10:00 AM
HÖRSAAL D

INVITED SESSION II:

INNOVATION IN PRACTICE: SMART SURVEY TECHNIQUES

SESSION CHAIR: STEFAN OGLESBY
 DATA IQ AG, SWITZERLAND

PRESENTATIONS

MEASURING THE IMPACT OF OOH ADVERTISING IN THE SWISS ALPS

BEAT FISCHER intervista AG, Switzerland

In 2024, intervista was commissioned by APGffISGA to investigate the advertising impact of campaigns in winter sports resorts in Switzerland. The innovative research design included two essential components. On the one hand, smartphone-based GPS tracking was used to continuously measure contacts with the advertising spaces and calculate campaign reach. On the other hand, a survey was conducted after the campaigns and the survey data was analyzed in combination with the measured contacts. This research approach made it possible to understand the advertising impact in relation to the contact dose. In addition to the methodology, the presentation will also show concrete results.

AI CHANGING THE INSIGHT GAME? A PRACTITIONER'S VIEW ON DEVELOPING AND IMPLEMENTING A RAG-BASED AUDIENCE SIMULATION AT SAMSUNG

FLORIAN BAUER Samsung Electronics Austria & Switzerland, Switzerland

GOALSETTING The primary goal of this project was to explore the broader use of Generative AI (GenAI) in consumer insights, enhancing accessibility, efficiency, and marketing decision-making. In preparation for Samsung's Galaxy Ring launch, we developed an AI model with two key purposes:

1. acting as a subject matter expert on the smart ring category, capable of answering business questions, simulating qualitative research, and predicting consumer behavior.
2. generating virtual personas to simulate potential buyers, supporting strategy and GTM activities.

As the project evolved, it became clear this initiative could reshape collaboration between consumer insights and marketing.

PROCESS Developed iteratively with a market research agency, the model was trained on proprietary research, stakeholder interviews, global marketing inputs, and social listening data. Validation tested its ability to replay ingested data, simulate buyer personas, and predict selected variables using holdout survey data. After successfully simulating responses on a nationally representative level, further validation focused on key buyer personas.

Samsung and the agency prioritized reliability and scalability, ensuring the model was ready for practical application. It was officially introduced to the Samsung Marketing team in November 2024.

RESULTS The model showed strong qualitative and quantitative validation ($R^2 > 0.9$ in most cases). Notably, it flagged an error in reporting where the holdout dataset was used, reinforcing its reliability. Confidence in its accuracy led to formal adoption within the marketing team. Ongoing usage is monitored, and findings will be shared at GOR 2025, alongside planned extensions like pre-testing ad assets.

ADDED VALUE This session provides a practitioner's perspective on developing an AI-driven, RAG-based model for marketing decision-making. It contributes to best practices in GenAI by reflecting on development, validation, and implementation. Additionally, it explores how AI-driven insights can transform marketing and consumer insights collaboration, fostering more data-informed, impactful decisions.

"IMPLICIT CONJOINT" – WHY LATENCY TIME IS IMPORTANT

PHILIPP FESSLER YouGov Schweiz AG, Switzerland

PETER KURZ bms marketing research + strategy

PIERPAOLO PRIMOCERI YouGov Schweiz AG, Switzerland

RELEVANCE & RESEARCH QUESTION Dual process theories suggest that the majority of human decisions are made by System 1, which works automatically and largely unconsciously. This system leads to highly habitualised behaviour in which consumption decisions are made relative to other options, regardless of the product category. Therefore, an optimal research design should not only create realistic decision contexts and be based on established experimental methods but also consider consumers' implicit decision-making processes. We will present an implicit Discrete Choice Model (DCM) that is suitable for modelling System 1-influenced decisions.

METHODS & DATA To this end, we had participants complete a DCM and, subsequently, evaluate the concepts chosen in the DCM tasks (i.e., winner concepts) by means of a Single Implicit Association Test (SIAT). This allowed us to specifically measure the respondents' response speed when indicating whether the concept represents a purchase option or not through a number of cognitive, affective and behavioural statements towards the respective concept. The SIAT is the most commonly used method for this purpose.

Participants' responses and response times measured in the SIAT were finally transformed into a linear function of associative strength and used to calibrate the utility estimation of the DCM. By incorporating subconscious associations into the choice model, this approach enhances predictive accuracy, bridging the gap between stated preferences and real-world purchasing behaviour.

RESULTS Two online studies – one on "city trips" and one on "smartphones" – demonstrate the impact of integrating SIAT into DCM. Simulations reveal that incorporating measured response times enhances the predictive accuracy of purchasing behaviour, capturing decision-making dynamics beyond traditional models. Since DCM is typically influenced by System 2 processes, this integration allows for the inclusion of more intuitive, implicit preferences. The results show that combining SIAT with DCM provides deeper, more meaningful insights into consumer decision-making, bridging the gap between stated and subconscious preferences.

ADDED VALUE Our approach enhances DCM by integrating SIAT, capturing subconscious biases often missed in traditional models and improving predictive accuracy.

WEDNESDAY, 02.04.2025
12:00 - 01:15 AM
HÖRSAAL D

INVITED SESSION III:
DGOF KI (AI) FORUM:
INSPIRATION SESSION (HELD IN ENGLISH)

OPPORTUNITIES & CHALLENGES IN APPLYING AI TO MARKET RE- SEARCH

SESSION CHAIR: YANNICK RIEDERJANSSEN EMEA, GERMANY
 AND OLIVER TABINO, 0 AGENTUR FÜR FORSCHUNG, GERMANY

Join us for an insightful session where experts test innovative solutions and explore the opportunities AI can bring to market research. Hear from industry experts about the current challenges in applying AI to market research and how they successfully overcome them.

PRESENTATIONS

GOR 2025 – AI INTERVIEW
TECHNOLOGY: ENHANCING QUALITY
DATA SETS WITH AI TOOLS IN THE
DIGITAL INTERVIEW PROCESS

HOLGER LÜTTERS HTW Berlin
MALTE FREKSA, LEONARD MACH horizoom GmbH Germany
JANNIS MARLAFEKAS, THOMAS OHDE Susi&James
DAVID RANFTLER Xelper

AI is about to transform industries, including market research. AI-assisted tools are enhancing both qualitative and quantitative data collection methods.

1. OBJECTIVES As primary researchers we launched an experiment for the usage of AI in the market research process by using 3 new technologies from three different tech partners to potentially replace the classic open ended question format. Feasibility and opportunities for AI integration in both quantitative and qualitative research approaches in questionnaires.

2. METHOD & APPROACHES Four approaches of Human Computer Interaction are compared with the topic of travel asking 12 identical questions in four different tech settings. For each of the technical settings a minimum of 50 responses was collected in the horizoom Panel. The four different modes of interviewing were:

3. CLASSIC TYPE-IN QUESTIONNAIRE with 12 open answers in a standard online questionnaire format

4. VOICE IN QUESTIONNAIRE (Questions in oral mode using AI transcriptions) HTW Berlin

5. AI CHATBOT INTERVIEW (Dialogue while typing) Xelper

6. AI ASSISTED TELEPHONE INTERVIEW (AICati: Telephone Call with AI Bot) Susi&James

7. RESULTS Once the different setups are up and running they all produce insightful responses with enriched data as respondents compared to the classic open ended question. The presentation will reveal the differences in the approaches.

8. IMPACT With the help of AI, we aim to support different MR tasks, augmenting the input from a user's perspective. The technologies used could potentially reduce costs and bias while enabling bigger data collections for both qualitative and quantitative research approaches which can be easily enriched with the different modes. One industry impact may also include the technology change by using non-MR technologies to perform MR tasks.

9. LIMITATIONS The different modes produce different interaction profiles which seem to be more related to the person answering than to the technology used. In future studies the interaction mode preference of respondents might be integrated into the market research process to ensure a fit with a respondent's preferred interview mode. The comparative studies focused on the AI supported data collection. AI could was also being used for the analytical parts of all four input types. The comparative studies focused on the AI supported data collection. AI could was also being used for the analytical parts of all four input types.

WHY DO LLMs ENTER MARKET RESEARCH IN COMPANIES SO SLOWLY

THOMAS LAUFEN AG, Germany

CURRENT SITUATION With the advent of LLMs, a significant impact on various industrial activities has been anticipated. However, their current application in market research within companies remains limited. The challenges associated with this approach lie beyond the capabilities of the method itself.

PROBLEM LLM-based methods compete with established techniques and existing data sources.

FUTURE The limited adoption of LLMs is not due to deficiencies in the method itself but rather to competition with existing data sources and analytical approaches. The use of LLMs is likely to grow when they enable answering questions that remain unresolved with current analytical approaches.

WEDNESDAY, 02.04.2025
02:30 - 03:45 AM
HÖRSAAL D

INVITED SESSION IV:
DGOF KI (AI) FORUM:
WORLD CAFÉ (SESSION HELD IN GERMAN)

OPPORTUNITIES & CHALLENGES IN APPLYING AI TO MARKET RE- SEARCH

SESSION CHAIR: GEORG WITTENBURG, INSPIRIENT AND
YANNICK RIEDERJANSSEN EMEA AND OLIVER TABINO,
AGENTUR FÜR FORSCHUNG – GERMANY

Engage in World Café discussions with fellow GOR participants. Discuss your personal challenges and barriers of integrating AI into market research projects and collaborate on strategies to overcome them. Share experiences and gain new insights on successfully implementing AI in market research projects.