26[™] GENERAL ONLINE RESEARCH CONFERENCE 21 - 23 FEBRUARY 2024 IN COLOGNE

SHAPING ONLINE RESEARCH



SIMON KÜHNE, BELLA STRUMINSKAYA, CARINA CORNESSE, OTTO HELLWIG, Stefan Oglesby, yannick Rieder (Eds.)

26TH GENERAL ONLINE RESEARCH CONFERENCE PROCEEDINGS COLOGNE 2024

 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 978-3-9822985-5-9

SIMON KÜHNE, BELLA STRUMINSKAYA, 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-Lena Kemper / LAYOUT AND TYPESET: Nicole Propach, Köln

TABLE OF CONTENTS 03



ORGANIZATION 04

DGOF BOARD

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

> **DR. CARINA CORNESSE** DIW Berlin, Germany

DR. STEFAN OGLESBY data IQ AG, Switzerland

PROF. DR. BELLA STRUMINSKAYA Utrecht University, The Netherlands

> YANNICK RIEDER Janssen-Cilag GmbH, Germany

PROF. DR. SIMON KÜHNE Bielefeld University, Germany

PROGRAMME COMMITTEE

PROF. DR. SIMON KÜHNE (Programme Chair), DGOF Board & Bielefeld University PROF. DR. BELLA STRUMINSKAYA

(Vice Programme Chair), DGOF Board & Utrecht University **DR. OTTO HELLWIG**

(Track A), DGOF Board, Bilendi & respondi, Germany DR. CARINA CORNESSE (Track A), DGOF Board & DIW Berlin DR. STEFAN OGLESBY (Track B), DGOF Board & Data IQ AG

PROF. DR. FLORIAN KEUSCH (Track B), University of Mannheim

CONFERENCE CHAIR

DR. OTTO HELLWIG Chairman DGOF Board, Bilendi & respondi, Germany DR. FELIX GAISBAUER (Track C), Weizenbaum-Institut e.V. DR. JOSEF HARTMANN

(TRACK C), Verian **YANNICK RIEDER** (GOR Best Practice Award 2024 Competition),

DGOF Board & Janssen-Cilag GmbH DR. OLAF WENZEL (GOR Thesis Award 2024 Competition), Wenzel Marktforschung DR. FREDERIK FUNKE

(GOR Workshops), datenmethoden.de & LimeSurvey GmbH

DGOF OFFICE

LOCAL PARTNER

Rheinische Fachhochschule Köln gGmbH

INTERNATIONAL BOARD 2024

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

DEAR GUESTS OF THE GOR 24!

A FAMILIAR CITY – A NEW PLACE

WELCOME TO THE 26TH EDITION OF THE GENERAL ONLINE RESEARCH CONFERENCE. WE ARE VERY HAPPY TO HOST THIS YEAR'S GOR IN COOPERATION WITH THE RHEINISCHE FACHHOCHSCHULE KÖLN.

This year marks the sixth time that the GOR has visited its founding city of Cologne. In 2019, we last met in Cologne's Südstadt district at our cooperation partner at the time, TH Köln. This year, the Rheinische Fachhochschule is opening its doors to us as a new partner of the GOR at its Campus Centre in Cologne-Ehrenfeld.

The former working-class and industrial district Ehrenfeld now has a very lively cultural scene and is becoming increasingly popular as a residential area - a growing number of theatres, restaurants and trendy pubs bear witness to this. 230 kilometres from the sea and within sight of the GOR, Ehrenfeld has one of three inland lighthouses in Germany. The Heliosturm was built at the end of the 19th century for test and demonstration purposes by the electrical engineering company Helios. The GOR is also sometimes a test and demonstration facility for online research and its latest varieties, a kind of lighthouse in the research landscape. Fits!

As in previous years, we have a great conference programme lined up for you including keynotes, discussions, presentations, awards, posters, workshops and much more. You can choose between four simultaneous conference tracks: Track A covers "Survey Research: Advancements in Online and Mobile Web Surveys". Track B deals with "Data Science: From Big Data to Smart Data". Track C features "Politics, Public Opinion and Communication" and Track D covers "Digital Methods in Applied Research".

In addition, we have two award competitions: The GOR Thesis Award 2024 for the best thesis in online research, and the GOR Poster Award 2024 for the best poster of the conference. The DGOF Best Paper Award 2024 for the best paper in online research will also be awarded at GOR.

Our keynotes this year take different looks at the development of our profession. On Thursday, 22 February, Dr. Martin Andree, Founder AMP Digital Ventures will talk about "Digital monopolies: How Big Tech stole the internet - and how we can reclaim it", and on Friday, 23 February, Prof. Annette Jäckle, professor of survey methodology at the University of Essex will talk about Data collection using mobile apps and how to increase participation. On the pre-conference day, Wednesday 21 February, five workshops will take place and



on Wednesday evening we will meet for the GOR 24 Get Together with drinks and snacks. Thursday evening is GOR party time. The GOR Get-Together will take place at "Franky's Bar" (Venloer Str. 403, 50825 Köln). The "Mach Bar" (Zülpicher Straße 40, 50674 Köln) opens its doors for the GOR Party on Thursday at 8pm.

We are particularly grateful for the enthusiastic support of and collaboration with our partners at Rheinische Fachhochschule Köln: Prof. Dr. Christian Bosau and his team. We would also like to thank our sponsors and media partners. And, of course, a big THANKS to you, the conference participants, presenters, and speakers at this event!

HAVE A GREAT TIME AT THE GENERAL ONLINE RESEARCH CONFERENCE 2024! DR. OTTO HELLWIG ABOUT DGOF 07

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: WWW.DGOF.DE

1 DGOF MEANS DEVELOPMENT:

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

2 DGOF CONNECTS:

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

3 DGOF IS DIVERSE:

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

4 DGOF IS INNOVATIVE:

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

5 DGOF IS DISRUPTIVE:

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



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

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

DR. OTTO HELLWIG



OTTO HELLWIG WAS THE CEO OF RESPONDI AG FROM ITS FOUNDATION IN 2005 UNTIL THE SALE OF THE RESPONDI GROUP TO BILENDI IN 2021.

Since then he has been Corporate Integration Director of Bilendi & respondi. He has been active in the field of market and social research since the early 1990s.

Dr Hellwig has a degree in social sciences, psychology and media studies. He worked for several years as a research assistant at the Institute for Applied Social Research at the University of Cologne and received his doctorate in 2000.

OTTO HELLWIG HAS BEEN CHAIRMAN OF THE DGOF BOARD SINCE MARCH 2013.

DR. CARINA CORNESSE



CARINA CORNESSE IS HEAD OF THE GERMAN SOCIO-ECONOMIC PANEL INNOVATION SAMPLE (SOEP-IS) AND SOCIAL COHESION PANEL (SCP) AT THE GERMAN INSTITUTE FOR ECONOMIC RESEARCH (DIW BERLIN).

She holds a doctoral degree in Sociology from the University of Mannheim as well as research affiliations with the University of Bremen and the 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.

PROF. DR. SIMON KÜHNE



SIMON KÜHNE IS PROFESSOR OF APPLIED SOCIAL DATA SCIENCE AT THE FACULTY OF SOCIOLOGY AT BIELEFELD UNIVERSITY.

His methodological research is in the areas of Survey Methodology and Computational Social Science, esp. social media. He also works on aspects of social inequality in the areas of sexual and gender diversity, racism and discrimination, and health.

Simon Kühne has been actively involved in shaping the GOR for several years. After serving as a reviewer and track chair, he has shared responsibility for the program as Vice Program Chair since 2022.

HE HAS BEEN A MEMBER OF THE DGOF BOARD Since 2023 and he is the programme chair of the Gor 24 conference.

DR. STEFAN OGLESBY



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

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

HE IS A BOARD MEMBER OF DGOF SINCE 2019.

PROF. DR. Bella struminskaya



BELLA STRUMINSKAYA IS AN ASSISTANT PROFESSOR IN METHODS AND STATISTICS AT UTRECHT UNIVERSITY IN THE NETHERLANDS AND AN AFFILIATED RESEARCHER AT STATISTICS NETHERLANDS.

She holds a doctoral degree in Survey Methodology from Utrecht University. Her research focuses on innovations in data collection methods, data quality in studies that use smartphone apps, sensors, and wearables, and combining survey data with other data sources. She has published on various aspects of data quality in sensor and survey data, nonresponse and measurement error, panel conditioning, and device effects.

BELLA STRUMINSKAYA HAS BEEN A BOARD MEMBER OF THE GERMAN SOCIETY FOR ONLINE RESEARCH (DGOF) SINCE 2017 AND IS THE VICE PROGRAMME CHAIR OF THE GOR 24 CONFERENCE.

YANNICK RIEDER



YANNICK RIEDER IS MARKET RESEARCH MANAGER AT JANSSEN, THE PHARMACEUTICAL DIVISION OF JOHNSON & JOHNSON, WHERE HE IS RESPONSIBLE FOR INTERNATIONAL STUDIES.

His passion is the implementation of innovative approaches to generate in-depth customer and market insights. He has been successfully implementing this for over 10 years, both on the agency and corporate side.

YANNICK IS A DGOF BOARD MEMBER SINCE 2023.

DEAR PARTICIPANTS OF THE GOR 24,

HAVING BEEN A MEMBER MYSELF OF THE DGOF FOR A LONG TIME AND HAVING BEEN PARTICIPATING IN ALMOST ALL OF THE LAST GOR CONFERENCES AS A RESEARCHER, IT IS A GREAT PLEASURE FOR ME TO BE THE LOCAL PARTNER FOR THE DGOF AND THE GOR CONFERENCE THIS YEAR.



I would like to welcome you to the General Online Research Conference (GOR) at the lovely and vibrant campus of the RFH (Rheinische Fachhochschule Köln gGmbH). Besides many other reasons for visiting the lovely City of Cologne, the GOR conference is now an even better reason to come to Cologne and to make a visit to the RFH, who is one of the major educational providers in Germany, offering a modern curriculum in teaching, research, and services. With a tradition spanning over 60 years, the RFH, deeply rooted in the Rhineland region, is one of the oldest universities of applied sciences in Germany.

RFH, as your gracious host, values not only excellent research but also a warm and open sense of community. According to our guiding principles, the RFH is committed to providing high-quality, practice-oriented education and places significant importance on supporting students in their personal and professional growth. Similarly, the RFH strives for close collaboration with the business sector and other relevant stakeholders to ensure that its programs are always tailored to the demands of the current job market.

Could there be any better fit to the central idea of the GOR, who likewise emphasizes the interdisciplinary nature of online research, the connection of researchers and practitioners and the exchange of new knowledge in a vibrant and friendly community?

The GOR Conference at RFH is not just an opportunity to listen to and present cutting-edge research; it's much more a celebration of collective learning, exchanging ideas, and building connections.

I encourage you not only to actively participate in the presentations and workshops during the conference but also to seize the opportunity during breaks and evening events to engage in relaxed conversations. Often, it's the unplanned discussions that spark the most beautiful ideas and collaborations.

In the upcoming days, you can expect not only a rich array of high caliber talks and workshops but also the chance to meet like-minded individuals, forge new friendships, and deepen existing connections. Human interaction, personal engagement, and the ability to meet face-to-face are treasures that should be cherished.

I cordially invite you to view these days not just as a conference but as an experience – a time when we not only share knowledge but also laugh, discuss, and inspire each other. I am confident that the value of this shared experience will extend beyond the conference's borders, creating lasting connections among the participants. Cologne as a city and RFH as your host provide the perfect backdrop for that.

Thank you for being part of this unique community, and I look forward to experiencing unforgettable days together at the GOR Conference at RFH.

WARM REGARDS,

PROF. DR. CHRISTIAN BOSAU

SPONSORS AND ORGANIZERS 11



21 / N 2 / 2 N 2 A

09:00 AM - 10:00 AM	BEGIN CHECK-IN		WORKSHOP 4:
10:00 AM - 01:00 PM	WORKSHOP 1: EMBRACING DIVERSITY: INTEGRATING QUEER PERSPECTIVES IN ONLINE SURVEY RESEARCH Instructors: Zaza Zindel, Lisa de Vries (Bielefeld University, Germany)		WHY DATA SCIENCE AND OPEN SCIENCE ARE KEY TO BUILD SMART BIG DATA: AN EXAMPLE BASED ON A DECADE RESEARCH ON HARD-TO-REACH POPULATION IN TAIWAN Instructor: Ji-Ping Lin (Academia Sinica)
01-00 PM - 01-20 PM	WORKSHOP 2: PROBABILITY THEORY IN SURVEY METHODS Instructor: Blanka Szeitl (HUN-REN, Hungary)		WORKSHOP 5: FLEXIBLE TEXT CATEGORISATION IN PRACTICE: USING AI MODELS TO ANALYSE OPEN-ENDED SURVEY RESPONSES Instructor: Ludger Keeting (Tivian)
01.001M-01.001M			
01:30 PM - 04:30 PM	WORKSHOP 3 – ONLINE: THE MAGNIFICENT SEVEN: IDENTIFY AND REDUCE COMMON DATA QUALITY ISSUES IN ONLINE SURVEYS Instructor: Raffael Meier (onlineumfragen.com; Switzerland)	04:30 PM - 05:00 PM 05:00 PM - 06:30 PM 07:00 PM - 08:00 PM 08:00 PM - 11:00 PM	BREAK DGOF MEMBERS GENERAL MEETING Early Career speed networking event Gor 24 get together



08:00 AM - 09:00 AM	BEGIN CHECK-IN	01:15 PM - 02:30 PM	LUNCH BREAK
10:00 AM - 10:15 AM	GOR 24 OPENING & KEYNOTE 1:	02:30 PM - 03:30 PM	POSTER SESSION
	DIGITAL MONOPOLIES: HOW BIG TECH STOLE THE INTERNET – AND HOW WE CAN RECLAIM IT	03:30 PM - 03:45 PM	BREAK
	Keynote Speaker:	03:45 PM - 04:45 PM	A3.1: Solutions for Survey Nonresponse /
	Martin Andree (AMP Digital Ventures, Germany)		sponsored by GESIS
10:15 AM - 10:45 AM	BREAK		A3.2: Survey Instruments / sponsored by GESIS B3: The Power of Social Media Data
10:45 AM - 11:45 AM	A1: Survey Methods Interventions 1 /		C3: Artificial Intelligence
	sponsored by GESIS		D3: Kl in der qualitativen Marktforschung
	B1: Large, Larger, LLMs	04·45 PM - 05·00 PM	BREAK
	C1: Media Consumption	04.43114-03.00114	DILAN
	D1: Best Practice Cases	05:00 PM - 06:00 PM	A4.1: Innovation in Interviewing & Coding /
	T1: GOR Thesis Award 2024 Competition 1 /		sponsored by GESIS
	sponsored by TIVIAN		A4.2: Data Quality Assessments 1 / sponsored by GESIS
11:45 AM - 12:00 PM	BREAK		B4: Willingness to Participate in
			Passive Data Collection Studies
10:45 AM - 11:45 AM	A2: Mixing Survey Modes / sponsored by GESIS		C4: Political Communication and Social Media
	B2: Al Tools for Survey Research 1		D4: Wissenschaft trifft Praxis.
	C2: Online Research, Attitudes,		Wann ist eine Onine-Stichprobe gut für welchen Bedarf
	Preferences, Behavior	08:00 PM - 11:00 PM	GOR 24 PARTY
	D2: Innovation in Practice:		
	LLMs and more		
	T2: GOR Thesis Award 2024 Competition:		

PhD / sponsored by TIVIAN

22/**02**/**2021**

09:30 AM - 10:00 AM	BEGIN CHECK-IN	12:45 PM - 02:00 PM	LUNCH BREAK
10:00 AM - 10:4 5AM 10:45 AM - 11:15 AM	KEYNOTE 2: DATA COLLECTION USING MOBILE APPS: WHAT CAN WE DO TO INCREASE PARTICIPATION? Keynote Speaker: Annette Jäckle (University of Essex, United Kingdom) GOR AWARD CEREMONY	02:00 PM - 03:00 PM	A6.1: Questionnaire Design Choices / sponsored by GESIS A6.2: Data Quality Assessments 2 / sponsored by GESIS B6.1: Automatic Analysis of Answers to Open-Ended Questions in Surveys B6.2: Al Tools for Survey Research 2 D6: KlEngum, KLCafé
11:45 AM - 12:45 PM	A5.1: Recruiting Survey Participants / sponsored by GESIS A5.1: Detecting Undesirable Response Behavior / sponsored by GESIS B5: To Trace or to Donate, That's the Question C5: Politics, Media, Trust D5: KI Forum: Impuls-Session – Chancen und Regulierungen	<mark>03:00 PM - 03.15</mark> PM 03.15 PM - 04:15 PM	LUNCH BREAK A7.1: Survey Methods Interventions 2 / sponsored by GESIS A7.2: Social Media Recruited Surveys / sponsored by GESIS B7: Mobile Apps and Sensors

WEDNESDAY 21/02/2024 10.00 AM - 01.00 PM

EMBRACING DIVERSITY: INTEGRATING QUEER PERSPECTIVES IN ONLINE SURVEY RESEARCH

Instructors: Zaza Zindel, Lisa de Vries (Bielefeld University, Germany) Workshop language: English

Description of the content of the workshop:

Political and social advancements have enhanced the acceptance and visibility of sexual and gender minorities in many Western countries. However, the ongoing challenge of accurately addressing their unique experience in online survey research remains. Researchers and survey providers often struggle to incorporate queer perspectives, leaving many surveys and research designs blind to these minority groups.

This workshop offers a comprehensive introduction to the integration of sexual and gender diversity within (online) survey research. It focuses on four key areas:

- Measurement of Sexual Orientation and Gender Identity: Exploring nuanced approaches for respectful and inclusive data collection on sexual orientation and gender identity.
- Integrating Queer Perspectives: Learning effective strategies to craft survey questions that resonate with and capture the experiences of sexual and gender minorities.
- 3. Sampling Methods:

Gaining insights into strategies and techniques for effectively reaching and engaging sexual and gender minority populations in online survey re-search.

4. Data Preparation and Analysis:

Equipping participants with the skills to sensitively manage and analyze data collected from diverse populations to draw valuable in-sights.

This dynamic workshop combines informative presentations, group discussions, and hands-on exercises, ensuring participants leaving with the confidence and skills to successfully integrate sexual and gender diversity into their research projects.

About the instructors:

LISA DE VRIES is a research associate in the Quantitative Methods of Empirical Social Research department at Bielefeld University. In her dissertation, she focused on the effects of discrimination on leadership positions and job preferences of sexual and gender minorities. Her research interests incluce discrimination and harassment, LGBTQI* parent families and the measurement of sexual orientation and gender identity (SOGI). In addition, she has several experiences with sampling LGBTQI*-people in probability and non-probability surveys as well as measuring SOGI.

ZAZA ZINDEL is a research associate and a doctoral candidate in sociology at Bielefeld University, specializing in survey research. Her dissertation centers around the use of social media as a means of recruiting rare populations for web surveys. Her research interests include survey methodology, the potential of social media for empirical social research, and exploring new technologies to improve statistical representation of marginalized, vulnerable, or rare populations.

WEDNESDAY 21/02/2024 10.00 AM - 01.00 PM

PROBABILITY THEORY IN SURVEY METHODS

Instructor: Blanka Szeitl (HUN-REN, Hungary) Workshop language: English

Description of the content of the workshop:

The workshop focuses on the role of probability theory in the development and assessment of new techniques for data collection in survey research. As survey research advances more quickly and vigorously, the importance of understanding the fundamentals of probability theory is becoming more evident. We will emphasize the fundamentals of probability theory and explore how it relates to the evolution of survey research in the academic and business worlds. The topics will be presented with examples, simulations, brief proofs, and calculations that are relevant to the application and evaluation of innovative methods.

Goals of the workshop:

Course objectives are:

- To explore the connection between probability theory and the application and evaluation of new methods in survey research;
- to provide a systematic overview of the use of mathematical tools in mixed-method surveys and;
- 3. to demonstrate the relevance of probability theory to this field.

About the instructor:

BLANKA SZEITL is a survey methodologist and PhD candidate in applied mathematics. She is lecturer at Department of Statistics at University of Eötvös Lorand and at Bolyai Institute of Mathematics at University of Szeged. She is the head of Survey Methods Room Budapest research group focusing on innovative sampling procedures and data correction methods. She is researcher at HUN-REN Centre for Social Sciences, where she analyzes data of the European Social Survey (ESS). She is a member of the ESS sampling and Weighting Expert Panel working on the implementation of the ESS sampling strategy for countries participating in the ESS data collection. She is co-founder of Panelstory Opinion Polls, which is the first mixed-method probability panel in Hungary. Her research interests are survey sampling, innovative methods, probability theory and assessment procedures. She loves to read about the history of probability and statistics.

WEDNESDAY 21/02/2024 01:30 PM - 04:30 PM

ONLINE WORKSHOP: THE MAGNIFICENT SEVEN: IDENTIFY AND REDUCE COMMON DATA QUALITY ISSUES IN ONLINE SURVEYS

Instructor: **Raffael Meier** (onlineumfragen.com GmbH & Pädagogische Hochschule Schwyz, Switzerland) Workshop language: **English**

Description of the content of the workshop:

Survey data is contaminated with nonsubstantial variations from countless sources such as response styles, socially desirable responses and (partially) falsified interviews. We discover several of these causes, some originating in the design of the questionnaire (before the field phase) and others due to the individual behaviour of participants (during the field phase). The magical number seven matches both important sources of contamination due to effects in the questionnaire and important sources of contamination caused by respondents during participation in surveys. The aim is therefore to find variations in the observed responses that are not related to actual differences in participants' opinions.

In the workshop, we will look together at various possible contaminations, analyse data using concrete examples, discuss how such data can be prevented or cleaned up and what recommendations can be derived for future survey conducting. The workshop will include theoretical and practical inputs, as well as numerous activities with concrete data. Questions and discussions as well as examples from your own experience and data practice are welcome.

Phenomena include speeders, outliers, acquiescence, extreme response styles, mid-point responding and stereotypical responses, patterns, misunderstanding of questions due to poor item construction, and falsified and partially falsified interviews by interviewers and staff of survey research organisations, different fieldwork standards, different question formats for the same topic across several surveys, different content-related or methodological contexts, e.g. in crossnational surveys or multi-year surveys, heterogeneous interpretation of the questions due to cultural differences.

About the instructor:

RAFFAEL MEIER is co-founder and CTO of onlineumfragen.com and works as a researcher for the professorship "Media and Computer Science Didactics" at the Institut für Medien und Schule (IMS) at University of Teacher Education Schwyz. With onlineumfragen.com, as a specialised online survey partner and survey consultant for the DACH region, he supports selected clients and manages the development of the survey platform. With an interdisciplinary background in psychology, educational science, computer science and methodology, he works entrepreneurially and scientifically on data quality in online surveys and privacy & data protection behaviour in children. He held several workshops on the topic of data quality in online surveys at the Lucerne University of Applied Sciences and Arts and for the Swiss market research association Swiss Insights.

WEDNESDAY 21/02/2024 01:30 PM - 04:30 PM

WHY DATA SCIENCE AND OPEN SCIENCE ARE KEY TO BUILD SMART BIG DATA: AN EXAMPLE BASED ON A DECADE RESEARCH ON HARD-TO-REACH POPULATION IN TAIWAN

Instructor: **Ji-Ping Lin** (Academia Sinica; Taiwan) Workshop language: **English**

Description of the content of the workshop:

The emerging availability of big data in the past decade has overcome traditional constraints in research, especially in the discipline of humanities & social sciences. Increasing availability of big data is changing our world and transforming conventional thoughts about decision-making. Data science aims to cope with issues of big data. By definition, it consists of three disciplines, i.e., hacking skills, advanced mathematics and statistics, and domain knowledge. Taking full advantage of big data requires not only knowledge about fundamentals of data science, but also the ability of implementation. Big data do not offer us enough insight and vision. We need to go further to build smart data through the process of enriching and integrating the quantity and quality from different sources of big data. In the meanwhile, open data and open science have emerged simultaneously in the past decade in light of growing calls for the need to examine research reproducibility.

This workshop aims at

- addressing how open data and smart data sets are built by integrating hacking skills, advanced math/statistics methods, and domain knowledge of various disciplines on the basis of data science and open science
- 2. the role of online open data repositories in promoting crowd collaboration.

In the three disciplines of data science, the workshop focuses solely on how hacking skills and advanced math/stat are applied to build big data and smart data In the context of extracting valuable information embedded in source individual data, enriching the extracted information through the processes of cleaning, cleansing, crunching, reorganizing, and reshaping the source data. The data enrichment processes produce a number of data sets that contain no individual information but retain most of the source data information. The enriched data sets thus can be open to the public as open data.

Because the corresponding domain knowledge about hard-to-reach population research and Taiwan Indigenous Peoples (TIPs) is not easy to understand for the audience, the instructor will make a very short introduction. The workshop uses a set of open data in TIPD (Taiwan Indigenous Peoples Open Research Data, for details, see https://osf.io/ e4rvz/) as an example to demonstrate big data, open data, smart data, data science, and open science. TIPD complies with FARE (Findable, Accessible, Interoperable, Reusable) data principle. It consists of the following categories of open data from 2007 to 2022:

- 1. categorical data,
- 2. multi-dimensional data,
- population dynamics
- 4. temporal geocoding data,
- 5. household structure data,
- 6. traditional TIPs community data,
- 7. generalized TICD query system as a smart,
- 8. genealogical data.

In the end, the workshop will briefly highlight the impact of open data on promoting crowd collaboration and that of smart data on making effective policy decision-making by using interactive migration dynamics derived from TIPD as an example.

About the instructor:

DR. JI-PING LIN received his B.Sc. in Geography from National Taiwan University (Taiwan) in 1988, M.Sc. in Statistics from National Central University (Taiwan) in 1990, and Ph.D. in Geography in 1998 from McMaster University (Ontario, Canada). His main research specialty and interests include migration and population studies, labor study, survey study, scientific & statistical computing, big & open data, data science, and open science. He is serving as associate research fellow at Academia Sinica, Taiwan. The instructor worked in Taiwan's Bureau of Statistics & Census as research scientist, with abundant real-world experiences in processing, integrating, and enriching various sources of large-scale raw data, as well as in survey planning, sampling design, and conducting surveys. Lin has been serving as consultant for a number of Taiwan's central government agencies. Since 2013, the instructor devotes himself to the research on hard-to-reach population (HRP) and Taiwan Indigenous Peoples (TIPs). Based on the fundamentals of computational social science, data science and open science, he has been building a number of big open data and smart data.

WEDNESDAY 21/02/2024 01:30 PM - 04:30 PM

FLEXIBLE TEXT CATEGORISATION IN PRACTICE: USING AI MODELS TO ANALYSE OPEN-ENDED SURVEY RESPONSES

Instructor: Ludger Kesting (Tivian; Germany) Workshop language: English

Description of the content of the workshop:

Gaining an understanding of a powerful but easy approachable text analysis model, the advantages and disadvantages, approach and background, analysis dashboard for text analytics, background knowledge of open source data. Learning about analysis representations, manual completion, analysis approach, text classification, zero shot text classification, visualisation of the analysis, practical application in Tableau.

About the instructor:

LUDGER KESTING'S educational background is empirical social science and statistics based on my studies of sociology, computer science and ethnology. His focus and interest were always a data driven understanding of connections between people, social groups and cultures now breaking this down to employee and customer experience to help companies to listen to their employees, empower their leaders to drive success.

DR. MARTIN ANDREE



MR. ANDREE IS AUTHOR OF THE BOOK "BIG TECH MUST LEAVE! THE DIGITAL GIANTS DESTROY OUR DEMOCRACY AND ECONOMY - WE WILL

DIGITAL MONOPOLIES: HOW BIG TECH STOLE THE INTERNET – AND HOW WE CAN RECLAIM IT

Data measurements prove crystal clear: The huge amount of domains on the internet is meaningless, it is only a handful of tech companies which are attracting the majority of digital traffic. The rest of the internet resembles a huge graveyard. Digital monopolies are currently bringing ever larger parts of our lives under their control. The platforms are increasingly dominating the formation of political opinion and at the same time are deleting our free market economy. We should ask ourselves: is this still legal? Why should we put up with it any longer? Media scientist Martin Andree shows how far the hostile takeover of our society by the tech giants has already progressed - and how we can reclaim the internet.

Moreover, the keynote will address the specific role of science and market research, which should take a clear position towards these problems. The current destabilization of Western democracies shows how much people depend on competent scientific media research to provide guidelines and orientation for society in order to save our democracy, especially in times of disinformation and fake news. It's time to take over responsibility and make a change for the better – as long as it is still possible.

STOP THEM" (2023). He is teaching digital media at the University of Cologne. Since more than 15 years, he is undertaking research projects on the dominance of the big tech companies. On top, he is founder and CEO of AMP Digital Ventures with focus on semantic A.L., digital innovation and performance marketing. He is interviewed, published and featured by leading German media, such as the national television channels, leading newspapers and conferences, such as the digital summit of the German government or Gamescom. He has written seven books, among others the celebrated "Atlas of the Digital World".

DATA COLLECTION USING MOBILE APPS: WHAT CAN WE DO TO INCREASE PARTICIPATION?

There are limits to what can be measured with survey questions: we can only collect information about things our respondents know, can recall, are willing to tell us – and that fit within a timeconstrained questionnaire. Increases in smartphone ownership and use, along with technological changes are creating new possibilities to collect data for surveys of the general population, for example, through linkage or donation of existing digital data, collection of bio-samples or -measures, or use of sensors and trackers. Surveys are therefore developing into systems of data collection: depending on the concept of interest, different methods are used to generate data of the required level of accuracy, granularity, and periodicity.

For example, Understanding Society: the UK Household Longitudinal Study supplements the annual questionnaire-based data with linked data and data derived from bio measures and bio samples. In addition, we are developing and testing protocols to collect data using mobile applications, activity and GPS trackers and air quality sensors. We have conducted a series of mobile app studies, collecting detailed information about household expenditure, daily data about relationships, stressors and wellbeing, detailed body measurements, and spatial cognition. However, in each case, only a sub-set of respondents invited to the mobile app study participated and provided data.

In this talk I will present research from a series of experimental studies carried out on the Understanding Society Innovation Panel, that aim to identify the barriers faced by respondents in participating in mobile app studies, provide evidence on how best to design data collection protocols to maximise participation and reduce selectiveness of participants, and examine the quality of data collected with mobile apps.

PROF. ANNETTE JÄCKLE



ANNETTE JÄCKLE IS PROFESSOR OF SURVEY METHODOLOGY AT THE INSTITUTE FOR SOCIAL AND ECONOMIC RESEARCH AT THE UNIVERSITY OF ESSEX, UK, AND ASSOCIATE DIRECTOR OF INNOVATIONS FOR UNDERSTANDING SOCIETY: THE UK HOUSEHOLD LONGITUDINAL STUDY. She leads a programme of research on multi-modal methods of data collection for population surveys. Current projects include mixed mode surveys, respondent consent to data linkage, and data collection using mobile apps and sensors.

THE MEMBERS OF THE GOR POSTER AWARD 2024 JURY ARE:



JAN KAREM HÖHNE Jury Chair & DZHW



MAREIKE OEHRL September Strategie & Forschung GmbH



ÁDÁM STEVCOVICS HUN-REN Centre for Social Sciences



SEBASTIAN VERFÜRTH RTL Deutschland

THURSDAY 22/02/2024 02:30 PM - 03:30 PM

Posters offer the opportunity to present late breaking research, short research findings or discuss work in progress at the GOR conference. The presented works will be evaluated by a jury. The GOR Poster Award 2024 comes with a prize money of \notin 500.

Posters are presented in a plenary session on Thursday. Access to the poster presentations will be possible during the whole conference. A poster may cover any topic of online research. All submissions in this category are considered for the GOR Poster Award 2024.

The poster award ceremony will take place on Friday, 23 February 2024 at 10.45.

PAST WINNERS OF THE GOR POSTER AWARD ARE:

GOR POSTER AWARD 2023: Joshua Claassen for his Poster "From Text to Voice: Innovating the Final Comment Question in a Web Survey" and Daniela Wetzelhütter & Dimitri for their Poster "Al-Enabled Data Analysis Quality: Addressing A Knowledge Gap"

GOR POSTER AWARD 2022: Lisa Marie Natter for her Poster "The Impact of Monetary incentives on Retention Rates in a Panel Study with Mixed Web and Mail Modes"

GOR POSTER AWARD 2021: Tobias Rettig, Bella Struminskaya and Annelies Blom for their Poster "Memory Effects in Online Panel Surveys: Investigating Respondents' Ability to Recall Responses from a Previous Panel Wave"

GOR POSTER AWARD 2020: Adrian Hoffmann, Julia Meisters and Jochen for their Poster "Indirect Questioning Techniques: An Effective Means to Increase the Validity of Online Surveys"

THE MEMBERS OF THE GOR THESIS AWARD 2024 JURY ARE:



DR. OLAF WENZEL Jury Chair & Wenzel Marktforschung



DR. FREDERIK FUNKE LimeSurvey & datenmethoden.de



DR. CARINA CORNESSE DIW Berlin



PROF. DR. BETTINA LANGFELDT Kassel University



DR. TABEA TESCH Royal Canin, Mars Inc.

The GOR Thesis Award 2024 is sponsored by



THURSDAY22/02/202410:45 - 11:45 AND12:00 - 01:15

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

All submissions relevant to online research are welcome. Presentations in the past years covered a broad range of topics, be it online surveys or research on the Internet or social aspects of the Web. Theses must have been submitted in 2022 or 2023, the Thesis language can be either English or German. The programme committee assigned three members of the Thesis jury to review each submission. The review process was completely anonymised and reviews were randomly distributed in a way that conflicts of interests were avoided.

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

NOMINATED FOR THE GOR THESIS AWARD 2024 ARE:

GEORG AHNERT University of Mannheim, Germany: FAIR SAMPLING FOR GLOBAL RANKING RECOVERY

ISABELLE HALSCHEID Technische Hochschule Köln, Germany; Murmuras GmbH, Germany: UNDERSTANDING THE MOBILE CONSUMER ALONG THE CUSTOMER JOURNEY: A BEHAVIOURAL DATA ANALYSIS BASED ON SMARTPHONE SENSING TECHNOLOGY

CONSTANZE ROEGER TH Köln, Germany: EFFECTS OF ACTIVE AND PASSIVE USE ON SUBJECTIVE WELL-BEING OF USERS OF PROFES-SIONAL NETWORKS

ORIOL J. BOSCH University of Oxford, United Kingdom; The London School of Economics, United Kingdom; Universitat Pompeu Fabra, Spain: CHALLENGING THE GOLD STANDARD: A METHODOLOGICAL STUDY OF THE QUALITY AND ERRORS OF WEB TRACKING DATA

TIMO KOCH University of St. Gallen, Switzerland; LMU Munich: THE LANGUAGE OF EMOTIONS: SMARTPHONE-BASED SENTIMENT ANALYSIS

JULIAN B. AXENFELD German Institute for Economic Research - DIW Berlin, Germany: IMPUTATION OF MISSING DATA FROM SPLIT QUES-TIONNAIRE DESIGNS IN SOCIAL SURVEYS

CAMILLA SALVATORE Utrecht University, The Netherlands: ESSAYS ON INFERENCE FOR NON-PROBABILITY SAMPLES AND SURVEY DATA INTEGRATION

THE MEMBERS OF THE DGOF BEST PAPER AWARD 2024 JURY ARE:



DR. EVA AIZPURUA Nation Centre for Social Research, UK



PROF. DR. FLORIAN KEUSCH Jury Chair, University of Mannheim, Germany



DR. CATHLEEN STÜTZER TU Dresden, Germany



PROF. DR. NICOLA DOERING Ilmenau University of Technology, Germany



DR. HENNING SILBER GESIS – Leibniz Institute for the Social Sciences, Germany



CHAN ZHANG, PHD Zhejiang University, China

The DGOF Best Paper Award 2024 is sponsored by



FRIDAY 23/02/2024 10:45 – 11:15 (AWARD CEREMONY)

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

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

The award is worth 500 Euro and will be presented at the annual GOR conference. An abstract (and, if available, a preprint) of the award-winning paper will be posted to the DGOF website (www.dgof.de). To be considered for the award, papers must have been published in an outlet that uses a peer-review process (e.g., peer-reviewed journal, full papers in peer-reviewed conference proceedings, refereed book chapter) at the time of submission. Papers written in German or English and published not earlier than 2022 (if the paper was published online-first, then the online-first publication date counts) were eligible to be submitted for the DGOF Best Paper Award 2024.

WINNER:

REITER, TOBIAS, SCHOEDEL, RAMONA (2023). NEVER MISS A BEEP: USING MOBILE SENSING TO INVESTIGATE (NON-)COMPLIANCE IN EXPERIENCE SAMPLING STUDIES. BEHAV RES (2023).

Accessible on: https://doi.org/10.3758/s13428-023-02252-9

ABSTRATCS

ABSTRACTS 24

THURSDAY, 22/FEB/2024

A1: SURVEY METHODS IN-TERVENTIONS 1

(SPONSORED BY GESIS)

SESSION CHAIR: ALMUTH LIETZ DeZIM, Germany

PROVIDING APPRECIATIVE FEEDBACK TO OPTIMIZING RESPONDENTS – IS POSITIVE FEEDBACK IN WEB SURVEYS EFFECTIVE IN PREVENTING NON-DIFFERENTIATION AND SPEEDING?

MAREK FUCHS, ANKE METZLER

Technical University of Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION

Interactive feedback to non-differentiating or speeding respondents has proven effective in reducing satisficing behavior in Web surveys (Couper et al. 2017; Kunz & Fuchs 2019). In this study we tested the effectiveness of appreciative dynamic feedback to respondents who already provide well-differentiated answers and who already take sufficient time in a grid question. This feedback was expected to elevate overall response quality by motivating optimizing respondents to keep response quality high.

METHODS & DATA

About N=1,900 respondents from an online access panel in Germany participated in a general population survey on "Democracy and Politics in Germany". In this study, two 12-item grid questions were selected for randomized field-experiments. Respondents were assigned to either a control group with no feedback, or to experimental group 1 receiving feedback when providing non-differentiated (experiment 1) or fast answers (experiment 2) or to experimental group 2 receiving appreciative feedback when providing well differentiated answers (experiment 1) or when taking sufficient time to answer (experiment 2). Interventions were implemented as dynamic feedback appearing as embedded text bubbles on the question page up to four times and disappearing automatically.

RESULTS

Results concerning non-differentiation confirm previous findings according to which dynamic feedback leads to overall higher degrees of differentiation. By contrast, appreciative feedback to well differentiating respondents seems to be effective in maintaining the degree of differentiation only for respondents with particular long response times. Dynamic feedback to speeders seems to reduce the percentage of speeders and increase the percentage of respondents exhibiting moderate response times. By contrast, appreciative feedback to slow respondents exhibits a contra-intuitive effect and results in significantly fewer respondents with long response times and yields shorter overall response times.

ADDED VALUE

Results suggest that appreciative feedback to optimizing respondents has only limited positive effects on response quality. By contrast, we see indications of deteriorating effects when praising optimizing respondents for their efforts. We speculate that appreciative feedback to optimizing respondents is perceived as an indication that they process the question more careful than necessary.

COMPARING VARIOUS TYPES OF ATTEN-TION CHECKS IN WEB-BASED QUESTION-NAIRES: EXPERIMENTAL EVIDENCE FROM THE GERMAN INTERNET PANEL AND THE SWEDISH CITIZEN PANEL

JOSS ROSSMANN (1), SEBASTIAN LUNDMARK (2), HENNING SILBER (1), TOBIAS GUMMER (1)

1: GESIS – Leibniz Institute for the Social Sciences, Germany 2: SOM Institute, University of Gothenburg, Sweden

RELEVANCE & RESEARCH QUESTION

Survey research relies on respondents' cooperation during interviews. Consequently, researchers have begun measuring respondents' attentiveness to control for attention levels in their analyses (e.g., Berinsky et al., 2016). While various attentiveness measures have been suggested, there is limited experimental evidence comparing different types of attention checks with regard to their failure rates. A second issue that received little attention is false positives when implementing attentiveness checks (Curran & Hauser, 2019). Some respondents are aware that their attentiveness is being measured and decide not to comply with the instructions in the attention measurement, leading to incorrect identification of inattentiveness.

METHODS & DATA

To address these research gaps, we randomly assigned respondents to different types of attentiveness measures within the German Internet Panel (GIP), a probability-based online panel survey (N=2900), and the non-probability online part of the Swedish Citizen Panel (SCP; N=3800). Data were collected in the summer and winter of 2022. The attentiveness measures included instructional manipulation checks (IMC), instructed response items (IRI), bogus items, numeric counting tasks, and seriousness checks, which varied in difficulty and the effort required to pass the task. In the GIP study, respondents were randomly assigned to one of four attention measures and then reported whether they purposefully complied with the instructions or not. The SCP study replicated and extended the GIP study in that respondents were randomly assigned to one early and one late attentiveness measure. The SCP study also featured questions about attitudes toward and comprehension of attentiveness measures.

RESULTS

Preliminary results show that failure rates varied strongly across the

ABSTRACTS 25

different attentiveness measures, and that failure rates were similar in both the GIP and SCP. Low failure rates for most types of attention checks suggest that respondents were generally attentive. The comparatively high failure rates for IMC/IRI type attention checks can be attributed to their high difficulty, serious issues with their design, and purposeful non-compliance with the instructions.

ADDED VALUE

We conclude by critically evaluating the potential of different types of attentiveness measures to improve response quality of web-based questionnaires and pointing out directions for their further development.

EVALUATING METHODS TO PREVENT AND DETECT INATTENTIVE RESPONDENTS IN WEB SURVEYS

LUKAS OLBRICH (1,2), JOSEPH W. SAKSHAUG (1,2,3), ERIC LEWANDOWSKI (4)

1: Institute for Employment Research (IAB), Germany 2: LMU Munich

3: University of Mannheim, 4NYU

RELEVANCE & RESEARCH QUESTION

Inattentive respondents pose a substantial threat to data quality in web surveys. In this study, we evaluate methods for preventing and detecting inattentive responding and investigate its impacts on substantive research.

METHODS & DATA

We use data from two large-scale non-probability surveys fielded in the US. Our analysis consists of four parts:

First, we experimentally test the effect of asking respondents to commit to providing high-quality responses at the beginning of the survey on various data quality measures (attention checks, item nonresponse, break-offs, straightlining, speeding). Second, we conducted and additional experiment to compare the proportion of flagged respondents for two versions of an attention check item (instructing them to select a specific response vs. leaving the item blank). Third, we propose a timestamp-based cluster analysis approach that identifies clusters of respondents who exhibit different speeding behaviors and in particular likely inattentive respondents. Fourth, we investigate the impact of inattentive respondents on univariate, regression, and experimental analyses.

RESULTS

First, our findings show that the commitment pledge had no effect on the data quality measures. As indicated by the timestamp data, many respondents likely did not even read the commitment pledge text. Second, instructing respondents to leave the item blank instead of providing a specific response significantly increased the rate of flagged respondents (by 16.8 percentage points). Third, the timestamp-based clustering approach efficiently identified clusters of likely inattentive respondents and outperformed a related method, while providing additional insights on speeding behavior throughout the questionnaire. Fourth, we show that inattentive respondents can have substantial impacts on substantive analyses.

ADDED VALUE

The results of our study may guide researchers who want to prevent or detect inattentive responding in their data. Our findings show that attention checks should be used with caution. We show that paradata-based detection techniques provide a viable alternative while putting no additional burden on respondents.

B1: LARGE, LARGER, LLMS

SESSION CHAIR: DANIELA WETZELHÜTTER

University of Applied Sciences Upper Austria, Austria

FREE TEXT CLASSIFICATION WITH NEURAL NETWORKS: TRAINING, PROCESS INTEGRATION AND RESULTS FOR ISCO-08 JOB TITLES

PATRICK MERTES (1), SOPHIE TSCHERSICH (2)

1: Inspirient GmbH, Germany 2: Verian Group, Germany

RELEVANCE & RESEARCH QUESTION

Manual classification of job titles into classification systems like ISCO-08 is a time-consuming and labor-intensive task that often suffers from inconsistencies due to differences in expertise. The naïve approach of using lookup tables, while effective for a limited number of samples, falls short when handling larger and heterogeneous datasets, having different spelling styles for the same job. To address these challenges, our research seeks to answer the question: Can we build a classifier that outperforms the naïve approach and significantly improves the process of job title classification into ISCO-08?

METHODS & DATA

After a rigorous exploration of various machine learning approaches, we identified feedforward neural networks as the most effective for our task. Our training dataset comprises a diverse collection of approximately 100,000 distinct job titles (German), each meticulously labeled with the corresponding ISC0-08 code by human experts.

RESULTS

In our classifier, around 60% of classifications demonstrate confidence levels of 90% or higher, achieving an impressive accuracy of approximately 94% for those classifications. This allows for swift review of about 60% of the data. The overall accuracy of the classifier is approximately 73%, when no low confidence classifications are discarded. Despite this, manual review remains crucial for ensuring correctness, but the classifier significantly speeds up the process. Furthermore, the

classifier provides three ISCO-08 recommendations with confidence scores for each, giving human experts a solid reference point for manual classification, even in cases of lower confidence.

ADDED VALUE

Our data science approach offers several valuable benefits. Firstly, it significantly reduces the time investment required for this process, enabling researchers to process larger datasets efficiently. Secondly, it enhances the consistency of job title classification into ISC0-08, mitigating the issues associated with completely manual classification. Lastly, our classifier's accuracy is expected to improve even further as it continues to learn from more examples and corrections over time, making it a valuable and ever-evolving tool for labor market research. First tests show that the same approach works very well with other national and international economic and social classifications (e.g. ISCED, KIdB2010).

WHERE DO LLMS FIT IN NLP PIPELINES?

PAUL SIMMERING, PAAVO HUOVIALA

Q Agentur für Forschung GmbH, Germany

RELEVANCE & RESEARCH QUESTION

Large language models (LLMs) can perform classic tasks in natural language processing (NLP), such as text classification, sentiment analysis and named entity recognition. It is tempting to replace whole pipelines with an LLM.

But the flexibility and ease of use of LLMs comes at a price: their throughput is low, they require a provider like OpenAI or one's own GPU cluster and have high operating cost.

This study aims to evaluate the practicality of LLMs in NLP pipelines by asking, "What is the optimal placement of LLMs in these pipelines when considering speed, affordability, adaptability, and project management?".

METHODS & DATA

This study utilizes a mixed-method approach of benchmarks, economic analysis and two case studies. Model performance and speed are assessed on benchmarks. Economic considerations stem from prices for machine learning workloads on cloud platforms. The first case study is on social media monitoring of a patient community. It is centered on an LLM that performs multiple tasks using in-context instructions. The second case is large-scale monitoring of cosmetics trends using a modular pipeline of small models.

RESULTS

Small neural networks outperform LLMs by over 100-fold in throughput and cost-efficiency. Yet, without parameter training, LLMs attain high accuracy benchmark scores through in-context examples, making them preferable for small scale projects lacking labeled training data. They also allow flexibility of labeling schemes without retraining, which helps at the proof-of-concept stage. Further, they can be used to aid or automate the collection of labeled examples.

ADDED VALUE

LLMs have only recently become available for many organizations and drew new practitioners to the field. A first instinct may be to treat LLMs as a universal solution for any language problem.

The aim of this study is to provide social scientists and market research-

ers with references that help them navigate the tradeoffs of using LLMs versus classic NLP techniques. It combines theory with benchmark results and practical experience.

SENTIMENT ANALYSIS IN THE WILD

ORKAN DOLAY (1), DENIS BONNAY (2)

1: Bilendi & respondi, France

2: Université Paris Nanterre, France

RELEVANCE & RESEARCH QUESTION

The development of digital has brought new promises to qualitative and conversational surveys, but also raised new challenges on the analysis side. How should we handle numerous verbatim records? Al, especially LLMs seem to come to the rescue. But how reliable are these?

The aim of this presentation is to present the specific opportunities and challenges for AI focusing on the concrete case of sentiment analysis and to compare different AI models.

METHODS & DATA

We used a dataset of 167K pairs of the form <comment, rating> in English and French where participants were routinely asked to provide 0-10 ratings to evaluate how much they appreciated a certain product AND to provide a comment explaining their thoughts in connection with the rating.

Then we did assess the different SA models on how good they can guess participants' rating based on their comment?

Since SA models typically predict SA on a five point/stars scale, the question is whether those stars somehow match the ratings given by the participant.

We compared the google NLP suite vs NLP Town, an open source multi-lingual neural network model based on BERT. Basically, a multilingual BERT finetuned for sentiment analysis on 629k amazon reviews. Then we finetuned NLP town with inhouse datasets and external open data like Cornell and wikipedia extracts. Before comparing the performances with chatgpt3.5 and 4.

RESULTS

The perforamance vary between 50% and 70%. It's possible to use neural networks as an alternative to quant interrogation in qual-at-scale and to develop more conversational surveys

In this context, qual/quant data is useful to fine-tune open-source models such as NLP Town.

But the risk of generalization failure is very real. It proved very useful to supplement our qual/quant data by bringing in some extra open data (Wikipedia extracts as paradigmatic descriptive verbatims).

However, few shot learnings and last gen LLMs such as ChatGPT 4 do change the game by providing an 'effortless' (but still not perfect) solution.

ADDED VALUE

- Fondamental research...assessing reliability of Al in NLP
- comparing the reliability of different Al models
- comparing the reliability of different Armode
- identifying challenges and limitations

C1: MEDIA CONSUMPTION

SESSION CHAIR: FELIX CASSEL University of Gothenburg, Sweden

ANXIETY AND PSYCHOLOGICAL DISTANCE AS A DRIVE OF MAINSTREAM AND ONLINE MEDIA CONSUMPTION DURING WAR

VERED ELISHAR, DANA WEIMANN-SAKS, YARON ARIEL

The Max Stern Yezreel Valley College, Israel

RELEVANCE & RESEARCH QUESTION

This study examines media consumption patterns among Israeli users, during the 2023 Israeli-Hamas war. Drawing from the extensive body of literature on media use during wartime, this study investigates how civilians utilize different channels and platforms to fulfill their needs and perspectives amid this violent conflict. S

pecifically, consumption patterns will be analyzed as a function of users' level of anxiety, and their psychological distance from the war. We hypothesized that (1) The extent of individual anxiety will predict differences in mainstream versus online media usage, and that (2) Psychological distance from the war will mediate the relationship between anxiety and media usage patterns.

METHODS & DATA

A structured questionnaire was delivered among a nationally representative sample of Jewish -Israelis aged 18 and above (n=500) during the third week of the war, October 2023. Maximum standard error was set at 4.5%. Sample size calculations conducted using G*Power were based on a medium-sized effect size to achieve 90% power in detecting significant differences.

RESULTS

To test our first hypothesis (H1), a multiple regression analysis assessed the impact of anxiety on the usage of mainstream versus online media. The results indicated that anxiety significantly predicted an increase in mainstream media usage (B = .039, p < .05) but had no significant impact on alternative media usage (B = -.097, p > .05). suggesting that higher levels of anxiety were associated with a preference for mainstream media.

The second hypothesis (H2) involved a mediation analysis using Hayes' PROCESS macro. The analysis showed full mediation; the direct effect of anxiety on media usage became nonsignificant when accounting for psychological distance (B = .012, p > .05). However, the indirect effect of anxiety on media usage through psychological distance was significant (B = .053, 95% CI [.023, .129]), indicating that psychological distance completely mediates the relationship between anxiety and media usage patterns during wartime.

ADDED VALUE

This study contributes to the current literature on media consumption during wartime, by focusing on war-related anxiety as a drive, and by adopting 'psychological distance' to this field, analyzing it as another relevant variable.

ENGAGEMENT DYNAMICS AND DUAL SCREEN USE DURING THE 2022 FIFA WORLD CUP

DANA WEIMANN-SAKS, VERED ELISHAR, YARON ARIEL

Max Stern Academic College of Emek Yezreel

RELEVANCE & RESEARCH QUESTION

In this era of digital convergence, our study examines how psychological factors such as engagement, transportation, enjoyment, and media event perception influence dual-screen usage during the 2022 FIFA World Cup. It aims to unravel the complex dynamics between these factors and assess their impact on viewers' interactions with match-related and unrelated content across dual screens.

METHODS & DATA

We surveyed a representative sample of 515 Israeli participants using a structured online questionnaire, which assessed variables including transportation, enjoyment, media event perception, and dual-screen usage. Our study utilized Pearson correlations and Hayes' PROCESS model for advanced statistical analysis, exploring psychological factors' direct and indirect effects on dual-screen usage patterns.

RESULTS

We found significant positive correlations between engagement, transportation, enjoyment, and media event perception with match-related and unrelated dual-screen usage. Specifically, the Pearson correlation coefficients were r = .56 for engagement with match-related dual-screen usage (p < .001) and r = .37 for engagement with match-unrelated dual-screen usage (p < .001), highlighting the strong association between these psychological factors and dual-screen behaviors.

Engagement significantly mediated the relationships between media event perception, transportation, enjoyment, and dual-screen usage. In particular, for match-related dual-screen usage, the indirect effect of media event perception through engagement was significant (95% Cl, 0.067-0.149; F[2, 498] = 128.53; p < .001). For match-unrelated content, while direct effects were significant, indirect effects through engagement were not (95% Cl [-.288, -.015] for direct; [-.014, .088] for indirect), indicating varied influence patterns for different content types. All independent variables were positively correlated with match-related dual-screen usage and negatively correlated with match-unrelated usage. This suggests that higher levels of psychological engagement lead to more dual-screen activity related to the sports event.

ADDED VALUE

This study shows how psychological factors influence dual-screen usage during major sports events like the FIFA World Cup. It provides critical insights for media producers, advertisers, and digital strategists in developing engagement strategies and content for dual-screen platforms. It enriches the discourse on media consumption patterns in the context of global sports events, significantly enhancing our understanding of contemporary media engagement in a multi-screen world.

ABSTRACTS 28

D1: BEST PRACTICE CASES

SESSION CHAIR: YANNICK RIEDER Janssen-Cilag GmbH, Germany

BRAVE NEW WORLD! HOW ARTIFICIAL INTELLIGENCE IS CHANGING EMPLOYEE RESEARCH AT DHL.

SVEN SLODOWY (1), NESLIHAN EKINCI (2)

1: r)evolution GmbH, Germany 2: DHL Group

RELEVANCE & RESEARCH QUESTION

How can Al help to make complex and resource-intensive research projects simpler, more insightful, faster and more effective?

METHODS & DATA

Use of various AI models (generative, pre-trained large language models (GPT), algorithmic machine learning models, AI-supported decision tree analyses) for process optimization, deepening of knowledge and impact prediction.

RESULTS

Based on qualitative and quantitative figures, the results of 5 pilot projects are presented to show the opportunities, limits and risks of using Al for different tasks in large research systems.

ADDED VALUE

We try to estimate the possible financial, personnel and time savings by using AI and we want to show how AI can improve the outcome of a research system.

ABSTRACT

In an annual online survey, DHL Group collects structured and open feedback from around 550,000 employees. This Employee Opinion Survey (EOS) is conducted worldwide in 55 languages and for 60,000 organizational units. Due to its size, the operational implementation of the survey requires the use of large financial and human resources. It is not surprising that various stakeholders expressed a desire to optimize the survey: HR departments wanted more automation and a more effective follow-up process.

Team heads wished more specific recommendations for action and top management wanted an optimized use of resources.

EOS project team then asked itself how AI could help to make the project simpler, more insightful, and more effective. All process steps were evaluated, and five AI pilots/projects were rolled out. **1. Al to automate survey setup.** The challenge is that the reporting structure cannot be derived directly from the formal line organization. And therefore, it's a major manual effort for the HR departments to assign 550,000 employees to their teams. To optimize the process, we used machine learning models to automatically assign employees to the reporting structure.

2. Al to improve online questionnaire. In online surveys, the answers to open questions often remain short, as there is no interviewer to explore. To fill this gap, we used a GPT model that reacts individually to the respondent's open answer and asks additional in-depth questions.

3. Al to speed up the open comment processing. We used Al to translate, anonymize and categorize the 142,000 open comments in a fully automated process.

4. Al to make results dashboard more user-friendly and effective. We implemented a chatbot in our reporting tool that uses the current OpenAl GPT model. The chatbot starts with an individual management summary and answers specific questions of the users.

5. Al to predict which follow-up measures are particularly effective. We evaluated the initiatives documented in the action planning tool from previous years with next year's EOS results using an Al-supported decision tree analysis.

Using these five AI projects as examples, we show the opportunities, limits and risks of using AI, especially for large research systems.

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

MATTHIAS TOBIES (1), LEA THÖNNES (2), ARMGARD ZINDLER (1)

1: Ipsos GmbH, Germany 2: Aktion Mensch e.V.

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 in this context.

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.

T1: GOR THESIS AWARD 2024 COMPETITION: BACHELOR/ MASTER

SESSION CHAIR: OLAF WENZEL Wenzel Marktforschung, Germany

FAIR SAMPLING FOR GLOBAL Ranking Recovery

GEORG AHNERT

University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION

Measuring human perception of attributes such as text readability [Crossley et al., 2023] or perceived ideology of politicians (Hopkins and Noel, 2022] is oftentimes difficult because rating scales are hard to interpret. Pairwise comparisons between candidates—for instance: which of these two politician is more conservative?—pose a viable alternative. Given such pairwise comparisons, the task is to recover a global ranking of all candidates. This is non-trivial because of its probabilistic nature, i.e., a "weaker" candidate might win a comparison by chance. Furthermore, resources are often limited and not all pairs can be compared until a satisfactory estimate of individual strength is reached. Therefore, pairs of individuals must be selected according to a specified sampling strategy.

In recent years, a subfield of machine learning has developed around the quantification of fairness. While not without criticism, researchers propose fairness metrics and integrate fairness targets into machine learning algorithms. Lately, algorithmic fairness research has expanded from classification tasks to ranking scenarios as well. Yet, the fairness of ranking recovery from pairwise comparisons remains largely unexplored. This is particularly relevant since measured human perceptions are likely biased. For instance in hiring, pairwise comparisons between candidates for a position might not lead to the identification of the ideal candidate in the presence of such biases.

To the best of my knowledge, no previous research is concerned with the combined influence of sampling strategies and ranking recovery methods on the accuracy and fairness of recovered rankings. I thus propose the following research questions:

1. What is the effect of the sampling strategies and ranking recovery methods on overall accuracy?

2. Under which conditions do ranking recovery methods put an unprivileged group at a disadvantage?

3. Can sampling strategies or ranking recovery methods mitigate the effects of existing biases?

METHODS & DATA

In this thesis, I present a framework that manipulates the sampling of individuals for comparison in the presence of bias. I simulate individuals with latent "skill scores" on a certain task. I then separate the individuals into two groups and subtract a bias from the scores of the "unprivileged" group. I implement three distinct sampling strategies for selecting individuals from both groups for comparison: (1) random sampling (2) oversampling the unprivileged group and (3) sampling by previous success. Using the Bradley-Terry model (Bradley and Terry, 1952), I then simulate pairwise comparisons between the sampled individuals. On the simulated pairwise comparison data, I apply various ranking recovery methods including basic heuristics (David, 1987) and a state-of-the art method that involves graph neural networks: GNNRank (He et al., 2022). F

urther, I recover rankings with Fairness-Aware PageRank (Tsioutsiouliklis et al., 2021)—an algorithm developed for a different task, that is, however, group-aware and aims at eliminating bias.

In order to evaluate the interaction between sampling strategies and ranking recovery methods, I propose a novel group-conditioned accuracy measure tailored towards ranking recovery. Using this measure, I am able to evaluate both the overall accuracy of the recovered ranking, but also its fairness as operationalized through group representation (exposure) and group-conditioned accuracy.

I provide a Python package under MIT license to facilitate replication of my findings as well as for further investigation of fairness in ranking recovery.

RESULTS

Regarding the effect of sampling strategies, I find that both oversampling and rank-based sampling harm the accuracy of the recovered ranking. This is surprising as we would expect oversampling to improve the ranking accuracy of the unprivileged group that is oversampled. However, since this group's ranking accuracy also depends on correct comparisons against the individuals of the other group, the oversampled group's accuracy suffers as well. Oversampling thus is not a good remedy against biased comparisons.

In scenarios where there is no bias present against the unprivileged group, the optimal choice of ranking recovery method depends on the sampling that was used before pairwise comparison. Under random sampling, more advanced methods add little to no benefit in accuracy compared to heuristics based methods (i.e., David's Score). When oversampling or rank-based sampling is applied, however, GNNRank outperforms the other methods.

In the presence of bias against the unprivileged group, Fairness-Aware PageRank outperforms all other ranking recovery methods. Not only does it mitigate group representation bias from the recovered ranking, it also improves the ranking's accuracy when measured against the unbiased, latent "skill scores". This highlights the importance of groupaware ranking recovery over marginal benefits observed between the other ranking recovery methods.

ADDED VALUE

This thesis bridges the gap between previous research on fairness in machine learning and ranking recovery from pairwise comparisons. It is the first to introduce a framework for systematic investigation of fairness in ranking recovery and focusses on real-world sampling strategies and existing ranking recovery methods. Further, I propose a novel group-conditioned accuracy measure tailored towards ranking recovery. The results highlight the importance of fairness-aware ranking recovery methods and I supply recommendations on which ranking recovery method to use under which circumstances.

UNDERSTANDING THE MOBILE Consumer along the customer Journey: A behavioural data Analysis based on smartphone Sensing technology

ISABELLE HALSCHEID (1,2)

1: Technische Hochschule Köln, Germany 2: Murmuras GmbH, Germany

RELEVANCE & RESEARCH QUESTION

Digitalisation is shaping a new consumption era characterised by high connectivity, mobility and a broad range of easily accessible information on products, prices and alternatives. The modern consumers are broadly connected via social media and more mobile than ever with their smart devices. This empowers consumers to make sophisticated buying decisions based on a comprehensive amount of easily accessible online information, while having a broad range of options to choose from. Moreover, they compare prices, ask for opinions online and are willing to choose alternative products or services if they fit better in their lifestyle and meet their needs. As a result, it becomes more difficult than ever to understand modern consumers along their complex and dynamic path to purchase. However, since the modern consumers are constantly online through their smartphones, they produce a notable amount of data about their mobile and online behaviour such as movement, social media activities, online purchases or google searches. This behavioural data is immensely valuable for companies because it allows them to get a deep understanding about the mobile consumption behaviour of their customers. Yet, there is no solution on how to use this data to follow the consumers on their mobile devices. Therefore, this thesis investigates the extent to which mobile data collected with sensing technologies is useful to describe mobile consumer behaviour. The goal was to propose a first approach on how mobile data can be analysed to understand mobile consumers along their customer journey. For this purpose, an explorative analysis was conducted based on the following research question: What analyses can be performed using data generated with smartphone sensing technology to understand mobile consumer behaviour along the customer journey?

METHODS & DATA

As a first step, a literature review on current customer journey analytics theories, models and practices was conducted as foundation for the explorative data analysis. Because there could not be found any reasonable research that focuses on analysing customer journeys from mobile consumers, a mobile customer journey model was developed by adapting current models that are used among practitioners in customer journey analytics.

For the data analysis, the author collaborated with Murmuras, which developed a smartphone sensing technology for collecting sensing data via an application on participants' mobile phones. The collection process adheres to GDPR compliance standards, with data exclusively stored on servers located in Germany. Importantly, no personal information is tracked; instead, only consumption-relevant data is recorded. The company runs an ongoing incentivised smartphone sensing panel with a constant participant basis of approximately 1.500 smartphone users in Germany. Because of this, the thesis could be provided with long-term data from 01.10.2021 to 31.08.2022. This mainly included app usage data as well as mobile browser data (e.g. google search terms, website visits, etc.) and specific in-app content such as advertisement in the Facebook and Instagram app and in-app shopping content from the Amazon shopping app.

The data was provided and analysed via the platform Metabase, which mainly uses SQL-programming for analysing data. As the author has previous experience working with the data and analytics platform during a student internship, this knowledge could be used to transfer the mobile customer journey model into analytics concepts. Based on that, an explorative data analysis was conducted to explore the full potential of sensing data in the context of customer journey analytics.

RESULTS

The results show that mobile sensing data can be used in three main research areas among customer journey analytics: examining the touchpoint performance of a brand across mobile apps, describing different target groups by their smartphone usage behaviour and deriving real customer journeys on users' devices. For these areas interactive dashboards using different types of sensing data were developed.

The first dashboard focuses on analysing the touchpoint performance across various sensing datasets, including general app usage, in-app advertising, browser data, and Amazon shopping data. Key Performance Indicators (KPIs) were calculated to assess both general and app-related touchpoint performance. The integrated mobile customer journey provides an overview of all brand touchpoints over time, with detailed analyses of ads, browser interactions, and shopping behaviour. The second dashboard dives into target group analysis, aiming to understand mobile behaviour and preferences by providing insights into demographics, smartphone usage habits, contact channels, and mobile shopping behaviours on Amazon. The last part of the analysis employs the dashboards to conduct a deep analysis of an individual brand customer. This involved identifying relevant touchpoints, observing intercorrelations between touchpoints, analysing phone and mobile shopping habits, and mapping the customer journey stages. The insights gained from this analysis contribute to a comprehensive customer journey map and offer opportunities for the brand based on a deeper understanding of the consumers' mobile life.

ADDED VALUE

Although the vast amount of sensing data and the complexity of its analysis in the context of customer journey analytics remains challenging, it could be shown that sensing data presents a big opportunity for companies and researchers in this research area. It is not only possible to follow the relevant customers on their complex path to purchase, but also act on it by having the knowledge on how and where exactly to interact with their customers in the mobile world. As this has been a blind spot for companies and researchers before, they now have the ability to decode the whole customer journey of target groups by combining existing data with the insights derived from mobile sensing data. As sensing technology and sensing capabilities as well as smart devices are constantly improving, it is expected that an even more complete picture of mobile customer journeys can be analysed, which will add further value to customer journey analytics in future.

EFFECTS OF ACTIVE AND PASSIVE USE ON SUBJECTIVE WELL-BEING OF USERS OF PROFESSIONAL NETWORKS

CONSTANZE ROEGER

TH Köln, Germany

RELEVANCE & RESEARCH QUESTION:

Over the past decade online networking platforms have become integral parts of everyday life for most people, reshaping the way individuals communicate and network both privately and professionally. The growing popularity of these sites has sparked both enthusiasm and apprehension, resulting in a heated debate on the negative consequences of social network site (SNS) use on users' well-being in both popular culture and academia. Almost simultaneously with the rise of private network sites such as Facebook, professional network sites (PNSs) including LinkedIn have gained popularity. Despite the great interest in usage patterns (active and passive use) and the negative effects of SNS use on users' well-being, relatively little research has been performed on PNSs. Especially the association between PNS use and well-being has received very little academic attention so far. In view of the increasing popularity of PNSs for both private users and organizations this is surprising. Examining the impact on well-being is important as PNSs become more popular, leading to an increasing number of users who may be affected by the potentially harmful consequences such as decreased satisfaction with life, increased depressive symptoms or loneliness some authors have previously attributed to SNS use.

The aim of this study was to transfer previous findings on SNS use to the context of PNSs, exploring the multifaceted relationship between usage patterns and users' well-being leading to the following research questions:

RQ1 What is the relationship between PNS usage type and users' subjective well-being?

RQ2 What factors play a role in determining the influence of PNS usage type on the subjective well-being of the users?

RQ2.1 How does bridging social capital influence the relationship between active use and users' subjective well-being?

RQ2.2 How do social comparison and envy influence the relationship between passive use and users' subjective well-being?

METHODS & DATA

A quantitative online survey was conducted which yielded an adjusted total sample of 526 LinkedIn users (173 male, 350 female, 2 diverse, 1 undisclosed) aged 19 to 65 (M = 28.69, SD = 8.66). A convenience sample was recruited using WhatsApp, LinkedIn and university mailing lists. Additionally, three survey sharing platforms (i.e. SurveyCircle, SurveySwap and PollPool) were used.

According to the active-passive model of SNS use (Verduyn et al., 2017), which was employed as the theoretical framework for this thesis and transferred to the context of PNSs for this purpose, the effects of active and passive use on users' subjective well-being are explained by three mediating variables: social capital for active and social comparison as well as envy for passive use. Followingly, participants were asked to fill out measures regarding their usage pattern on LinkedIn, their subjective well-being, their tendency to engage in social comparison behavior, their experiences with envy as well as their levels of social capital.

Three mediation analyses were run using the PROCESS add on (Hayes, 2013) for IBM SPSS 28.0.1.0. To test the relationship between active LinkedIn use and subjective well-being, which was predicted to be mediated by bridging social capital, a simple mediation model was tested (model 1). Next, a serial mediation analysis was run to test the relationship between upward social comparison and envy as mediators in the relationship between passive LinkedIn use and subjective well-being (model 2). The same procedure was repeated, replacing upward comparison by downward social comparison (model 3).

RESULTS

Results of the mediation analyses revealed an indirect positive relation between active use of LinkedIn and well-being. Conversely, a negative indirect relation was found between passive use of LinkedIn and subjective well-being.

Bridging social capital fully mediated the relationship between active LinkedIn use and well-being (significant positive indirect effect ab = .0624, 95%-CI [.0303; .0999] and insignificant direct effect c' = .0967, p = .1237, 95%-CI [-.0191; .1585]).

As predicted, social comparison and envy acted as serial mediators in the relation between passive LinkedIn use and subjective well-being [model 2: a1d21b2 = -.0347, 95%-CI [-.0583; -.0120]; model 3: a1d21b2 = -.0101, 95% CI [-.0217; -.0009]].

Though, results of the two mediation models examining passive Linkedln use indicated possible omissions of other mediating variables as the direct effect between passive Linkedln use and subjective well-being (model 2: c' = .1692, p < .001, 95%-Cl [.0904; .2481]; model 3: c' = .1433, p < .001, 95%-Cl [.0651; .2215]) remained significant after the mediator variables were added to the model.

ADDED VALUE

The results of this thesis further expand upon previous research by examining users of PNSs. This study extends prior findings of other studies in two ways. First, it advances literature on online networking site use and well-being as it explores PNS use. Previous research mainly examined the relation between SNS use and well-being with special attention to Facebook. Moreover, prior studies have mainly focused on examining either passive or active use, while this study examined both usage patterns at once.

While results of this study are preliminary and should not be generalized, findings suggest that SNSs and PNSs share similarities, that lead to similar effect patterns when examining the relationship between usage patterns and well-being. Testing the active-passive model of SNS use (Verduyn et al., 2017) in the context of PNSs, revealed appropriate applicability. Results of this thesis also have practical relevance for both users and creators of platforms like Linkedln. Active use behavior should be promoted and encouraged as it has been associated with positive affects on users' well-being. When being educated on the different effects of usage patterns, users can proactively change their behaviors, positively affecting their well-being.

A2: MIXING SURVEY MODES

(SPONSORED BY GESIS)

SESSION CHAIR: JESSICA DAIKELER GESIS – Leibniz Institute for the Social Sciences, Germany

NAVIGATING THE DIGITAL SHIFT: INTEGRATING WEB IN IAB (PANEL) SURVEYS

JAN MACKEBEN

Institut für Arbeitsmarkt- und Berufsforschung, Germany

RELEVANCE & RESEARCH QUESTION

In the realm of social and labor market research, a noteworthy transformation has unfolded over the past few years, marking a departure from conventional survey methods. Traditionally, surveys were predominantly conducted through telephone interviews or face-to-face interactions. These methods, while effective, were time-consuming and resource-intensive. However, with the rapid advancement of technology, there has been a significant paradigm shift towards utilizing online modes for data collection.

The emergence of the web mode has revolutionized the landscape of surveys, offering a more efficient and cost-effective means of gathering information. Online surveys provide researchers with a broader reach, enabling them to engage with diverse populations across geographical boundaries. Moreover, the convenience and accessibility of web-based surveys have contributed to increased respondent participation.

As we navigate the digital age, the web mode has become increasingly integral in shaping the methodologies of social and labor market research. Its versatility, speed, and ability to cater to a global audience underscore its growing importance in ensuring the accuracy and comprehensiveness of data collection in these vital fields.

METHODS & DATA

In this paper, we focus on the largest panel surveys conducted by the Institute for Employment Research. These include the Panel Labor Market and Social Security (PASS), the IAB Establishment Panel (IAB-EP), the Linked Personnel Panel (LPP), consisting of both employer and employee surveys, and the IAB Job Vacancy Survey. Historically, all these surveys employed traditional data collection methods. However, in recent years, they all have undergone a transition by incorporating or testing the inclusion of the web mode.

RESULTS

In the presentation, I will provide an update on each survey's current status, illustrating how the web mode has been integrated and examining its impact on response rates and sample composition.

ADDED VALUE

The incorporation of the web mode in key Institute for Employment Research panel surveys is crucial in the digital age. This transition enhances efficiency, reduces costs, and broadens participant diversity, ensuring studies remain methodologically robust and adaptable to the evolving digital landscape.

EFFECT OF INCENTIVES IN A MIXED-MODE SURVEY OF MOVERS

MANUELA SCHMIDT

University of Bonn, Germany

RELEVANCE & RESEARCH QUESTION

The use of incentives to reduce unit nonresponse in surveys is an established and effective practice. Prepaid incentives have been shown to increase participation rates, especially for postal surveys. As surveys keep moving online and response rates keep dropping, the use of incentives and its differential effect on survey modes need to be further investigated.

In our experiment, we investigate both the effects of survey mode and incentives for participation rates in a postal/web mixed-mode survey. In particular, we aim to answer the questions:

- i) In which sociodemographic groups do incentives work (particularly well)?
- ii) Is the effect of incentives affected by survey mode?
- iii) How does data quality differ between incentivized and non-incentivized participants?

METHODS & DATA

Our data is based on a random sample of all residents who moved from two neighborhoods of Cologne, Germany, between 2018 and 2022. Addresses were provided by the city's Office for Urban Development and Statistics. We were also provided with the age and gender of all selected residents as reported on their official registration.

For the experiment, we randomly selected 3000 persons. Of those, 2000 received postal invitations to a web survey, while 1000 received a paper questionnaire with the option to participate online. In both groups, 500 participants were randomly selected to receive a prepaid incentive of 5 euros cash with the postal invitation.

RESULTS

Our design yielded a good response rate of around 35% overall (47% with incentives and 26% without). Over 80% participated in the online mode. As we have information on the age and gender of the whole sample, including non-responders, detailed analyses on the effectiveness of incentives and their possible effect on data quality (measured by the share of "non-substantive" answers, response styles, and the amount of information provided in open-ended questions), will be presented.

ADDED VALUE

With this paper, we contribute to the literature on the effect of incentives, particularly on the comparison of survey modes. As our data is based on official registration and we have reliable information on non-responders, our results on the effects of incentives are of high quality.

MODE MATTERS MOST, OR DOES IT? INVESTIGATING MODE EFFECTS IN FACTORIAL SURVEY EXPERIMENTS

SOPHIE KATHARINA HENSGEN (1), ALEXANDER PATZINA (2), JOE SAKSHAUG (1,3)

- 1: Institute for Employment Research, Germany
- 2: University of Bamberg, Germany
- 3: Ludwig-Maximilians University Munich, Germany

RELEVANCE & RESEARCH QUESTION

Factorial survey experiments (FSEs), such as vignettes, have increased in popularity as they have proven to be of great advantage when collecting opinions on sensitive topics. Generally, FSEs are conducted via self-administered interviews in order to allow the participant to understand and assess the given scenario entirely. However, many establishment panels, such as the BeCovid establishment panel in Germany, rely on interviewer-administered data collection (e.g. telephone interviews), but could also benefit from using FSEs when interested in collecting opinions on more sensitive topics. Thus, the question emerges whether FSEs conducted via telephone result in similar results compared to webbased interviews. Furthermore, it would be of great interest to know whether these modes differ in their answer behavior for FSEs, such as straightlining, extreme responding or item non-response.

METHODS & DATA

To shed light on this issue, a mode experiment was conducted in the BeCovid panel in which a random subset of telephone respondents was assigned to complete a vignette module online (versus continuing with the telephone mode). Respondents were given a set of 4 vignettes varying in six dimensions followed by two follow-up questions regarding the person's success in the application process. Additional to various descriptive analyses we run multilevel regressions (random intercept model) to take the multiple levels into account.

RESULTS

The analysis shows no overall difference in the results of the random intercept model when controlling for the mode. However, there are significant differences between the modes regarding specific dimensions of the vignette, which could be described as sensitive. Furthermore, CATI shows an increase of straightlining as well as extreme responding, but no influence on the probability of acquiescence bias or central tendency bias. Lastly, respondents interviewed via telephone lead to more item non-response.

ADDED VALUE

This study shows that conducting FSEs through telephone interviews is feasible, but is associated with certain limitations. Depending on the subject matter, these interviews might fail to accurately capture genuine opinions, instead reflecting socially accepted responses. Additionally, they may result in diminished data quality due to satisficing and inattention.

TOWARDS AN OPTIMAL SOURCE PROPORTION IN MIXED-MODE STUDIES: GUIDELINES BASED ON A META-STUDY

JENNY GEIGER, ANUJA HARIHARAN, ALEXANDRA WACHENFELD-SCHELL GIM Gesellschaft für Innovative Marktforschung mbH, Germany

RELEVANCE & RESEARCH QUESTION

[mixture, optimal proportions, representativity, observation error] In mixed-modes survey research, determining the best proportion of data whilst ensuring representativity and low observation error (OE) for the intended research question(s), is not a trivial problem. The sampling method (non-probability/probability) and recruiting method (offline/ online) further influences representativity & OE. Our research question: Is there a generalized approach to compute an optimal proportion of online and telephonic sources minimizing representativity error & OE in a desired variable of interest?

METHODS & DATA

[mixed-mode, CATI/CAWI, recruited online/offline, [Non]-probability] We conducted a meta-analysis of several mixed-mode studies with varying proportions of CATI/CAWI. Panelists were recruited offline and/ or online. Non-probability & probability sampling techniques were used and weights were applied. Sample sizes ranged from [1000 - 4000]. Demographic data on age, gender, federal state, education, profession were available. We compared the following across selected studies and 9 CATI/CAWI splits [10/90 - 90/10]:

1) measurement invariance; 2) representativity error; 3) OE by regressing on the variable "online usage behavior".

RESULTS

[measurement invariance, representativity, optimal source proportion]

Our analysis confirmed measurement invariance across demographic variables in both modes –with significant similarity of means & variances in t-tests or Chi-square tests, or a small Cohen's effect size (<0.2). Representativity error was stable across all splits, except for one study where "profession" was coded differently. OE was measured by comparing R² of the regressions - a steady increase in R² with higher proportion of CATI was observed. Moreover, employing an offline-recruited panel yielded a significantly higher R² than employing an online-recruited panel. In addition, an offline-recruited panel necessitated a lower proportion of CATI to achieve a comparable R² to the online-recruited panel. Thus, clear recommendations can be made for the path to an optimal source proportion in mixed-methods studies.

ADDED VALUE

[machine-learning analyses, cost-benefit]

Having a high CATI proportion afforded by cost & time, offline-recruited panels for CAWI, ensuring measurement invariance, and a suitable weighting procedure was shown to be most effective for maximizing representativeness and explanatory power in mixed-mode studies. Further work will explore the added value of suitable machine-learning analyses to learn and predict the optimal source proportion.

B2: AI TOOLS FOR SURVEY RESEARCH 1

SESSION CHAIR: TIMO LENZNER

Control Con

IN SEARCH OF THE TRUTH. ARE SYNTHETIC, AI GENERATED DATA THE FUTURE OF MARKET RESEARCH?

BARBARA VON CORVIN, ANNELIES VERHAEGHE Human 8 Europe, Belgium

RELEVANCE & RESEARCH QUESTION

Generative AI is the most talked-about topic among insight professionals, with 93% of researchers seeing it as an opportunity for the industry. With the rise of generative AI also came the rise of synthetic data. This data is artificially generated through machine learning techniques, rather than being observed and collected from real-world sources. Are we at the start of an era? What if synthetic data is taking over – being faster, cheaper, better (also in terms of privacy)?

To understand the potential of generative AI systems, Human8 has been on a journey to conduct what we like to call 'research on research'. Our primary focus has been on understanding how generative AI impacts qualitative research.

METHODS & DATA

By developing a personal AI research assistant using ChatGPT as an algorithm, we have been able to experiment with AI on also confidential research data and put generative AI to the test.

After we had conducted research with an online community of n=86 human participants, we created a synthetic pendant. With the help of Al and internet data, we developed an online community with 86 synthetic participants that were statistically and structurally identical to the original training data. We asked them the same research questions we had asked the human participants before. We created an artificial dataset with the same characteristics as the real-world dataset but without including any real-world data. And, we compared the outcome.

RESULTS

The results show very clearly how the conditions of ChatGPT influence study results. We will share the findings of our experiments and the learning for using synthetic data moving forward.

ADDED VALUE

Our presentation will help better understand the differences between

ш

data collected from human beings and synthetic, Al generated data. We will explain the reasons-why and provide some guidance on the cases in which the use of synthetic, Al generated data can be beneficial and, where the use of Al involves a risk.

CHATGPT AS A DATA ANALYST: Focus on the benefits and risks

DANIELA WETZELHÜTTER (1), DIMITRI PRANDNER (2)

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

RELEVANCE & RESEARCH QUESTION

Simple descriptive results can (now) be generated with ChatGPT in a relatively resource-efficient way - e.g. by generating a syntax code at the push of a button and then supporting the user in interpreting the output. The skills required to formulate the prompts and validate the descriptive results generated are still rather "manageable" compared to more complex analysis procedures (e.g. classification analysis). Errors that can occur start with the use of inappropriate analysis methods and range from the use of incorrect syntax to incorrect interpretation of results. This leads to the question: what are the benefits - and what are the risks of generating (and possibly using) incorrect results - of the new possibilities offered by Al-based data analysis?

METHODS & DATA

Based on this, the article will focus on the following aspects (using replication datasets with available, tested syntax codes) in the course of replication of already published studies

- Errors in the generated syntax code (e.g. omitting important steps, suggesting inappropriate statistical tests)
- Number of trials required (until an acceptable result is obtained or it is determined that no 'acceptable' result can be obtained)
- Usefulness of the results (e.g., clarity of interpretation, compactness)

RESULTS

The findings can be summarised as follows. The use of tools such as ChatGPT

- (1) is convincing for generating a decision basis for the choice of analysis method.
- (2) can support simple descriptive data analysis, result description and interpretation in a resource-efficient way.
- (3) is only advisable for generating a syntax for carrying out a complex procedure (e.g. MDS) with the appropriate expertise (possible for checking and adapting various specifics) – as it is very error-prone.

ADDED VALUE

The presentation focuses on the application of Al-supported data analysis in 'everyday research', which can be amateurish in nature, and emphasises the need to have the necessary skills to ensure the required quality of results. The aim of the resulting research is to develop a specific strategy for efficient 'scientific use'.

CHATBOT DESIGN AS AN ALTERNATIVE TO A MOBILE FIRST DESIGN IN WEB SURVEYS: DATA QUALITY AND RESPONDENT EXPERIENCE

CEYDA ÇAVUŞOĞLU DEVECI, MAREK FUCHS, ANKE METZLER

Technical University of Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION

The increasing use of smartphones in Web surveys requires an optimized questionnaire design for smartphones (Dillman et al., 2104). Since instant messaging is an integral part of every-day communication, this study investigates whether an instant messaging interface (chatbot design) can be used to improve the respondents' experience with Web surveys and whether data quality is comparable to Web surveys using a mobile-first design.

METHODS & DATA

In 2020, a survey on "Implications of COVID-19 on the Student's life" has been administered to a sample of 280 university students in Germany. Participants were randomly assigned to either a chatbot design or a mobile-first design. About half of the respondents in each design was invited to use their Smartphone to answer the survey while the other half was instructed to use a large screen device.

RESULTS

Results concerning the respondents' experience with the chatbot design indicate that even though the perceived level of difficulty of the chatbot design was rated significantly higher compared to the mobile-first design, the chatbot design was rated significantly more inventive and entertaining. Results concerning data quality were incoherent: Overall, item-missing rates of the two Web survey designs were on equal levels. In terms of number of characters of answers to narrative open-ended questions, there was no significant difference between the two designs. By contrast, the chatbot design yielded a lower degree of differentiation in one of two grid questions. Finally, results of the overall survey duration suggest that in the chatbot design group smartphone respondents exhibited marginally shorter response times than respondents using large screen devices.

ADDED VALUE

According to this study, the use of a chatbot design improves the respondents' experience with Web surveys even though the new chatbot design is still rated more difficult compared to the more traditional mobile-first design. Also, a chatbot design has the potential to minimize the burden (response time) of smartphone respondents. Results concerning data quality exhibit a so far inconsistent picture of the chatbot design that asks for a comprehensive assessment.

C2: ONLINE RESEARCH, ATTITUDES, PREFERENCES, BEHAVIOR

SESSION CHAIR: DANA WEIMANN SAKS The Max Stern Yezreel Valley College, Israel

CORRELATING ABORTION ATTITUDE MEASURES ACROSS SURVEYS: A NOVEL APPROACH TO LEVERAGING HISTORICAL SURVEY DATA

JOSH PASEK

University of Michigan, United States of America

RELEVANCE & RESEARCH QUESTION

The wealth of survey data amassed over the last century represents an invaluable tool for understanding human beliefs, attitudes, and behaviors and how these have evolved. But although thousands of datasets are available to researchers, scholars are often unable to use more than a handful for any given project. One challenge is that many questions, even those asking about similar topics, employ different wordings and response options. Hence, it is often difficult to tell whether differences between responses to questions are indicative of items that track subtly different topics, methodological choices, or changes over time. Instead, scholars examining trends often limit analyses to the subset questions asked identically at multiple time points. The current study proposes a novel solution to identifying common questions across data collections.

METHODS & DATA

Using microdata from over 2000 distinct probability US surveys of abortion attitudes, we produce a vector of means for each abortion measure at the intersections of age, gender, race, religion, and location. These can then be correlated across surveys (with appropriate weighting) to determine how similar the measures are and to identify measures that appear to capture similar underlying constructs (through clustering and other dimension reduction). We then parameterize how estimates of that similarity shift depending on the data collection methods, survey firms, and the temporal distance between surveys.

RESULTS

We show that this technique allows us to (1) identify the different types of historical questions that exist to measure views on abortion, (2) dis-

cern the similarity of those different types of questions, and (3) estimate how attitudes toward different types of questions have trended over time both overall and within population subgroups. We also find that stability of measures is relatively consistent for relations between items asked within 100 days of one-another, whereas it drops notably with longer time differences between measures.

ADDED VALUE

The study opens up novel methods for analysis of historical survey data.

DOES SURVEY RESPONSE QUALITY VARY BY RESPONDENTS' POLITICAL ATTITUDES? EVIDENCE FROM THE GGGS 2021

ALICE BARTH University of Bonn, Germany

University of Bonn, Germany

RELEVANCE & RESEARCH QUESTION

In standardized surveys, the quality of responses is essential. Numerous studies discuss how respondents' care and effort in answering survey questions is linked to personality, cognitive ability and socio-demographic variables such as age, education, and income, but only few researchers have studied the effect of political attitudes on response quality. Voogt & van Kempen (2002) find that differences between survey respondents and non-respondents in terms of political attitudes and behaviour exceed socio-demographic differences, and Barth & Schmitz (2016) argue that response quality systematically varies with ideological positions. Therefore, in this study we ask whether political attitudes are related to response quality in a standardized survey of the general population in Germany.

METHODS & DATA

German General Social Survey (ALLBUS/GGGS 2021), https://doi. org/10.4232/1.14002. The GGGS is a biennial survey based on a random sample of the German population. In 2021, due to the Covid-19 pandemic, it was fielded as a mail / web survey for the first time. The questionnaire was distributed in three randomized split versions. In the first step, indicators for response quality are constructed separately for each split version.

These include non-substantial "response styles", such as extreme responding and non-differentiation, as well as the proportion of item non-response. Subsequently, we conduct regression analyses with political attitudes (e.g. political interest, positions towards cultural and economic issues, intention to vote in upcoming election) as explanatory factors of response quality while controlling for socio-demographic variables, survey mode and number of contact attempts.

RESULTS

The analyses show that differences in response quality in the GGGS 2021 are systematically related to age, education and political interest as well as other political attitudes.

ADDED VALUE

Research on the nexus between response quality and political attitudes is highly relevant, as systematic relationships between response quality and substantive variables of interest may seriously compromise
surveys' ability to capture "public opinion". Whereas the GGGS is an offline-recruited population survey, effects of political attitudes on response quality are likely to be even more pronounced in (non-probability) online panels.

BUILDING THE CITY: A NOVEL STUDY ON ARCHITECTURAL STYLE PREFERENCES IN SWEDEN

FELIX CASSEL, ANDERS CARLANDER

University of Gothenburg, Sweden

RELEVANCE & RESEARCH QUESTION

Understanding citizens' architectural style preferences are important for aesthetically pleasing and sustainable urban environments. However, the opinions of the people are seldom considered in contemporary urban planning. In recent years, an intense debate has unfolded in Sweden about the role of politicians versus architects versus urban planners on how future urban landscapes should be built and, specifically, in what architectural style tradition. Thus, we explore architectural style preferences (classic vs modernist) of Swedish citizens. We model how preferences are predicted by sociodemographic and political factors.

METHODS & DATA

Data consisted of a non-probability sample (N=3119) and a probability sample (N=2125) from the online Swedish Citizen Panel at the University of Gothenburg. Participants were asked to state their preference for classic or modernist architecture and to associate design elements (building materials, level of details in facades and costs) with each tradition. Trust in various professional groups involved in urban planning, including architects, were also assessed.

RESULTS

Findings demonstrate a general preference for classic over modernist architecture in the non-probability and probability-based sample (p <.001). Notably, no differences among different party supporters where observed, indicating low political polarization on architectural preferences. Further, a logit regression model underscored the negative association between classic architectural preferences and social and political trust, as well as trust in architects (p <.001).

ADDED VALUE

We show that Swedish citizens have a clear preference towards classical architecture and that this support is stable across sociodemographic groups and party preference. The results provide insights for policy makers in urban planning. Replication and extension of these findings is currently being collected in a large population representative mixedmode survey on Swedish citizens. Results from the additional data collection will be presented at the conference.

FREQUENCY MATTERS? ASSESSING THE IMPACT OF ONLINE INTERRUPTIONS ON WORK PACE

EILAT CHEN LEVY (1), SHEIZAF RAFAELI (2), YARON ARIEL (1)

1: Max Stern Academic College of Emek Yezreel

2: Shenkar College of Engineering, Design and Art

RELEVANCE & RESEARCH QUESTION

With the increasing prevalence of digital work environments, understanding the impact of online interruptions on work efficiency becomes crucial. This study probes into how online interruptions' frequency and information richness affect the pace of work. It specifically examines two hypotheses: whether slow-frequency interruptions lead to a more efficient work pace compared to fast-frequency ones and how the nature of interruption information (lean: text-only vs. rich: image + text) influences the speed of work-related tasks.

METHODS & DATA

A 2 \times 2 factorial experimental design was implemented, involving 250 participants in a simulated online trading game in which participants had to gain profits. The experiment manipulated two main variables: interruption frequency (slow vs. fast) and information richness (lean vs. rich). Participants' task completion times were recorded to measure work pace. Statistical analyses, including ANOVA and post hoc tests, were conducted to determine the effects of these variables on work efficiency.

RESULTS

The ANOVA revealed significant main effects for interruption frequency on work pace, F[1, 246] = 8.97, p < .01, and for information richness, F[1, 246] = 6.54, p < .05. Participants dealing with slow-frequency interruptions had a mean work pace of 12.29 tasks per hour (SD = 4.21) compared to 13.87 tasks per hour (SD = 4.02) for those with fast-frequency interruptions. Surprisingly, lean information interruptions resulted in a faster work pace (M = 12.68, SD = 3.82) than rich information (M = 13.48, SD = 4.50). The interaction effect was significant, F[1, 246] = 9.33, p < .01, indicating that the most efficient work pace occurred under slow-frequency and lean information interruptions.

ADDED VALUE

This research sheds light on the nuanced effects of online interruptions in digital workplaces, challenging prevailing notions in media richness theory. By demonstrating that not only the frequency but also the type of information in interruptions can significantly influence work pace, it provides actionable insights for designing more productive digital work environments. These findings have implications for human-computer interaction designers, organizational psychologists, and workplace strategists aiming to optimize productivity in multi-tasking settings.

D2: INNOVATION IN PRACTICE: LLMS AND MORE ...

SESSION CHAIR: STEFAN OGLESBY data IQ AG, Switzerland

BEYOND REPORTS: MAXIMIZING CUSTOMER SEGMENTATION IMPACT WITH AI-DRIVEN PERSONA CONVERSATIONS

THEO GERSTENMAIER, KRISTINA SCHMIDTEL

Factworks, Germany

RELEVANCE & RESEARCH QUESTION

Segmentation is a challenging, core strategic research task, enabling businesses to understand the many needs of diverse customer groups. Yet, its effectiveness lies in its adoption within an organization. Despite Al's pervasive influence, its potential benefit in segmentation studies remains underutilized. This prompts us to explore: To what extent can Al help us socialize segmentation research, enabling stakeholder interaction with data and driving organizational adoption to influence business outcomes?

METHODS & DATA

Our research introduces an innovative approach leveraging Al-driven persona chatbots, tapping into Language Model-based systems like ChatGPT. Our aim is to create an interactive chatbot that can be shared across organizational departments, facilitating in-depth familiarization with customer segments. To do that, we train a GPT model on a comprehensive dataset combining quantitative and qualitative research findings from a study on segmenting online travel booking site users. This will enable us to evaluate its potential as a tool to humanize research findings and share accurate information about identified segments.

RESULTS

Our research plans to assess the chatbot's capacity to uphold factual accuracy based on its training data while also exploring its ability to generate creative yet aligned responses consistent with the characteristics of the segmented customer groups it represents. Initial assessments showcase promising signs of the chatbot's capacity to navigate between factual accuracy and creative engagement, aligning well with the segmented customer profiles it represents. However, its effectiveness heavily depends on well-engineered prompt design.

ADDED VALUE

In embracing this innovative approach, our goal is to create a tool that aids organizations in unlocking the full potential of segmentation. By en-

couraging greater immersion and fostering deeper empathy with consumer segments, our persona chatbot aims to make research findings more accessible to wider, less research-savvy audiences and enables them to get to know segments in a more playful and engaging way.

HOW GOOD ARE CONVERSATIONAL AGENTS FOR ONLINE QUALITATIVE RESEARCH?

DENIS BONNAY (1), ORKAN DOLAY (2), MERJA DAOUD (2)

1: Université Paris Nanterre, France

2: Bilendi

RELEVANCE & RESEARCH QUESTION

Conversational agents such as chatGPT open up new ways for online qualitative research. On the analysis side, they may be used to extract key ideas and to provide participants' quotes illustrating those. On the field management side, they may be used for moderation, to help dig deeper into what participants think. However, beyond the obvious advantages in terms of feasibility, numbers, speed and costs, the question how Al supplemented research design fares compared to purely human driven research appears as a pressing and hard to address question. Our goal in this research is more precisely to assess the quality of Al supplemented qualitative research for analysis and moderation, by comparison with human standards.

METHODS & DATA

We shall compare results obtained with and without a chatGPT powered Al assistant on a recently launched qualitative research platform www. bilendi.de/static/bilendi-discuss enabling the use of such an assistant. Regarding analysis, we qualitatively compare the results of a pure human analysis with those of the ChatGPT powered analysis. Regarding moderation, we quantitatively compare the response rate of participants to human moderators vs the chatGPT powered moderator. The data will consist in two data sets, a first study run in Finland in September 2023 gathering 22 participants, and a second study run in France and Germany and the UK over November-December 2023 gathering 30 participants per country. A pilot was run for this second demo in France in November 2023 with 225 participants.

RESULTS

In the Finnish study (analysis only), ideas provided by the ChatGPT powered assistant were found to be 70% consistent with those of the human analysis, 20% consistent but 'not usable as such' and 10% inconsistent. In the pilot study for the second demo (moderation only), response rate to human moderators was 85% and response rate to the ChatGPT powered assistant was 74.63%.

ADDED VALUE

Recent research by Chopra and Haaland ('Conducting Qualitative Interviews with AI', Cesifo Working Paper, 2023) provides encouraging evidence in terms of participants engagement and generated insights. The present research develops on those results by coming with systematic comparisons between human and machine performance.

SMARTPHONE APP-BASED Mobility Research

BEAT FISCHER

intervista AG, Switzerland

Thanks to GPS tracking with a smartphone app, a person's mobility behavior can be tracked in great detail. The information obtained on stages, routes, transport use and mobility purposes offers real added value in many areas of research. In this presentation, Beat Fischer explains the methodology, provides insights into the data science behind it and shows case studies with data from the Swiss Footprints Panel.

T2: GOR THESIS AWARD 2024 Competition: PHD

SESSION CHAIR: OLAF WENZEL Wenzel Marktforschung, Germany

CHALLENGING THE GOLD STANDARD: A METHODOLOGICAL STUDY OF THE QUALITY AND ERRORS OF WEB TRACKING DATA

ORIOL J. BOSCH (1,2,3)

1: University of Oxford, United Kingdom 2: The London School of Economics, United Kingdom

3: Universitat Pompeu Fabra, Spain

RELEVANCE & RESEARCH QUESTION

The advent of the Internet has ushered the social sciences in a new era of data abundancy. In this era, when individuals engage with online platforms and digital technologies, they leave behind digital traces. These digital traces can be collected for scientific research through innovative data collection methods.

One of these methods, web trackers, has gained popularity in recent years. This approach hinges on the utilization of web tracking technologies, known as meters, encompassing a diverse array of solutions that participants can install onto their devices. These meters enable the tracking of various traces left by participants during their online interactions, such as visited URLs. Historically, web tracking has been upheld as the de facto gold standard for measuring online behaviours. This thesis studies whether this prevailing notion holds true. Specifically, it explores the following questions: is web tracking data affected by errors? If so, what is the prevalence of these errors? To what extend do these errors introduce bias to web tracking measures? What is the overall validity and reliability of web tracking measures? And can we do something to limit the impact of web tracking data on the measurement quality of its measures?

METHODS & DATA

To explore these questions, this thesis uses data from the TRI-POL project. The TRI-POL is three-wave survey, conducted between 2021 and 2022, matched at the individual level with web tracking data. Data were collected through the Netquest opt-in metered panels in Spain, Portugal, and Italy, which consist of individuals who have meter(s) already installed in their devices and who can be contacted to conduct surveys. Cross quotas for age and gender, and for educational level, and region were used in to ensure a sample matching on these variables to the general country online populations.

The thesis is composed of three interconnected papers. The first paper, "When survey science met web tracking: Presenting an error framework for metered data", develops and present a Total Error framework for digital traces collected with Meters (TEM). The TEM framework (1) describes the data generation and the analysis process for metered data and (2) documents the sources of bias and variance that may arise in each step of this process. Using a case study, the paper also shows how the TEM can be applied in real life to identify, quantify, and reduce metered data errors.

The second paper, "Uncovering digital trace data biases: tracking undercoverage in web tracking data," adopts an empirical approach to address tracking undercoverage. This is a key error identified in the TEM: the failure to capture data from all the devices and browsers that individuals utilize to go online. The paper uses a new approach to combine self-reported data on participant's device usage, and paradata about the devices tracked, to identify undercoverage. Moreover, the paper estimates the bias introduced by different undercoverage scenarios, through the use of a Monte Carlo simulations.

The third and last paper, "Validity and Reliability of Digital Trace Data in Media Exposure Measures: A Multiverse of Measurements Analysis," explores the validity and reliability of web tracking data when used to measure media exposure. To do so, the paper uses a novel multiverse of measurements analysis approach, to estimate the predictive validity and true-score reliability of more than 7,000 potentially designable web tracking measures of media exposure. The reliability of the multiverse of measurements is estimated using Quasi-Markov Simplex Models, and the predictive validity of the measures is inferred as the association between media exposure and political knowledge (gains). Furthermore, the paper estimates the effect of each design choice on the reliability and validity of web tracking measures using Random Forests.

RESULTS

The TEM in the first paper suggest that web tracking data can indeed be affected by a plethora of error sources and, therefore, statistics computed with this data might be biased. Hence, caution should be taken when using metered data for inferential statistic. By clearly showing how web tracking data is collected and analysed, and identifying the errors of web tracking data, the framework allows to develop approaches to quantify those errors, and strategies to minimise them.

..........

Furthermore, the thesis shows - in the second paper- that tracking undercoverage is highly prevalent in commercial panels. Specifically, it reveals that across the countries examined, 74% of the panellists studied had at least one device they used for online activities untracked. Additionally, the simulations prove that web tracking estimates, both univariate and multivariate, are often substantially biased due to tracking undercoverage. As an example, across the different scenarios tested, undercoverage can inflate the proportion of participants identified as news avoiders by 5-21 percentage points, an overestimation of 29-123%. This represent the first empirical evidence demonstrating that web tracking data is biased. Moreover, it exposes deficiencies in the practices and procedures followed by both online fieldwork companies and researchers.

Focusing on the measurement properties of web tracking measures, the third paper shows that the median reliability of the entire universe of measurements explored is high but imperfect (≈ 0.86). Hence, in general, the explored measures of media exposure capture around 86% of the variance of their true score. Conversely, the predictive validity of the measures is low, given that overall the association between being exposed to media and gaining political knowledge is null. Although most self-reported measures of media exposure have been criticized precisely because of their lack of predictive power, results suggest that this is not limited to self-reports.

Hence, with the current evidence, web tracking measures of media exposure cannot be considered an improvement to self-reports. Additionally, results from the Random Forests suggest that the design decisions made by researchers when designing web tracking measurements can have a substantial impact on their measurement properties.

ADDED VALUE

Collectively, this thesis challenges the prevailing belief in web tracking data as the gold standard to measure online behaviours. It shows that web tracking data is affected by errors, which can substantially bias the statistics produced, as well as harm the reliability and validity of the resulting measures. In addition, the thesis demonstrates that high-quality measures can only be achieved through conscious design decisions, both when collecting the data (e.g., making sure all devices are tracked), and when defining how to construct the measurements. Methodologically, the thesis illustrates how a combination of traditional survey and computational methods can be used to assess the quality of digital trace data.

THE LANGUAGE OF EMOTIONS: SMARTPHONE-BASED SENTIMENT ANALYSIS

TIMO KOCH (1,2)

1: University of St. Gallen, Switzerland 2: LMU Munich

RELEVANCE & RESEARCH QUESTION

In an era transformed by artificial intelligence (AI) and the surge of voice assistants, chatbots, and other text or speech-based systems generating massive volumes of language data, automated emotion recognition and sentiment analysis have become integral across disciplines ranging from online marketing to user experience research. However, a main challenge has constrained previous research in this field: differentiating subjective emotional experience ("How do I feel in this moment?") from observable emotional expressions ("How do I express my feelings through language?"). While recognizing subjective emotions is of great scientific and practical relevance, the empirical difficulty of obtaining data on subjective emotional experiences and concurrent real-time language samples has limited the research. As a consequences, prior studies and deployed algorithms mainly relied on datasets composed of text or speech data either rated by participants for their emotional content or provided by actors, thereby focusing on emotion expression.

Here, the advent of conventional smartphones has provided a novel research tool, enabling the collection of self-reports on subjective emotional experience via apps and the gathering of everyday speech data through the smartphone's keyboard and built-in microphone. The present work leverages the ubiquity of smartphones, utilizing those capabilities to gather authentic text and speech samples, along with self-reported emotional states, bridging the gap between subjective emotional experiences and their linguistic expressions.

Thereby, the present dissertation addresses the research question if subjective emotional experience can be associated with and predicted from features in spoken and written natural language. Moreover, it identifies specific language characteristics, such as the use of certain word categories or voice parameters, associated with one's subjective emotional experience. Finally, this work examines the influence of the context of language production on emotional language.

METHODS & DATA

The present dissertation unfolds across two pivotal studies, employing everyday smartphones to collect rich datasets of both spoken and written language as well as self-reports on momentary emotional experience.

Study 1 analyzes subjective momentary emotion experience in more than 23,000 speech samples from over 1,000 participants in Germany (Study 1.1) and the US (Study 1.2). In Study 1.1, participants uttered predetermined sentences with varying emotional valences (positive/ neutral/negative) into their smartphones' microphones and self-reported on their momentary emotional states through an app. From the voice logs, vocal parameters (e.g., loudness, pitch, frequency) were algorithmically extracted. On the contrary, in Study 1.2, participants were given the freedom to express their current thoughts and feelings during the speech recordings alongside the emotion self-reports. Here, not only acoustic parameters, but also state-of-the-art word embeddings based on a Large Language Model (LLM) were extracted from participants' speech. Then, machine learning algorithms were employed to predict self-reported emotional experience from the extracted voice parameters and word embeddings. Also, interpretable machine learning methods were employed to identify the most important vocal features for emotion predictions.

Study 2 leverages a dataset of over 10 million typed words from 486 participants to investigate traces of subjective emotion experience in text data. Here, the smartphone's keyboard was utilized to log data on typing dynamics (e.g., typing speed), word use based on sentiment dictionaries and indirect emotion markers (e.g., use of first person singular), and emoji and emoticon use. Moreover, the logged data were enriched with contextual information on the app where the respective text had been produced as well as the input prompt text (e.g., "Was gibt's Neues?" on Twitter). This allowed to distinguish between private com-

munication, for example sending a message on WhatsApp, and public communication, like posting on Facebook. As in study 1, self-reported momentary emotional states and overall stable trait emotionality were assessed through an app. Then, descriptive correlations between self-reported emotion measures and language characteristics as well as machine learning models were investigated for different communication contexts and time aggregations (e.g., daily emotional experience vs. momentary emotions).

RESULTS

Results from study 1 indicate that while scripted speech offers limited emotional cues, spontaneous speech significantly enhances the prediction accuracy for emotions. Further, speech content showed a superior predictive performance compared to vocal acoustics in the employed machine learning models.

Also, for both prompted and spontaneous speech, the emotional valence of the spoken content had no effect on the algorithmic recognition of emotions from vocal features. Finally, interpretable machine learning methods revealed vocal features related to loudness and spectral fluctuation to be most relevant for emotion predictions from vocal parameters.

Study 2 reveals that sentiment dictionaries capture subjective emotion experience for large time windows, such as for overall trait emotionality or weekly emotion experience, but are limited for shorter periods, like momentary emotions. Besides those time effects, findings indicate that the context of language production has a significant impact on distinct emotion-related language variations.

Most prominently, the use of first-person singular words (e.g., "I," "me") correlated significantly stronger with negative trait emotionality in public communication than in private communication while the use of first person plural (e.g., "we") had a higher correlation with positive trait emotionality for private communication than public communication.

ADDED VALUE

In conclusion, the present dissertation sheds light on the complex interplay between language and subjective emotion experience. The two studies that underpin this dissertation are among the first pieces of research to collect and scientifically investigate everyday spoken and written language using conventional smartphones over an extended period, illustrating the promises of personal devices as a new data collection tool.

Moreover, the present work emphasizes the significance of the context of language production in emotion detection, demonstrating the potential for nuanced context-aware sentiment recognition systems to understand consumer sentiment and enhance user experience.

Finally, by highlighting the challenges of current emotion-recognition methodologies, this dissertation contributes to the academic discourse as well as the development of privacy-conscious sentiment detection technologies.

IMPUTATION OF MISSING DATA FROM SPLIT QUESTIONNAIRE DESIGNS IN SOCIAL SURVEYS

JULIAN B. AXENFELD

German Institute for Economic Research (DIW Berlin), Germany

RELEVANCE & RESEARCH QUESTION

In face of declining response rates and escalating costs in social survey research, more and more survey projects are switching from traditional face-to-face interviews to much less expensive self-administered online surveys. However, online surveys have comparatively narrow limits in questionnaire length due to a higher susceptibility for breakoffs. Thus, moving online may force survey designers to cut down on the number of questions asked in a survey, potentially resulting in the cancellation of important research projects due to limited resources. In this context, survey projects increasingly adopt innovative new data collection designs promising to reduce questionnaire length without dropping questions entirely from the survey, such as split questionnaire designs.

This is achieved by presenting each respondent only randomly assigned subsets of the questionnaire with the goal of imputing the planned missing data originating from this procedure thereafter. This dissertation addresses the imputation of social survey data from split questionnaire designs and the methodological decisions associated with implementing such surveys to facilitate imputation, asking how split questionnaires may be designed and how the resulting data may be imputed such that estimates could be achieved with a satisfying accuracy in practice based on the imputed data.

METHODS & DATA

Through a series of Monte Carlo simulations, drawing on real social survey data from the German Internet Panel and the European Social Survey, this research assesses the accuracy of estimates across various scenarios, encompassing the implementation of both the split questionnaire design and the subsequent imputation. It delves into the impacts of different split questionnaire module construction strategies, varying imputation techniques, the interplay between planned missingness and conventional item nonresponse, and the implications of general-purpose versus analysis-specific imputation on the accuracy of estimates for a multivariate model. In each simulation run, a split questionnaire design is simulated by allocating items to modules, randomly assigning a number of modules to each survey participant, and deleting all data from the modules not assigned. Thereafter, the data are multiply imputed and estimates calculated based on the imputed data. These estimates are then compared to benchmarks calculated from the complete data to assess their accuracy.

RESULTS

Main findings from this research involve:

1. With respect to the imputation, each respondent should receive a selection of questions from a large variety of topics rather than all questions from a selection of topics, as the latter leads to estimates with lower accuracy.

2. One may need to simplify imputation models with respect to the applied imputation methods and predictor sets to prevent highly inaccurate estimates, especially for relations between variables. For example, the imputation may benefit from excluding variables with near-zero correlations to the imputed variable from imputation models, or from

.........

ABSTRACTS 42

applying dimensionality reduction techniques on the predictor space to reduce the effective number of predictors.

3. Additional conventional item nonresponse by respondents may challenge the imputation especially if this implies large amounts of missing data from both sources combined, even if the nonresponse is missing completely at random. In this study, especially combined amounts of missing data exceeding 40% appeared harmful to the accuracy of estimates. Thus, even though a split questionnaire design allows for collecting data on more items than are presented to each individual respondent, there seem to be practical limitations on how much questionnaire length can be reduced without negative repercussions on data quality.

4. If the data are imputed for general research purposes to be supplied to a variety of third-party data users, the imputed data appear well-suited to be used for analyses of continuous relations in the entire survey sample. Conversely, estimating models with strongly non-continuous relationships (such as interactions or quadratic terms) or models based only on a subset of the survey sample could result in considerable biases, given the current state-of-the-art imputation procedures. For such analyses, the data would need to be imputed once more for this specific research objective, rather than for general purposes.

ADDED VALUE

The insights gleaned from these simulations thus offer valuable guidance and recommendations for future implementations of split questionnaire designs in online surveys: Split questionnaire survey designers should take care to present questions from preferably all survey topics to each respondent and make sure the split questionnaire design does not result in too large amounts of missing data, also taking into account their expectations about additional unplanned nonresponse.

Furthermore, researchers applying imputation to these data may need to reduce complexity in the imputation models to some extent, as for example through dimensionality reduction. Finally, if the data are imputed for general purposes, it should be communicated clearly for which kinds of analyses the imputed data could be used and for which analyses an analysis-specific imputation may be needed.

ESSAYS ON INFERENCE FOR NON-PROBABILITY SAMPLES AND SURVEY DATA INTEGRATION

CAMILLA SALVATORE

Utrecht University, The Netherlands

RELEVANCE & RESEARCH QUESTION

Probability sample surveys, which are the gold standard for population inference, are facing difficulties due to declining response rates and related increasing costs. Fielding large size probability samples can be cost prohibitive for many survey researchers and study sponsors. Thus, moving towards less expensive, but potentially biased, non-probability sample surveys or alternative data sources (big or digital trace data) is becoming a more common practice.

While non-probabilistic data sources offer many advantages (convenience, timeliness, exploring new aspects of phenomena), they also come with limitations. Drawing inference from non-probability samples is challenging because of the absence of a known sampling frame and random selection process. Moreover, digital trace data are often unstructured and require additional analysis to extract the information of interest. Additionally, there is no unique framework for evaluating their quality, and the lack of a benchmark measure can be a problem when studying new phenomena. Furthermore, it is important to evaluate the construct being measured, as it may be different from the one measured by traditional data sources. Thus, from a statistical perspective, there are many challenges and research questions that need to be addressed, such as the possibility of doing inference with non-probabilistic data, the quality of these data, and whether these data sources can replace or supplement traditional probability sample surveys.

The focus of this work is on answering three research questions: 1) What is the evolution of the field of survey data integration and what new trends are emerging?, 2) Can probability and non-probability sample surveys be combined in order to improve analytical inference and reduce survey costs?, and 3) How can traditional and digital trace data be combined to augment the information in traditional sources and better describe complex phenomena?

METHODS & DATA

The three research questions are addressed by three different studies. The first study presents an original science mapping application using text mining and bibliometric tools. In addition to characterizing the field in terms of collaboration between authors and research trends, it also identifies research gaps and formulates a research agenda for future investigations. From this research, it appears evident that data integration is a broad and diverse field in terms of methodologies and data sources. Thus, the second and third studies explore whether using non-probabilistic data can improve inference or can allow to study new aspects of a complex phenomenon.

The second study focuses on the structured and more traditional volunteer web surveys. In order to address the second research question, the paper presents a novel Bayesian approach to integrate a small probability sample with a larger online non-probability sample (possibly affected by selection bias) to improve inferences about logistic regression coefficients and reduce survey costs. The approach can be applied in different contexts. We provide examples from socioeconomic contexts (volunteering, voting behavior, trust) as well as health contexts (smoking, health insurance coverage).

The third study relates the analysis of traditional data in combination with unstructured textual data from social media (Twitter, now X). It shows how digital trace data can be used to augment traditional data, thus feeding smart statistics. On this purpose we propose an original general framework to combine traditional and digital trace based indicators. We show an application related to business statistics but it can be applied to all cases where traditional and new data sources are available.

RESULTS

In the second study, through the simulation and the real-life data analysis we show that the Mean Squared Errors (MSEs) of regression coefficients are generally lower when implementing data integration with respect to the case of no data integration.

Also, using assumed probability and non-probability sample costs, we show that potential cost savings are evident. This work is accompanied by an online application (Shiny App) with replication code and an interactive cost-analysis tool. By entering probability and non-probability (per-unit) sample costs, researchers are able to compare different scenarios of costs. These results can be used as a reference for survey researchers interested in collecting and integrating a small probability sample with a larger non-probability one.

The third study results in the development of a general framework to combine traditional and digital trace data. This framework is modular and it is composed of three layers, each describing the steps necessary for the technical construction of a smart indicator. The modularity of the framework is a key feature, as it allows for flexibility in its application. In fact, researchers can use the framework to explore different methodological variants within the same architecture, and potentially carry out improvements to specific modules or test for sensitivity of the results obtained at the different levels.

ADDED VALUE

Research in the field of survey data integration and inference for non-probability samples is expanding and becoming increasingly dynamic. Combining different data sources, especially traditional and innovative ones, is a powerful way to gain a comprehensive understanding of a topic, exploring new perspectives, and can result in new and valuable insights.

This work significantly contributes to the current debate in the literature by presenting original methodological findings and adopting a broad perspective in terms of analytical tools (text mining, Bayesian inference and composite indicators) and data sources (volunteer web surveys and textual data from social media).

Addressing the three research questions, it: a) enhances understanding of existing literature, identifying current trends and research gaps for future investigations, b) proposes an original Bayesian framework to combine probability and non-probability online surveys in a manner that improves analytic inference while also reducing survey costs, and c) establishes a modular framework that allows for building composite smart indicators in order to augment the information available in traditional sources through digital trace data.

The added value of this work lies in its presentation of diverse perspectives and case studies on data integration, showcasing how it can provide enhanced statistical analysis.

P: POSTERSESSION

FEAR IN THE DIGITAL AGE – HOW Nomophobia together with fomo and extensive smartphone use Lowers social and psychological Wellbeing

CHRISTIAN BOSAU, PAULA MERKEL

Rheinische Fachhochschule gGmbH (RFH), Germany

RELEVANCE & RESEARCH QUESTION

While FoMO (Fear of Missing Out) is already well known as an important factor that leads to extensive smartphone use (ESU) and lowers wellbeing (WB), research starts to look at the new phenomenon Nomophobia (the fear of being separated from one's smartphone and not being connected and reachable, e.g. Yildirim & Correia, 2015). However, it still remains unclear, how Nomophobia lowers wellbeing – social as well as psychological wellbeing – over and above the already known factors FoMO and ESU.

METHODS & DATA

This study (ad-hoc-sample: N=132) combines all factors in one design and investigates to what extend Nomophobia (measured by: NMP-Q-D Coenen & Görlich, 2022) is an additional factor that causes negative effects on wellbeing (measured by: FAHW, Wydra, 2020) over and above FoMO (measured by FoMO, Spitzer, 2015) as well as ESU (measured by: SAS-SV, Randler et al., 2016). Several regression analyses calculated the effect sizes for the main effects as well as the interaction effects for the different factors – controlled for age and gender.

RESULTS

Interestingly, different effects can be found regarding psychological wellbeing compared to social wellbeing. Whereas ESU (beta=-.31, p<.01) but not nomophobia lowered the psychological wellbeing quite a lot, instead Nomophobia (beta=-.18, p<.10) but not ESU lowered the social wellbeing significantly. FoMO was similarly a negative factor for psychological (beta=-.22, p<.05) as well as social wellbeing (beta=-.21, p<.05). Interaction effects between all of these factors were tested but could not be found. All in all, quite a part of the variance can be explained only by these three factors: 16% of the variance of psychological wellbeing and 12% of the variance of social wellbeing.

ADDED VALUE

This study extends the knowledge about the factors that causes negative effects on the wellbeing of people in the digital age. Smartphones are so prominent and important nowadays that the fear of losing them can cause additional harm. The results show, that they serve as an important connection tool for social relationships, losing them creates stress and their exorbitant use lowers the wellbeing of people.

IS LESS REALLY MORE? THE IMPACT OF SURVEY FREQUENCY ON PARTICIPATION AND RESPONSE BEHAVIOUR IN AN ONLINE PANEL SURVEY

JOHANN CARSTENSEN, SEBASTIAN LANG, HEIKO QUAST

German Centre for Higher Education Research and Science Studies (DZHW), Germany

RELEVANCE & RESEARCH QUESTION

Online surveys offer the possibility of interviewing participants of a panel more frequently at reasonable costs. A higher contact frequency might thereby lead to a lower rate of unsuccessful contact attempts through increased bonding with the respondents and address maintenance. If life history data is collected, a higher survey frequency also offers the advantage of shorter reporting periods and decreased time lag for the retrospective collection of these data (Haunberger 2010). This should reduce recall errors and the cognitive burden for respondents.

Nevertheless, more frequent interviews also increase the response burden or survey fatigue and could thus lead to a reduced willingness to participate (Haunberger 2010; Schnauber and Daschmann 2016; Stocké and Langfeldt 2003; Nederhof 1986). Until now there is insufficient empirical evidence for survey makers to decide on an optimal design when implementing online panel surveys (see most recently Zabel 1998 for very short wave intervals). Furthermore, existing evidence on survey frequency is limited to CATI and face to face interviews, constraining the validity of possible conclusions about online surveys. We are therefore analysing how the response rate changes when the survey frequency in an online survey is increased.

METHODS & DATA

To examine the effect of the survey frequency we implemented an experiment in a panel of secondary school graduates that surveys respondents every two years. To vary the survey frequency, an additional wave was conducted one year after the second wave for a random sample of participants. Both, control and treatment group, were interviewed again two years after the second wave. We compare response rates between these two groups in the latest wave.

RESULTS

We find a minimally higher response rate with a biennial survey – but without statistical significance. Thus, with a higher expected data quality, no (significant) losses in terms of response seem to be expected if the survey frequency is increased from biennial to annual.

ADDED VALUE

Our results serve as a guideline for survey makers on how to implement online panel surveys aiming for the sweet spot between optimized contact strategies, response burden, and high quality online panel data.

DIGITALISATION: CATALYZING THE TRANSITION TO A CIRCULAR ECONOMY IN UKRAINE

TETIANA GOROKHOVA

Centre for Advanced Internet Studies, Germany

RELEVANCE & RESEARCH QUESTION

Digitalization can contribute to the shift towards a sustainable circular economy (CE). Digitalization not only refines business processes but also emphasizes waste curtailment, prolonging product life, and slashing transaction costs. However, fully leveraging this integration presents challenges, with clear gaps hindering the fluid adoption of digital-backed circular business models.

Despite its significance, there's a dearth of comprehensive literature on digitalization's potential and challenges. This research aims to explore the main benefits and obstacles of applying digitalization in CE business models in Ukraine, focusing on identifying opportunities and challenges and finding ways to overcome these hurdles.

METHODS & DATA

The study involved interviews with business representatives, researchers, NGOs, and students (a total of 36 participants) during the thematic training course according to the Erasmus+ program in an online format in Ukraine. One of the activities during the training course was finding answers to four questions relevant to the research aim in small groups for 40 minutes for 7-8 participants.

RESULTS

I identified challenges related to the integration of circular principles into existing business models, data ownership, data sharing, data integration, collaboration, and competence requirements. The post-war rebuilding and modernizing of industries towards sustainability, visualization, and innovation in product design, enhancement in resource efficiency, and optimization of logistics process collaboration with stakeholders and implementation of digital technologies were noted as main opportunities in adopting business models based on CE in the Ukrainian perspective.

ADDED VALUE

This research uncovered less recognized or previously unexplored prospects linked to digitalization in the context of transitioning to a CE. One of the new opportunities is virtualization in business models can influence on reduce costs, conserve resources, and provide reliable data.

The research underscores the significant role of digitalization in enabling the transition towards a circular economy in Ukraine's business sector. While there are considerable opportunities for innovation and modernization, the challenges of integration, collaboration, data management, and skill gaps cannot be overlooked. Addressing these challenges through targeted educational programs, strategic partnerships, and supportive policies will be pivotal in harnessing the full potential of digitalization in advancing circular economy models.

DEVICE USE IN A FACE-TO-FACE Recruited Neighborhood Survey.

YFKE ONGENA, MARIEKE HAAN

University of Groningen, Netherlands, The

RELEVANCE & RESEARCH QUESTION

Due to the overall presence of smartphones and the ease of use of these devices, understanding the impact of device choice on survey data quality is becoming increasingly important. This study delves into the intricacies of a community survey conducted through both a paper flyer in the mail box and face-to-face recruitment by students. The primary objective is to explore the correlation between demographic characteristics and the selection of devices for survey completion.

Additionally, the study investigates variations in data quality, measured through completion time and response patterns such as straightlining, acquiescence bias, and midpoint responding.

METHODS & DATA

The target population consisted of all 5475 residents of a neighborhood in Groningen, living in 4035 households. In December 2023, a total of 3500 flyers were distributed to every address that was recognized as a home address with a separate mail box. Subsequently, students visited homes, encouraging residents to participate in the survey. Students referred to the flyer delivered in the mail box, but presented residents with a new flyer in case the flyer got lost. Participants were given the option to engage via a QR code (i.e., completion on a smartphone) or a concise URL (i.e., completion on a pc), with a sweet incentive of a cake as compensation for their contribution.

RESULTS

Within two weeks, 605 residents completed the questionnaire, resulting in a response rate of 17%. Notably, the QR code emerged as the preferred method for survey completion, with 85% opting for it, while the URL accounted for 15%. Interestingly, both students and individuals aged over 65 demonstrated a higher likelihood of using the URL. However, no significant associations were uncovered between completion time and the type of device chosen for survey participation.

ADDED VALUE

This study boasts a unique inclusion of all addresses within a single neighborhood in the recruitment sample, transforming it into a comprehensive population survey. In addition, in this study door-to-door recruitment and use of flyers that respondents use to decide their type of device distinguishes this study from earlier work.

LONG TERM ATTRITION AND SAMPLE Composition over time: 11 years of the german internet panel

TOBIAS RETTIG, ANNE BALZ

University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION

Longitudinal- and panel studies are based around the repeated interviewing of the same respondents. However, all panel studies are confronted with the loss of respondents who stop participating over time, i.e., panel attrition. Few studies have had the opportunity to observe attrition in the context of a panel study that features both frequent interviews and has been conducted over a long period of time and therefore offers many data points. In this contribution we investigate attrition rates over time and changes in sample composition for three samples in a probability-based online panel over a period of eleven years and 68 panel waves.

METHODS & DATA

We analyze participation data and respondent characteristics (e.g., socio-demographics) from 68 waves of the German Internet Panel (GIP) covering a time period from September 2012 to present. The GIP is the longest-running probability-based online panel in Germany and allows us to observe respondents from three recruitments samples drawn in 2012, 2014, and 2018, respectively.

RESULTS

Preliminary results indicate a high attrition rate over the first panel waves and a slower yet steady loss of respondents in the long term. On average, about 25% of recruited respondents were lost over the first year. The average annual attrition rate across all samples then falls to around 10% for the second and third year and a single-digit percentage for every year after that. Over time, a larger proportion of respondents in the remaining sample are married and hold academic degrees. The sample also slightly shifts towards a higher proportion of female respondents and persons living in single households. The proportion of respondents living in east or west Germany, their mean year of birth and employment status remain relatively unchanged.

ADDED VALUE

For longitudinal research and panel practitioners, it is important to understand how much attrition to expect over time and which groups of respondents are especially at risk. These insights aid in guiding researchers in determining how many respondents to recruit, when to refresh the sample and which respondents should be especially targeted with strategies for improving recruitment rates or reducing attrition.

ABSTRACTS 46

SAMPCOMPR: A NEW R-PACKAGE For sample comparisons and Bias analyses

BJÖRN ROHR, HENNING SILBER, BARBARA FELDERER

GESIS - Leibnitz Institute for Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION

The steady trend in declining response rates and the rise of non-probability surveys makes it increasingly important to conduct nonresponse and selection bias analyses for social science surveys or conduct robustness checks to evaluate if the results are robust across population subgroups. Although this is important for any research project, it can be very time-consuming. The new R-Package SampcompR was created to provide easy-to-apply functions for those analyses and make it easier for any researcher to compare their survey against benchmark data for bias estimation on a univariate, bivariate, and multivariate level.

METHODS & DATA

To illustrate the functions of the package, we compare three web surveys conducted in Africa in March 2023 using Meta advertisements as a recruitment method (Ghana n = 527, Kenya n = 2,843, and South Africa n =313) to benchmarks from the cross-national Demographics and Health Survey (DHS). The benchmarks will be socio-demographics and health-related variables such as HIV knowledge. In univariate comparison, bias is measured as the relative bias for every variable and, on an aggregated level, the average absolute relative bias (AARB). In bivariate estimation, we compare Pearson's r values against each other, and in multivariate comparison, different regression models are compared against each other.

RESULTS

Our poster will show examples of output from the package, including visualizations and tables for each comparison level. While the focus will be on figures, tables can also be useful for documentation and more detailed inspection. As to the specific content of our example, we will see that the social media surveys show a high amount of bias on a univariate level. In contrast, the bias is less pronounced on a bivariate or multivariate level. We will also report country differences in sample accuracy.

ADDED VALUE

Our R-Package will provide an easy-to-use toolkit to perform bias analyses and survey comparisons and, therefore, will be a valuable tool in the social research workflow. Using the same or similar procedures and visualizations for the various comparisons will increase comparability and standardization. The visualization is based on the commonly used R-package ggplot2, making it easily customizable.

ASK A LLAMA – CREATING VARIANCE IN SYNTHETIC SURVEY DATA

MATTHIAS ROTH

GESIS – Leibniz-Institut für Sozialwissenschaften in Mannheim, Germany

RELEVANCE & RESEARCH QUESTION

Recently there has been a growth of research on whether Large Language Models (LLM) can be a source for high quality synthetic survey data. However, research has shown that synthetic survey data produced by LLMs underestimates the variational and correlational patterns that exist in human data. Additionally, the process of creating synthetic survey data with LLMs inherently has a lot of researcher's degrees of freedom which can impact the distribution of the synthetic survey data. In this study we assess the problem of underestimated (co-)variance by systematically varying three factors and observe their impact on synthetic survey data: (1) The number and type of covariates a LLM sees before answering a question, (2) the model used to create the synthetic survey data and (3) the way we extract responses from the model.

METHODS & DATA

We use five socio-demographic background questions and seven substantive questions from the 2018 German General Social Survey as covariates to have the LLM predict one substantive outcome, the satisfaction of the respondent with the government. To predict responses to the target question we use LLama2 in its chat and non-chat variant, as well as two versions finetuned on German text data to control for differences between LLMs.

RESULTS

First results show that the (co-)variance in synthetic survey data changes depending on (1) the type and quantity of covariates the model sees, (2) the model used to generate the responses and (3) whether we simulate from the model implied probability distribution or only look at the most likely response option. Especially (3), simulating from the model implied probability distribution, improves the estimation of standard deviations. Covariances estimates, however, remain underestimated.

ADDED VALUE

We add value in three ways: (1) We provide information on which factors impact variance in synthetic survey data. (2) By creating German synthetic survey data, we can compare findings with results from research that has mostly focused on survey data from the US. (3) We show that using open-source LLMs enables researchers to obtain more information from the models than relying on closed-source APIs.

TO SHARE OR NOT TO SHARE? ANALYZING SURVEY RESPONSES ON SMARTPHONE SENSOR DATA SHARING THROUGH TEXT MINING.

MARC SMEETS, VIVIAN MEERTENS, JELDRIK BAKKER Statistics Netherlands, Netherlands, The

RELEVANCE & RESEARCH QUESTION

In 2019, Statistics Netherlands (CBS) conducted the consent survey, inviting respondents to share various types of smartphone data, including location, personal photos and videos, and purchase receipts. The survey particularly focused on understanding the reasons behind the reluctance to share this data. This study explores the following research question: What classifications of motivations and sentiments can be identified for unwillingness to share data with CBS, using a data-driven text mining approach?

RESULTS

This research applies multiple text mining techniques to detect underly-

ing sentiments and motivations for not sharing sensor measurements with CBS. The manually classified responses from the survey serve as valuable training and test data for our text mining algorithms. Our findings provide a comprehensive comparison and validation of manual and automated classification methods, offering insights into the effectiveness of text mining.

ADDED VALUE

The study underscores the potential of text mining as an additional tool for analyzing open-text responses in survey research. By using this technique, we detect sentiments and motivations, enhancing the understanding of respondents' perspectives on data sharing. This approach not only contributes to applying text mining in understanding attitudes towards data privacy and consent, but expands the methodology of survey research for analyzing open-ended questions and text data in general.

CAN SOCIALLY DESIRABLE RESPONDING BE REDUCED WITH UNIPOLAR RESPONSE SCALES?

VAKA VÉSTEINSDÓTTIR, HAUKUR FREYR GYLFASON

University of Iceland, Iceland

RELEVANCE & RESEARCH QUESTION

It is well-known that the presentation and length of response scales can affect responses to questionnaire items. However, less is known about how different response scales affect responses and what the possible underlying mechanisms are. The purpose of this study was to compare bipolar and unipolar scales using a measure of personality (HEXACO-60) with regard to changes is response distributions, social desirability and acquiescence.

METHODS & DATA

Four versions of the HEXACO-60 personality questionnaire were administered online via MTurk to 1,000 participants, randomly assigned to one of four groups, each containing one of the four versions. The first group received the HEXACO with its original response options (a fivepoint bipolar response scale), the second group received the HEXACO with a five-point unipolar agreement response scale, and the third group also received a unipolar agreement response scale but with three response options (the original response scale without the disagree response options). The fourth group was asked to rate the social desirability scale value (SDSV) of each of the 60 HEXACO items on a seven-point response scale (from very undesirable to very desirable). An index of item desirability was created from the SDSV and a measure of acquiescence was created by selecting items with incompatible content from the HEXACO to produce item pairs where agreement to both items would indicate acquiescence.

RESULTS

The three versions of the HEXACO-60 were analyzed with regard to distributions, social desirability of item content and acquiescence. The results show differences in the distribution of responses between the three response scales. Compared to the bipolar scale, the unipolar scales increased agreement with items rated as undesirable, which would indicate less socially desirable responding on unipolar scales. However, the use of unipolar scales increased overall agreement to items, which could indicate either increased acquiescence or different

interpretations of the question in relation to the response options. The results and possible interpretations will be discussed.

ADDED VALUE

The study provides added understanding of the effects of changing response scales from bipolar to unipolar and aids in understanding of the mechanisms underlying responses.

A3.1: SOLUTIONS FOR SURVEY NONRESPONSE

SESSION CHAIR: ZAZA ZINDEL Bielefeld University, Germany

DOES DETAILED INFORMATION ON IT-LITERACY HELP TO EXPLAIN NONRESPONSE AND DESIGN NONRE-SPONSE ADJUSTMENT WEIGHTS IN A PROBABILITY-BASED ONLINE PANEL?

BARBARA FELDERER (1,) JESSICA HERZING (2)

1: GESIS, Germany 2: University of Bern

RELEVANCE & RESEARCH QUESTION

The generalizability of inference from online panels is still challenged by the digital divide. Newer research concludes that not only individuals who do not have Internet access are under-represented in online panels but also those who do not feel IT-literate enough to participate which is potentially leading to nonresponse bias. Weighting methods can be used to reduce bias from nonresponse if they include characteristics that are both correlated to nonresponse and the variable(s) of interest. In our study we assess the potential of asking nonrespondents about their IT-literacy in a nonresponse follow-up questionnaire on improving nonresponse weighting and reducing bias. Our research questions are:

1.) Does including information on IT-literacy collected in the recruitment survey improve nonresponse models for online panel participation compared to standard nonresponse models including socio-demographics only?

2.) Does including IT-literacy improve nonresponse adjustment?

METHODS & DATA

Data are collected in the 2018 recruitment of a refreshment sample of the probability-based German Internet Panel (GIP). Recruitment was conducted by sending invitation letters for the online panel by postal mail. Sampled individuals who were not willing or able to participate in the recruitment online were asked to fill in a paper-and-pencil questionnaire asking about their IT-literacy. The questionnaire was experimentally fielded in the first invitation or reminder mailings. The control group did not receive a paper questionnaire.

RESULTS

We find IT-literacy to explain nonresponse to the GIP over and above the standard socio-demographic variables frequently used in nonresponse modeling. Nonresponse weights including measures of IT-literacy are able to reduce bias for variables of interest that are related to IT-literacy.

ADDED VALUE

Online surveys bear the risk of severe bias for any variables of interest that are connected to IT-literacy. Fielding a paper-and-pencil nonresponse follow-up survey asking about IT-literacy can help to improve nonresponse weights and reduce nonresponse bias.

YOUTH NONRESPONSE IN THE UNDERSTANDING SOCIETY SURVEY: INVESTIGATING THE IMPACT OF LIFE EVENTS

CAMILLA SALVATORE, PETER LUGTIG, BELLA STRUMINSKAYA Utrecht University, The Netherlands

RELEVANCE & RESEARCH QUESTION

Survey response rate are declining worldwide, particularly among young individuals. This trend is evident in both cross-sectional and longitudinal surveys, such as Understanding Society, where young people exhibit a higher likelihood of either missing waves or dropping out entirely.

This paper aims to explore why young individuals exhibit lower participation rates in Understanding Society. Specifically, we investigate the hypothesis that young people experience more life events such as a change in job, relationship status and a move of house, and it is the occurrence of such life events that are associated with a higher likelihood to not participate in the survey.

METHODS & DATA

The data source is Understanding Society, a mixed-mode probability-based general population panel study in the UK. We analyze individuals aged 18-44 at Understanding Society's Wave 1, and we follow them until Wave 12. We consider four age groups: 18-24 (youth), 25-31 (early adulthood), 32-38 (late adulthood) and 39-45 middle age (reference group for comparison). In order to study the effect of life events on attrition, we applied the Discrete-Time Multinomial Hazard Model. In this model the time is entered as a covariate and the outcome variable is the survey participation indicator (interview, noncontact, refusals or other). The outcome is modeled as a function of lagged covariates, including demographics, labor market participation, qualifications, household structure and characteristics, marital status and mobility, as well as binary indicators for life event-related status changes.

RESULTS

Consistent with existing literature, our findings reveal that younger respondents, as well as those with an immigration background, lower education, and unemployment status, are less likely to participate. We also demonstrate that changes in job status and relocation contribute particularly to attrition, with age remaining a significant factor.

ADDED VALUE

As many household surveys are moving online to save costs, the findings of this study will offer valuable insights for survey organizations. This paper enriches our understanding of youth nonresponse and presents practical strategies for retaining them. This project is funded by the Understanding Society Research Data Fellowship.

EXPLORING INCENTIVE PREFERENCES IN SURVEY PARTICIPATION: HOW DO SOCIO-DEMOGRAPHIC FACTORS AND PERSONAL VARIABLES INFLUENCE THE CHOICE OF INCENTIVE?

ALMUTH LIETZ, JONAS KÖHLER

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

RELEVANCE & RESEARCH QUESTION

Incentives for survey participants are commonly used to tackle declining response rates. It was shown that cash incentives are particularly effective in increasing response rates. However, the feasibility of cash incentives for publicly funded research institutions is not always guaranteed. As a result, other forms such as vouchers or bank transfers are often used. In our study, we aim to identify the extent to which socio-demographic and personal variables influence individuals' preference for either vouchers or bank transfers. In addition, we examine differences in preferences concerning specific vouchers from different providers.

METHODS & DATA

We draw on data from the DeZIM.panel - a randomly drawn, offline recruited online access panel in Germany with an oversampling of specific immigrant cohorts. Since 2022, regular panel operation has taken place with four waves per year, supplemented by quick surveys on current topics.

So far 9 regular waves have already been carried out. Within the surveys, we offer compensation in form of a \notin 10 postpaid incentive. Respondents can choose between a voucher from Amazon, Zalando, Bücher.de, and a more sustainable provider called GoodBuy. Respondents can also provide us with their bank account details and we transfer the money.

RESULTS

Analysis reveals that over half of the respondents who redeemed their vouchers chose an Amazon voucher and around 40 percent preferred to receive the money by bank transfer. Only a small proportion of 7 percent chose one of the other vouchers. This pattern can be seen across all waves. Initial results of logistic regressions show a significant preference for vouchers among those with higher net incomes. Additionally, we will examine participants who, despite not redeeming their incentive, continue to participate regularly in the survey.

ABSTRACTS 49

ADDED VALUE

Understanding which incentives work best for which target group is of great relevance when planning surveys and finding an appropriate incentive strategy.

A3.2: SURVEY INSTRUMENTS

SESSION CHAIR: CORNELIA NEUERT

GESIS — Leibniz Institute for the Social Sciences, Germany

UNMAPPED POTENTIALS: MEASURING AND CONSIDERING THE SELF-DEFINED RESIDENTIAL AREA OF INDIVIDUALS

MAXIMILIAN SPRENGHOLZ (1), ZERRIN SALIKUTLUK (2), Christian Hunkler (3)

1: Humboldt-Universität zu Berlin

- 2: DeZIM-Institut, Humboldt-Universität zu Berlin
- 3: Humboldt-Universität zu Berlin

RELEVANCE & RESEARCH QUESTION

Many research questions would benefit from information about individuals' self-defined residential area, i.e., where they spend their daily lives. However, the location information collected (or available to researchers given pseudonymization requirements) are typically larger aggregates, such as the postcode area. Even if respondents' addresses were available, we still would not know where they actually run their errands, go to the doctor, or spend family time on the playground.

However, this is the area substantially affecting their lives, and often the area they care about most. We collected data on the self-defined residential areas for research on anti-Muslim racism and examine the determinants of opposition to the opening of new Muslim-read establishments in those areas.

METHODS & DATA

As part of a representative online survey in Germany implemented by Kantar (n = 17.500), respondents drew their residential area as a polygon on a map using the open map tools offered by OpenStreetMap, Leaflet, and Leaflet Draw. Besides a wide range of socio-demographic and attitudinal measures, we also asked about several aspects of the self-defined residential area, e.g., the number of mosques, Turkish/Arabic-read restaurants and supermarkets. Moreover, we asked if respondents or their neighbors would oppose new establishments being built/ opened. We then merged the actual number of establishments in that area fetched via the Google Places API.

RESULTS [preliminary, new results to be expected by Nov 30]

Our results show that about 40 percent of respondents drew a plausible residential area (validated by postcode, shape, and size). First analyses indicate that respondents particularly oppose the opening of new Muslim-read establishments if there are none or very few in their residential area (e.g., not a single mosque).

ADDED VALUE

Although still preliminary, it appears that collecting information about respondents' residential areas works reasonably well in online surveys with tools already available. Once the area information is collected, it is easy to add point-referenced geographical data, e.g., from Google Places or OpenStreetMap, and crosswalks can be used to add data corresponding to other geographical units – all of which may offer valuable additional perspectives.

PARTNERSHIP BIOGRAPHIES IN SELF-Administered surveys: The effect Of screening-in information on Survey outcomes

LISA SCHMID, THERESA NUTZ, IRINA BAUER

GESIS - Leibniz-Institute for the social sciences, Germany

RELEVANCE & RESEARCH QUESTION

Many cross-sectional and panel studies survey retrospective biographical information, covering areas such as educational and occupational careers, fertility biographies, or former intimate relationship experiences. In interview-administered surveys, the use of event history calendars (EHC) has been established by collecting biographical data retrospectively and is evident to reduce recall effects and, thus, improve data quality. However, to handle raising survey costs, recent survey programs are increasingly conducted in self-administered modes, e.g., as web or mail surveys, or combine both in a mixed-mode approach. Additionally, the share of participants using their smartphone to respond to surveys is high and rising in self-administered modes. The lack of an interviewer, as well as displaying questionnaires on small screens impede the use of approved EHCs in self-administered surveys and at the same time increase the need for user-friendly survey tools and less complex questions.

Based on data from a survey experiment, we test whether visual feedback in complex survey modules improves the data quality of biographical information, i.e., relationship biographies. We set out to investigate:

 Is displaying of information from previous questions on partnership(s) related to survey outcomes regarding non-response and interview duration?

(2) Does the number of partnerships and the partnership duration vary with the display of information from previous questions?

 $\left(3\right)$ is the display of information from previous questions related to respondent burden?

METHODS & DATA

We run anova models on a sample of 3,446 respondents from a web survey conducted in December 2022. Within this survey, we vary the

display of the question list on relationships using information from previous questions on partnerships as visual feedback. As visual feedback, we used the names of respondents' (ex-)partners, their relationship status, and respective dates. The control group runs through the question lists without visual feedback.

RESULTS

Preliminary results do not show differences between the experimental and control group regarding non-response and interview duration. However, our results hint at differences in the number of relationships reported and respondents' perceived burden.

ADDED VALUE

Our study adds knowledge on how complex survey modules can be conducted without the presence of an interviewer.

CONSIDERING RESPONDENTS' PREFERENCES: THE EFFECTS OF SELF-SELECTING THE CONTENT IN WEB SURVEY QUESTIONNAIRES

KATHARINA PFAFF, SYLVIA KRITZINGER Universität Wien, Austria

RELEVANCE & RESEARCH QUESTION

Previous research has shown that the willingness to participate in surveys increases with the individual salience of the survey. In practice, this is usually taken into account to the extent that the survey topic and its relevance are presented in invitation letters hoping that it motivates a large and representative group to participate. In this study, we investigate how the willingness to participate in a survey and respondents' survey satisfaction change when they can choose from different topics.

METHODS & DATA

Our exploratory data analysis compares response rate, panel recruitment, and survey experience of 2,735 respondents. Descriptive statistics and Pearson's Chi-square test were used for the analysis. The sample is stratified by region and has been recruited offline from the Austrian Central Population Register. Respondents are randomly assigned to one of two questionnaire designs. In one design, the number and order are predetermined. Respondents assigned to the other design have the flexibility to decide how many topic blocks (modules) they would like to answer and in what order. Amongst others, the analysis examines the number, thematic preferences, and order of the selected modules. We also evaluate respondents' verbal feedback regarding the module selection choice.

RESULTS

Announcing that respondents can choose among different survey topics does not attract more or different respondents as opposed to an invitation letter to a survey, in which this option is not mentioned. There is also no difference regarding panel recruitment. Yet, the share of respondents being very satisfied with the survey is higher among those, who answer more modules than required for the incentive. Answers to an open-ended feedback question mirrors this satisfaction. While modules on politics are less often selected than others, most respondents choose all six modules.

ADDED VALUE

This study examines the effect of a dynamic online survey design, in which respondents can flexibly select - at least partially - the content of their survey. Although in practice it may not always be possible to adjust the entire survey's content to any respondent's preferences, the study highlights advantages and disadvantages of letting respondents choose parts of the survey.



SESSION CHAIR: ÁDÁM STEFKOVICS HUN-Ren Centre for Social Sciences, Hungary

BRIDGING SURVEY AND TWITTER DATA: UNDERSTANDING THE SOURCES OF DIFFERENC<u>ES</u>

JOSH PASEK (1), LISA SINGH (2), TRIVELLORE RAGHUNATHAN (1), CEREN BUDAK (1), MICHAEL JACKSON (3), JESSICA STAPLETON (3), LETICIA BODE (2), LE BAO (2), MICHAEL TRAUGOTT (1), NATHAN WYCOFF (2), YANCHEN WANG (2)

- 1: University of Michigan, United States of America
- 2: Georgetown University, United States of America
- 3: SSRS, United States of America; josh@joshpasek.com

RELEVANCE & RESEARCH QUESTION

For years, researchers have attempted to use social media data to generate inferences typically produced using surveys. But Twitter data and other social media traces do not consistently reflect contemporary survey findings. Two explanations have been proposed for why this might be the case: one posits that the set of people producing data on social media sites differs from those recruited to surveys; the other asserts that data generating processes are sufficiently different that it does not make sense to compare their social media and survey outputs directly.

METHODS & DATA

This study links a probability US sample of survey respondents with those same individuals' Twitter data as well as with decahose Twitter data. We compare four datasets to understand links between samples and data generating processes. These include survey responses on three topics for (1) a probability sample of the US public (N=9544); (2)

the same survey responses for the subset of individuals who use Twitter, consent to access, and tweet about the topics of interest (N=246); (3) tweets for this set of linked individuals who tweeted about the topic of interest; and (4) tweets from US individuals sampled from the Twitter decahose (N=7,363 after removing bots and non-individual accounts). Open-ended survey questions and social media posts are topic modeled using a guided topic modeling approach within topic areas to identify vaccination behaviors/attitudes, economic evaluations, and parenting challenges during the COVID pandemic.

RESULTS

We find that the subset of individuals who used Twitter and consented to linkage differed slightly in demographic composition, but mentioned similar distribution subtopics in response to open-ended survey questions about all three areas. In contrast, individuals with survey and Twitter data provided similar data across these two modes for one of our three topics (economics) and different data across the other topics (vaccinations and parenting). Tweets from consented users and the decahose sample, in contrast, provided similar distributions of topics for vaccinations and parenting, but not economics.

ADDED VALUE

This suggests that motivation to post and posting frequency may be more important for data acquired than who is represented.

PHYSICAL PROXIMITY AND DIGITAL Connections: the impact of geographic location on twitter user interaction

LONG NGUYEN (1), ZORAN KOVACEVIC (2)

1: Bielefeld University 2: ETH Zürich

RELEVANCE & RESEARCH QUESTION

In the context of an online social network where geographical distance is often assumed to be inconsequential, this study examines how physical proximity relates to Twitter user interaction. In line with previous findings, the central hypothesis is that individuals who live in closer physical proximity are more likely to engage with one another, despite the virtual nature of Twitter. Moreover, the extent of this impact is expected to be contingent on the specific topic under discussion.

METHODS & DATA

Employing a multi-layered approach, the study integrates techniques from natural language processing, network analysis, and spatial analysis. A dataset of over 500 million geolocated German tweets (including retweets) forms the basis of the analysis. First, a BERT-like language model is trained on the tweets to categorise them into thematically similar groups, enabling a granular exploration of topic-specific interactions.

Subsequently, retweet and reply networks are constructed for each thematic group as well as for the entire tweet corpus. Community detection algorithms are then used to identify clusters of users who frequently retweet and reply to each other. Spatial analysis is then applied to examine the correlation between users' physical proximity and their clustering as identified by community detection.

RESULTS

Preliminary results indicate a corpus-wide positive correlation between the spatial proximity of users and their clustering based on retweet and reply communities. However, the strength and significance of the correlation varies across the different topics discussed within the Twitter dataset. Notably, the geographical aspect of discussions can be found not only among local topics, but also in topics with a more universal appeal.

ADDED VALUE

This study offers a methodologically complex investigation of the interplay between geography and online social networks. By revealing the nuanced relationship between spatial proximity and Twitter user interaction based on topics, the study extends our understanding of online social dynamics. The findings contribute to the broader discourse on social media by highlighting the importance of local context and regional differences as a determinant of online interaction patterns.

GENDER (SELF-)PORTRAYAL AND STEREOTYPES ON TIKTOK

DORIAN TSOLAK (1,2,3), STEFAN KNAUFF (1,2,3), LONG H. NGUYEN (1,2), Rian Hedayet Zaman (1), Jonas Möller (1), Yasir Ammar Mohammed (1), Ceren Tüfekçi (1)

1: Bielefeld University, Germany

2: Bielefeld Graduate School in History and Sociology, Bielefeld 3: Institute for Interdisciplinary Research on Conflict and Violence, Bielefeld

RELEVANCE & RESEARCH QUESTION

Women and men are portrayed differently in advertising and on social media, as research on gender (self-)portrayal has shown. Most studies in this area analyzed small samples of static images to examine gender stereotypes conveyed through images on social media.

We study gender (self-)portrayal on TikTok, in particular which dynamic expressions are more often used by individuals passing as women or men. For this, we present a novel method to analyze large amounts of video data with computational methods.

METHODS & DATA

Our data encompasses approximately 36,000 unique videos extracted from the top 1000 trending TikTok videos in Germany over a consecutive 40-day period in 2021, supplemented by 973,000 metadata entries. Each video is processed using YOLOv8 pose detection, which dissects the videos into frames and annotates 17 key points per frame. We group the data into commonly used dynamic expressions (i.e., sequences of body movement).

We employ HDBSCAN and DTW to deal with differences of sequence and video length and to handle 'valid' missing data, e.g., from certain body parts not being visible in the footage.

RESULTS

Sequences are grouped into prototypes of dynamic expressions. Using manually annotated information, we can distinguish certain types of movement that are more commonly used by one gender. Utilizing metadata and expressions in the videos, we are able to explain a part of the variance of how a video performs, i.e. how many likes it gets or how long it stays within the top 1000 trends. A qualitative assessment of the

..........

prototypes of the most gender-biased expressions allows for integration with sociological theory on gender stereotypical body posing and provides insight into why some poses might perform better regarding likes and views.

ADDED VALUE

We extend the framework for analyzing gender stereotypical posing from static social media images to dynamic social media videos, which is an important endeavor to adapt to the trend of video-based social media content (Snapchat, TikTok, Instagram reels) becoming the de facto default type of content, especially for younger generations. Regarding methods, we offer a tractable way to analyze body posing on social media.

C3: ARTIFICIAL INTELLIGENCE

SESSION CHAIR: SANDRA WALZENBACH Konstanz University, Germany

AI: FRIEND OR FOE? CONCERNS AND WILLINGNESS TO EMBRACE AI TECHNOLOGIES IN ISRAEL

VLAD VASILIU (1), GAL YAVETZ (2)

1: Academic College of Emek Yezreel, Israel 2: Bar-Ilan University, Israel

RELEVANCE & RESEARCH QUESTION

Research on Al has a long history, spanning seven decades (Jiang et al., 2022), but only recently have scholars began exploring Al's impact on everyday activities (Ertal, 2018). Over the last two years, one could witness a surge in the use of large language models like ChatGPT, Bard, and Dall-e2. This study investigates people's concerns about Al replacing their roles and their willingness to embrace these technologies, focusing on traditional predictors of fear and adoption: income, education, and age.

METHODS & DATA

A representative survey of the adult (18+) Jewish population in Israel was conducted (n=502) via an internet panel (iPanel) in the beginning of 2023. It was comprised of demographic and perspectives on Al technologies questions.

RESULTS

Results indicate a significant negative correlation between income, education, and age with fears of AI replacing jobs (rs = -.179, p < .001; rs = -.108, p < .01; rs = -.096, p < .05). Additionally, a borderline significant positive correlation between willingness to adopt AI models and educa-

tion (rs = .071, p = .055) and a significant negative correlation with age (rs = .088, p < .05) were found. No correlation was observed between income and the willingness to adopt these technologies (rs = .019, p > .05).

ADDED VALUE

Notably, this research reveals a unique finding; Contrary to previous studies showing negative correlations between fear of technology and income or education, the fear of adopting new technologies is inversely related to age. As people grow older, their fear of adopting technology diminishes, likely because these tools offer a user-friendly interface resembling existing chat bots, requiring no new technological literacy. Another possible explanation is that the respondents feel secure in their workplace positions regardless to the new technologies.

Moreover, the lack of a correlation between income and willingness to adopt may stem from the low (sometimes free) cost associated with these technologies. In an era of rapid AI development and integration into daily life, studies like this one hold significance in understanding public sentiments surrounding these tools and their implications for personal and professional life.

HUMAN ACCURACY IN IDENTIFYING AI-GENERATED CONTENT

HOLGER LÜTTERS (1), MALTE FRIEDRICH-FREKSA (2), OSKAR KÜSGEN (3)

1: HTW Berlin, Germany

- 2: horizoom GmbH, Germany
- 3: pangea labs GmbH, Germany

RELEVANCE & RESEARCH QUESTION

The research addresses a significant question in the era of advanced digital technology: "Are humans ready to detect Al-generated content?" This question is pivotal as it explores human perception and understanding in the face of rapidly evolving Al capabilities in times of deep fakes on all media platforms.

METHODS & DATA

The empirical approach is using a digital interview with n>1000 Germans exposed to a variety of Al-generated and human-created content. In three categories (pictures, audio, videos) the participants were asked to identify the source of each piece of content, whether it was produced by Al or by a human. The content itself was created using Al Tools and stock content sources. The questionnaire is using implicit measurement and pairwise comparisons using the Analytic Hierarchy Process (AHP) methodology.

RESULT

The findings reveal intriguing insights into the human ability to discern Al-generated content. A significant proportion of participants are heavily challenged in correctly identifying the nature of the content, with varying degrees of accuracy across different types of media. These results highlight the sophistication of current Al technology in mimicking human creativity and the challenges faced by individuals in distinguishing between the two.

ADDED VALUE

This study adds substantial value to the discourse on Al and human interaction. It provides empirical evidence on the current state of human perception regarding Al-generated content in Germany, offering a foundation for further research in this area. The findings have implications for fields ranging from digital media and communication to AI ethics and policy-making, emphasizing the need for increased awareness and understanding of AI capabilities among the general public.

INDUSTRY STUDY: EXPERIENCES, EXPECTATIONS, HOPES AND CHALLENGES OF WORKING WITH AI IN QUALITATIVE RESEARCH.

RUTH ANNA WAKENHUT

KERNWERT, Germany

RELEVANCE & RESEARCH QUESTION

The use of various AI technologies in market research has increased significantly in recent years, and 2023 was a special year: industry publications clearly show that since the beginning of the year, large language models have also been used and new application areas have been tested. These new models are often described as game changers, particularly in qualitative research and analysis. However, there has been little cross-industry sharing of lessons learned. There is a limited understanding of how qualitative researchers use and experience these technologies in their day-to-day work, and how their work may change as a result. Our study aims to fill these gaps by collecting experiences and identifying concerns and challenges. We want to find out what qualitative researchers are actually doing after this year and how the sector has evolved. The aim of our study is to learn what experiences have been gathered so far and what methodological implications, expectations, challenges and opportunities exist.

METHODS & DATA

German-speaking qualitative researchers in the fields of market, social and UX research are invited to take part in the study. The survey consists of open and closed questions to capture different perspectives on the topic and takes approximately 7 minutes to complete. Questions cover experiences, methods, workflows and the real benefits of Al. Participants will answer completely anonymously so that experiences can be shared openly. Invitations will be sent out via newsletters, social media and industry media to reach as wide an audience as possible.

RESULTS

The results are not yet available, but we will be able to present them at the conference. The data will be collected in December and January.

ADDED VALUE

The use of AI poses several challenges for our industry. Sharing experiences is essential to properly assess the potential and develop common standards. We will make the results available to interested parties and communicate them through a variety of channels to encourage a dialogue within the industry.

D3: VIRTUAL RESPONDENTS AND AUDIENCES - IS THIS THE FUTURE OF SURVEY RE-SURVEY RE-SEARCH?

SESSION CHAIR: HOLGER GEISSLER marktforschung.de, Germany

PANELISTS:

DIRK HELD Co-Founder & Managing Director of DECODE Marketing and Co-Founder of Aimpower

LOUISE LEITSCH Director Research of Appinio FRANK BUCKLER Founder & CEO of Success Drivers & Supra Tools FLORIAN KÖGL Founder & CEO of ReDem

SESSION ABSTRACT

Virtual respondents and audiences are on the rise. There are more and more Al tools that generate insights based on human training data or evaluate specific advertising materials. For example, Aimpower's software can be used to evaluate packaging materials and TV commercials. The New Zealand start-up Yabble simulates entire target groups and their reaction to specific questions. The Ad Alliance's SignatureCode tool decodes the DNA of an advertisement. There is also Al-based eye tracking that simulates human perception.

We want to discuss with experts what the future of virtual respondents could look like. How can this approach be used and where are the limitations?

The following people will take part in the discussion:

- Dirk Held, Co-Founder & Managing Director of DECODE Marketing and Co-Founder of Aimpower
- Louise Leitsch, Director Research of Appinio
- Frank Buckler, Founder & CEO of Success Drivers & Supra Tools
- Florian Kögl, Founder & CEO of ReDem

The discussion will be moderated by Holger Geißler, Managing Director and Head of Editorial at marktforschung.de

ABSTRACTS 54

......

A4.1: INNOVA-TION IN INTER-VIEWING & COD-ING

SESSION CHAIR: JESSICA DONZOWA

Max Planck Institut für demographische Forschung, Germany

EXPLORING EFFECTS OF LIFE-LIKE VIRTUAL INTERVIEWERS ON RESPONDENTS' ANSWERS IN A SMARTPHONE SURVEY

JAN KAREM HÖHNE (1,2), FREDERICK G. CONRAD (3), CORNELIA NEUERT (4), JOSHUA CLAASSEN (1)

1: German Center for Higher Education Research and Science Studies (DZHW)

2: Leibniz University Hannover

3: University of Michigan

4: GESIS – Leibniz Institute for the Social Sciences

RELEVANCE & RESEARCH QUESTION

Inexpensive and time-efficient web surveys have increasingly replaced survey interviews, especially conducted in person. Even well-known social surveys, such as the European Social Survey, follow this trend. However, web surveys suffer from low response rates and frequently struggle to assure that the data are of high quality. New advances in communication technology and artificial intelligence make it possible to introduce new approaches to web survey data collection.

Building on these advances, we investigate web surveys in which questions are read through life-like virtual interviewers and in which respondents answer through selecting options from rating scales, incorporating features of in-person interviews in self-administered web surveys.

This has the great potential to improve data quality through the creation of rapport and engagement. We address the following research question: Can we improve data quality in web surveys by programming lifelike virtual interviewers reading questions aloud to respondents?

METHODS & DATA

For this purpose, we are currently conducting a smartphone survey (N ff 2,000) in Germany in which respondents are randomly assigned to virtual interviewers that vary in gender (male or female) and clothing (casual or business casual) or a text-based control interface (without a virtual interviewer). We employ three questions on women's role in the

workplace and several questions for evaluating respondents' experience with the virtual interviewers.

RESULTS

We will examine satisficing behavior (e.g., primacy effects and speeding) and compare respondents' evaluations of the different virtual interviewers. We will also examine the extent to which data quality may be harmed by socially desirable responding when the respondents' gender and clothing preference match those of the virtual interviewer.

ADDED VALUE

By employing life-like virtual interviewers, researchers may be able to deploy web surveys that include the best of interviewer- and self-administered surveys. Thus, our study provides new impulses for improving data quality in web surveys.

API VS. HUMAN CODER: COMPARING THE Performance of speech-to-text Transcription Using Voice Answers From A Smartphone Survey

JAN KAREM HÖHNE (1,2), TIMO LENZNER (3)

1: German Center for Higher Education Research and Science Studies (DZHW)

2: Leibniz University Hannover

3: GESIS – Leibniz Institute for the Social Sciences

RELEVANCE & RESEARCH QUESTION

New advances in information and communication technology, coupled with a steady increase in web survey participation through smartphones, provide new avenues for collecting answers from respondents. Specifically, the built-in microphones of smartphones allow survey researchers and practitioners collecting voice instead of text answers to open-ended questions.

The emergence of automatic speech-to-text APIs transcribing voice answers into text pose a promising and efficient way to make voice answers accessible to text-as-data methods. Even though there are various studies indicating a high transcription performance of speechto text APIs, these studies usually do not consider voice answers from smartphone surveys. We address the following research question: How do transcription APIs perform compared humans?

METHODS & DATA

In this study, we compare the performance of the Google Cloud Speech API and a human coder. We conducted a smartphone survey (N = 501) in the Forsa Omninet Panel in Germany in November 2021 including two open-ended questions with requests for voice answers. These two open questions were implemented to probe two questions from the modules "National Identity" and "Citizenship" of the German questionnaires of the International Social Survey Programme (ISSP) 2013/2014.

RESULTS

The preliminary results indicate that human coder provides more accurate transcriptions than the Google Cloud Speech API. However, the API is much more cost- and time-efficient than the human coder. In what follows, we determine the error rate of the transcriptions for the API and distinguish between no errors, errors that do not affect the interpretabil-

ity of the transcriptions (minor errors), and errors that affect the interpretability of the transcriptions (major errors). We also analyze the data with respect to error types, such as misspellings, word separation error, and word transcription error. Finally, we investigate the association between these transcription error forms and respondent characteristics, such as education and gender.

ADDED VALUE

Our study helps to evaluate the usefulness and usability of automatic speech-to-text transcription in the framework of smartphone surveys and provides empirical-driven guidelines for survey researchers and practitioners.

CAN LIFE-LIKE VIRTUAL INTERVIEWERS INCREASE THE RESPONSE QUALITY OF OPEN-ENDED QUESTIONS?

CORNELIA NEUERT (1), JAN HÖHNE (2), JOSHUA CLAASSEN (2)

1: GESIS – Leibniz Institute for the Social Sciences, Germany 2: DZHW; Leibniz University Hannover

RELEVANCE & RESEARCH QUESTION

Open-ended questions in web surveys suffer from lower data quality compared to in-person interviews, resulting in the risk of not obtaining sufficient information to answer the research question. Emerging innovations in technology and artificial intelligence (AI) make it possible to enhance the survey experience for respondents and to get closer to face-to-face interactions in web surveys. Building on these innovations, we explore the use of life-like virtual interviewers as a design aspect in web surveys that might motivate respondents and thereby improve the quality of the responses.

We investigate the question of whether a virtual interviewer can help to increase the response quality of open-ended questions.

METHODS & DATA

In a between-subjects design, we randomly assign respondents to four virtual interviewers and a control group without an interviewer. The interviewers vary with regard to gender and visual appearance (smart casual vs. business casual). We compare respondents' answers to two open-ended questions embedded in a smartphone web survey with participants of an online access panel in Germany (n=2,000).

RESULTS

The web survey will run in November 2023. After data collection, we analyze responses to the open-ended questions based on various response quality indicators (i.e., probe nonresponse, number of words, number of topics, response times).

ADDED VALUE

The study provides information on the value of implementing virtual interviewers in web surveys to improve respondents experience and data quality, particularly for open-ended questions.

A4.2: DATA QUALITY ASSESSMENTS 1 SPONSORED BY GESIS

SESSION CHAIR: PATRICIA HADLER GESIS – Leibniz Institute for the Social Sciences, Germany

DATA QUALITY IN A LONG AND COMPLEX Online-only survey: The UK genera-Tions and gender survey (ggs)

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

RELEVANCE & RESEARCH QUESTION

Long surveys present high burden for respondents. For a long time the rule of thumb for length of self-completion surveys was not to exceed 15-20 minutes. More surveys are moving towards self-completion designs due to increasing survey costs and also due to high rates of device ownership and internet access in the UK and other countries. For some social surveys 20 minutes is not enough to continue collecting high quality data required. Some studies experimented with longer questionnaires and obtained reassuring results, for example, European Social Survey (ESS). However, more evidence is needed in this under-researched area. We collected the first wave of the UK Generations and Gender Survey (GGS) where only online mode of data collection was available to respondents. The median time spent on the questionnaire is around 40 minutes which is much higher than the advice given to survey practitioners in the past. It is important to assess different aspects of data quality of the long questionnaire. The main research question is: is long questionnaire associated with poorer data quality?

METHODS & DATA

We analyse the GGS collected in the UK. The GGS is a part of a global data collection infrastructure focused on population and family dynamics. The GGS collects demographic, economic, and social attitude data on young and mid-life adults (18-59) as they enter into adulthood, form partnerships, and raise children. We assess different data quality indicators: break-off rate, item nonresponse, different response style behaviours, consent to participation in the second wave of the survey among other indicators. We first conduct descriptive analysis and then use different logistic regression models to investigate data quality in the UK GGS.

RESULTS

The results are reassuring and suggest that even though the GGS questionnaire is long and complex and interviewers are not there to guide the respondents through the process, the data quality is not poor.

ADDED VALUE

This study contributes to the under-researched area of long online questionnaires. This assessment suggests that, when carefully designed, long questionnaires do not represent risk to data quality and can be successfully implemented in self-completion surveys.

SCREENING THE SCREENS: Comparing Sample Profiles And Data Quality Between PC And Mobile Respondents

EVA AIZPURUA (1), GIANMARIA BOTTONI (2)

1: National Centre for Social Research, United Kingdom 2: European Social Survey Headquarters - City, University of London

RELEVANCE & RESEARCH QUESTION

In an era where smartphones have become ubiquitous, the use of mobile devices for survey completion has become prevalent, underscoring the need to understand their impact on data quality. Long gone are instructions for respondents to use alternative devices. While earlier research suggested that mobile device responses resulted in lower data quality, recent studies based on mobile-first survey designs challenge this view, indicating that smartphone usage does not inherently degrade data quality.

METHODS & DATA

This study contributes to this evolving body of research by examining smartphone survey completion in CRONOS-2, the ESS probability-based panel fielded across 12 European countries from November 2021 to March 2023.

RESULTS

We investigate response patterns over time and analyse demographic differences between smartphone and PC respondents, discussing how these insights might be used for targeted respondent recruitment. In addition, we explore survey completion times and data quality indicators, including item non-response and satisficing behaviors, drawing comparisons between PC and smartphone respondents. In this study, we also examine potential differences in break-off rates between smartphone and PC respondents. This approach is informed by previous research, which suggests that smartphones might lead to higher break-offs. Should this be the case, our goal is to identify any problematic items that could lead to such outcomes.

ADDED VALUE

The findings from our research are intended to assist survey researchers and practitioners in the design and execution of high-quality online surveys in a mobile-centric world.

EXPLORING DEVICE DIFFERENCES: ANALYZING SAMPLE COMPOSITION AND DATA QUALITY IN A LARGE-SCALE SURVEY

ALEXANDRA ASIMOV, SARAH THIESEN, MICHAEL BLOHM

GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION

More and more large-scale surveys offer web mode as part of a mixedmode design or as the only mode. Respondents can access the web survey from various devices including desktop computer, tablet, and smartphones. The proportion of respondents using smartphones to complete web surveys is growing. Some studies indicated that the breakoff rate is higher for respondents completing on a smartphone than on other devices. Especially in the context of large-scale social surveys with long completion times, this could lead to more breakoffs at an early stage of the survey. There is also evidence that other data quality indicators vary across devices. Those variations may increase for questions asked later in a survey, as respondents experience a greater burden toward the end of a survey. However, different devices may also address different individuals, which would result in a more balanced sample. We investigate the impact of the growing use of smartphones in large-scale surveys.

METHODS & DATA

We use data from the German General Social Survey (ALLBUS) 2021 and 2023. ALLBUS is a register-based cross-sectional survey of persons aged 18 and older in Germany with a completion time of around 50 minutes. Both surveys were conducted in a self-administered mixed-mode design (mail and web). We compare the sample composition and data quality between different devices. The data quality includes breakoffs, the final question before breakoff, item nonresponse, completion time, and providing an answer to an open question.

RESULTS

The proportion of smartphones used to participate in the survey increases from 5.7% in ALLBUS 2021 to 17.3% in ALLBUS 2023. Preliminary results show that smartphone users restart the survey more often than desktop users. Moreover, breakoff occurs earlier and at a greater rate for smartphone users compared to desktop users. There are significant socio-demographic differences between smartphone and desktop users.

ADDED VALUE

Examining differences in devices improves our understanding of how respondents behave and the potential impact on data quality, especially in the later stages of large-scale surveys.

ABSTRACTS 57

B4: **WILLINGNESS TO PARTICIPATE IN PASSIVE DATA** COLLECTION **STUDIES**

SESSION CHAIR: JOHANNES VOLK

Destatis, Germany

THE INFLUENCE OF CONDITIONAL AND UNCONDITIONAL INCENTIVES ON THE WILLINGNESS TO PARTICIPATE IN WEB TRACKING STUDIES

JUDITH GILSBACH, JOACHIM PIEPENBURG, FRANK MANGOLD, SEBASTIAN STIER, BERND WEISS

GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION

Linking web tracking and survey data opens new research areas. We record participants browsing behavior via a browser plugin. Tracking allows for measuring behavior that individuals tend to recall inaccurately and reduces the survey burden. However, previous studies found that participants were reluctant to participate in tracking studies. To increase participation rates, monetary incentives are widely used. These can be granted unconditionally, conditional on participation or as a combination of both. It is, however, unclear (1) how large conditional incentives should be and whether unconditional incentives can additionally increase participation rates. Additionally, we are interested in, (2) whether these effects are the same for a convenience sample and a probability-based sample.

METHODS & DATA

To answer our research questions, we conduct a 2x3 factorial experiment with approximately 2600 panelists of a new panel. Panelists are recruited via Meta-ads and via a German general population survey (ALLBUS). The first factor is whether panelists receive a prepaid incentive of 5 Euro or not. The second factor is the amount of the postpaid incentive (10,25 or 40 Euro), conditional on 30 out of 60 active days in the tracking period.

We investigate (1a) consent for participation in a web tracking study, (1b) actual installation of the browser plugin. We will present logistic regression models. Additionally, (2) we will investigate the differences between Meta and ALLBUS recruited participants.

RESULTS

Using a smaller dataset of our first field period including only participants recruited via Meta-ads, we find that the unconditional incentive has a positive relation with consent but not with installation. For the amount of the conditional incentive, we cannot see an effect yet. We will analyze a larger dataset for the results that will be presented at the conference. Results for our second research question will be available by the time of the conference.

ADDED VALUE

Little is known on how incentives and other factors impact participation in tracking studies as most research only investigates hypothetical consent. Our study adds to the knowledge on the incentive amount needed to recruit participants via Meta and via a general population survey into a web tracking study.

INTENTIONS VS. REALITY. VALIDATING WILLINGNESS TO PARTICIPATE MEASURES IN VIGNETTE EXPERIMENTS USING RE-AL-WORLD PARTICIPATION DATA

ÁDÁM STEFKOVICS (1,2,3), ZOLTÁN KMETTY (1,4)

1: HUN-REN Centre for Social Sciences, Hungary

2: IQSS, Harvard University

3: Századvég Foundation

4: Eötvös Loránd University

RELEVANCE & RESEARCH QUESTION

Vignette and conjoint experiments are extensively utilised in the field of social sciences. These methodologies assess preferences in hypothetical scenarios, exploring decision-making in complex choices with varied attributes, aiming to align survey responses more closely with real-world decisions. However, survey experiments are only valid externally to the extent that stated intentions align with real-world behaviour. This study uses a unique dataset that allows us to compare the outcomes of a vignette experiment (which assessed willingness to participate in a social media data donation study) to actual participation in a real data donation study involving the same survey respondents.

METHODS & DATA

A vignette experiment embedded in an online survey of a non-probability-based panel was conducted in Hungary in May 2022 (n=1000). Respondents expressed their willingness to participate in hypothetical data donation studies. In a mixed factorial design, five treatment dimensions were varied in the study descriptions (platform, range of data, upload/download time, monetary-, and non-monetary incentive). In February 2023, the same participants were invited to a real data donation study which had almost the same characteristics as the ones described in the vignettes.

RESULTS

The correlation between the self-reported willingness and actual participation was only 0.29. Moreover, the drivers of willingness and actual participation were different. For instance, education was one of the strongest predictors of willingness, yet was not significantly associated with actual participation. We also found differences regarding the effect of privacy beliefs or the Big-Five personality traits.

ADDED VALUE

This study contributed to the literature by validating the results of a vignette experiment using within-person comparisons from behavioural data. The results suggest that vignette experiments may strongly suffer from hypothetical or other biases, at least in scenarios when the personal risk and burden are high, and underscore the importance of improving external validity of such experiments.

WHO IS WILLING TO PARTICIPATE IN AN APP- OR WEB-BASED TRAVEL DIARY STUDY?

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

1: Utrecht University

2: Statistics Netherlands

RELEVANCE & RESEARCH QUESTION

Using apps as a survey mode offers promising features. The use of passive measurements on smartphones can be beneficial for response burden, by replacing traditional survey questions, and for data quality as it can reduce recall bias. However, not everyone is able or willing to participate in an app-based study, causing coverage issues and nonresponse. We investigate whether a mixed-mode design can be effective for our goals by analyzing who chooses to participate in an app study and who prefers a web questionnaire.

METHODS & DATA

We report on a study by Statistics Netherlands (winter 2022-2023) for which we invited 2544 individuals from a cross-sectional sample of the Dutch population. We asked individuals to use a smartphone app to collect their travel data, or participate in a web questionnaire. We combine a concurrent mixed-mode design with a "push-to-app" design by offering the web questionnaire at different moments: directly in the invitation letter or in one of the reminders. Invitees are asked to participate in one mode. We assess whether participation is related to individual characteristics with registry data.

RESULTS

More people register in the app [11.5%] than in the questionnaire [7.0%]. Total registration rates are higher when the web questionnaire is offered directly [19.8%] than in the first [18.5%] or second reminder (15.8%). The app registration rate does not increase much when the web questionnaire is offered later, suggesting that certain people have a mode-preference for the app. Most striking is the age effect. The app attracts younger participants while older participants are overrepresented in the web questionnaire. Combining the two yields a more balanced sample.

ADDED VALUE

We show that with a mixed-mode design, we can attract more respondents than with an app-only design in a probability based sample. With the use of population registries we are able to improve our understanding of who participates in app- and web-studies. Additionally, our analysis can contribute to the design of future diary studies combining a smartphone app and a web questionnaire.

C4: POLITICAL COMMUNICATION AND SOCIAL MEDIA

SESSION CHAIR: JOSEF HARTMANN Verian, Germany

MAPPING NEWS SHARING ON TWITTER: A BOTTOM-UP APPROACH BASED ON NETWORK EMBEDDINGS

FELIX GAISBAUER (1), ARMIN POURNAKI (2,3), JAKOB OHME (1)

1: Weizenbaum-Institut e.V., Germany 2: Max-Planck-Institut für Mathematik in den Naturwissenschaften, Germanu

3: Sciences Po, médialab, Paris, France

RELEVANCE & RESEARCH QUESTION

News sharing on digital platforms is a crucial activity that determines the digital spaces millions of users navigate. Yet, we know little about general patterns of news sharing – previous studies have focused on sharing of misinformation or specific/partisan outlets. To address this gap, we utilize a combination of three data sources to elucidate the extent to which sharing patterns of certain political user groups consist of specific outlets/topics/articles or have unknown diversity. Which types of news are shared in different political regions of Twitter? Are there news that are shared across the political spectrum?

METHODS & DATA

We combine multiple data sources via state-of-the-art network embedding methods and automated text analysis:

- we collected all tweets which contained a link to one of 26 legacy of alternative news outlets for March/2023 (2.5M tweets).
- we crawled the full texts of the articles if available (30K texts); articles were assigned topics with a paragraph-based BERTopic model.
- we collected the follower network of German MPs; we embedded all followers and MPs in a latent political space using correspondence analysis; CA reveals two clearly interpretable dimensions: one shows a clear distinction between AfD and MPs of all other parties; in the other dimension, all parties except AfD are arranged on a left-right axis.

RESULTS

We investigate which types of articles are shared in which political regions of the latent space. We observe interesting, partly counterintuitive sharing patterns: Left-leaning outlets are shared by users in differ-

ent political regions if the topic serves their political cause (qualitative example: an article of Die Zeit critical of working conditions at Deutsche Bahn was shared mostly by users following AfD or CDU/FDP politicians). On the other hand, soft/non-political news seem to be shared only by users in the 'mainstream' political region of the network (example: article on Lena Meyer-Landrut (Bild-Zeitung) with thousands of shares was not shared a single time by AfD followers). We explore these patterns systematically.

ADDED VALUE

We use digital trace data from a broad selection of news outlets. This is one of the first works that combine network embeddings with automated full-text analysis of news.

INDIVIDUAL-LEVEL AND PARTY-LEVEL FACTORS OF GERMAN MPS' GENERAL AND MIGRATION-RELATED POLITICAL COMMUNICATION IN PARLIAMENT AND ON FACEBOOK BETWEEN 2013 AND 2017

PHILIPP DARIUS

Hertie School, Germany

RELEVANCE & RESEARCH QUESTION

Facebook allows for direct communication with voters in the electorates. An issue that is divisive or polarizing on social media and political discourse is migration. This raises the guiding research question, of whether MPs who have positive or negative attitudes toward migration are more likely to speak in parliament on the issue or post about it on Facebook.

METHODS & DATA

This study compares the classical form of political speeches in parliament with social media communication on Facebook by members of parliament of the 18th German Bundestag (2013-2017). While prior studies compared political speech in parliamentary speeches and on social media focused on Twitter messages, this study uses a unique data set linking parliamentary speeches with election data, a candidate survey (GLES), and MPs' social media communication on Facebook. The linked data allows to control for a number of candidate characteristics and test the influence of party or migration-attitudes on speaking and posting behaviour.

The first part of the analysis examines factors associated with general political communication activity in parliament and on Facebook and deploys a generalized linear quasi-Poisson mode, whilst the second part identifies migration-related speeches and posting using a dictionary approach and also analyses the association with candidate characteristics in a quasi-Poisson model.

RESULTS

The first part of the analysis finds that party differences and candidacy play a role in speech activity, whereas being from a 'left-centrist' party (DIE LINKE, SPD, GRÜNE) is positively associated with the number of Facebook messages issued by MPs.

The second part focuses on migration-related communication activity. Against the expectation that MPs with negative migration stances might have used Facebook more intensively to post about migration, the findings indicate that MPs who are in favour of migration were more likely to speak about migration-related issues in parliament and post about it on Facebook.

ABSTRACTS 59

ADDED VALUE

The study uses a unique linked data set combining candidate studies with social media data and parliamentary speech data. The analysis could be improved by using contemporary large language models instead of a dictionary approach and the author would like to discuss added value with fellow conference attendees.

D4: WISSENSCHAFT TRIFFT PRAXIS. WANN IST EINE ONLINE-STICH-PROBE GUT FÜR WELCHEN BEDARF

SESSION CHAIR: OTTO HELLWIG Bilendi & respondi, Germany

IMPULSE VON:

DR. CARINA CORNESSE (GERMAN INSTITUTE FOR ECONOMIC RESEARCH, GERMANY & DGOF VORSTAND) UND MENNO SMID (VORSTANDSVORSIT-ZENDER (CEO) BEI INFAS HOLDING AG)

WEITERE DISKUTANTEN:

BEATE WAIBEL-FLANZ

Business Insights - Market research Manager bei REWE GROUP & stellv. Sprecherin des BVM Regionalrates

DR. BARBARA FELDERER

Teamleiterin Survey Design & Methodology Survey Statistics bei Gesis)

SESSION ABSTRACT

Diese Session widmet sich der Herausforderung, die Güte von Forschungsergebnissen in einer sich ständig wandelnden Welt sicherzustellen. Im Forschungsalltag der Marktforscher fast ausschließlich und



in der Sozialforschung häufig werden Daten aus nicht zufallsbasierten Online Access Paneln gewonnen. Nach wissenschaftlichen Standards verbietet es sich aber, diese Daten auf die Grundgesamtheit zu verallgemeinern. Betreiben die Forscher hier sinnlose Forschung oder wie kann man die Frage nach der Güte einer Stichprobe für spezifische Forschungsbedürfnisse beantworten. In zwei kurzen Impulsen und anschließender Diskussion bringt die Session Wissenschaft und Praxis zusammen fokussiert auf das Konzept von "Fit for Purpose" als eine potenzielle Alternative zur herkömmlichen Repräsentativität und untersucht, inwiefern diese Herangehensweise den dynamischen Anforderungen verschiedener Forschungspraktiken entspricht.

WANN IST EINE STICHPROBE "FIT-FOR-PURPOSE"?

CARINA CORNESSE

German Institute for Economic Research, Germany

In den Medien werden immer wieder Fälle von Studienerkenntnissen diskutiert, die sich bei genauerer Betrachtung als falsch herausstellen. Nicht selten ist der Grund die Datenbasis, welche die proklamierten Schlüsse aufgrund der Selektivität der Stichprobe nicht zulässt. Dies betrifft in besonderem Maße Studien, die von nicht-zufallsbasierten Online-Stichproben auf die deutsche Gesamtbevölkerung schließen wollen. Diese Rückschlüsse beruhen in der Regel auf Annahmen, die weder explizit kommuniziert werden, noch in vielen Fällen haltbar sind. Die Stichproben stellen sich dann also als nicht geeignet für den Zweck heraus, den sie erfüllen sollen und sind somit nicht "fit-for-purpose". Dieser Impulsvortrag beschreibt die Annahmen, die Inferenzschlüssen basierend auf Nicht-Zufallsstichproben zugrunde liegen und diskutiert Umstände, unter denen diese Annahmen halten können. Der Fokus liegt dabei auf der Frage, ob (und wenn ja wann) eine (hoch)selektive nicht-probabilistische Online-Stichprobe fitness-for-purpose für einen bestimmten Forschungszweck aufweisen kann.

FRIDAY, 23/FEB/2024

A5.1: RECRUITING SURVEY PARTICIPANTS

(SPONSORED BY GESIS)

SESSION CHAIR: OLGA MASLOVSKAYA University of Southampton, United Kingdom

RECRUITING ONLINE PANEL THROUGH FACE-TO-FACE AND PUSH-TO-WEB SURVEYS.

BLANKA SZEITL, VERA MESSING, ÁDÁM STEFKOVICS, BENCE SÁGVÁRI HUN-REN Centre for Social Sciences, Hungary

RELEVANCE & RESEARCH QUESTION

This presentation focuses on the difficulties and solutions related to recruiting web panels through probability-based face-to-face and push-toweb surveys. It also compares the panel composition when using two different survey modes for recruitment.

METHODS & DATA

As part of the ESS SUSTAIN-2 project, a webpanel was recruited in 2021/22 through a face-to-face survey of ESS R10 in 12 countries. Unfortunately, the recruitment rate was low and the sample size achieved in Hungary was inadequate for further analysis.

To increase the size of the webpanel (CRONOS-2), the Hungarian team initiated a probability-based mixed-mode self-completion survey (push-to-web design). Respondents were sent a post inviting them to go online or complete a questionnaire, which was identical to the interviewer-assisted ESS R10 survey.

RESULTS

We will present our findings on how the type of survey affects recruitment to a web panel through probability sampling. We will begin by introducing the design of the two surveys, then discuss the challenges encountered in setting up the panel, and finally compare the composition of the panel recruited through the two surveys (interviewer-assisted ESS R10 and push-to-web survey with self-completion).

Our research provides valuable insight into how the type of survey and social and political environment affect recruitment to a web panel.

ADDED VALUE

This analysis focuses on the mode effect on the recruitment of participants for a scientific research panel. Our findings highlight the effect of the social and political environment, which could be used as a source of inspiration for other local studies.

INITIATING CHAIN-REFERRAL FOR VIRTUAL RESPONDENT-DRIVEN SAMPLING – A PILOT STUDY WITH EXPERIMENTS

CARINA CORNESSE (1,2), MARIEL MCKONE LEONARD (3), JULIA WITTON (1), JULIAN AXENFELD (1), JEAN-YVES GERLITZ (2), OLAF GROH-SAM-BERG (2), SABINE ZINN (1)

1: German Institute for Economic Research

2: University of Bremen

3:German Center for Integration and Migration

RELEVANCE & RESEARCH QUESTION

RDS is a network sampling technique for surveying complex populations in the absence of sampling frames. The idea is simple: identify some people ("seeds") who belong or have access to the target population, encourage them to start a survey invitation chain-referral process in their community, ensure that every respondent can be traced back along the referral chain. But who will recruit? And whom? And which strategies help initiate the referral process?

METHODS & DATA

We conducted a pilot study in 2023 where we invited 5,000 panel study members to a multi-topic online survey. During the survey, we asked respondents whether they would be willing to recruit up to three of their network members. If they agreed, we asked them about their relationship with those network members as well as these people's ages, gender, and education and provided unique survey invitation links to be shared virtually. As part of the study, we experimentally varied the RDS consent wording, information layout, and survey link sharing options. We also applied a dual incentive scheme, rewarding seeds as well as recruits.

RESULTS

Overall, 624 initial respondents (27%) were willing to invite network members. They recruited 782 people (i.e., on average 1.25 people per seed). Recruits were mostly invited via email (46%) or WhatsApp (43%) and belonged to the seeds' family (53%) and friends (38%). Only 20% of recruits are in contact with the seed less than once a week, suggesting recruitment mostly among close ties. We find an adequate gender balance (52% female) and representation of people with migration background (22%) in our data, but a high share of people with college or university degrees (52%) and high median age (52 years). The impact of the experimental design on recruitment success is negligible.

ADDED VALUE

While in theory, RDS is a promising procedure, it often fails in practice. Among other challenges, this is commonly due to the fact that seeds will not or only insufficiently start the chain-referral process. Our project shows in which target groups initiating RDS may work and to what extent UX enhancements may increase RDS success.

A5.2: DETECTING UNDESIRABLE RESPONSE BEHAVIOR

(SPONSORED BY GESIS)

SESSION CHAIR: JAN-LUCAS SCHANZE

GESIS – Leibniz Institute for the Social Sciences, Germany

WHO IS GOING BACK AND WHY? USING SURVEY NAVIGATION PARADATA TO DIFFERENTIATE BETWEEN POTENTIAL SATISFICERS AND OPTIMIZERS IN WEB SURVEYS

DANIIL LEBEDEV (1), PETER LUGTIG (2), BELLA STRUMINSKAYA (2)

1: GESIS – Leibniz-Institut für Sozialwissenschaften in Mannheim, Germany

2: Utrecht University, Netherlands

RELEVANCE & RESEARCH QUESTION

Survey navigation paradata presents a unique opportunity to delve into the web survey completion behavior of respondents, particularly actions like revisiting questions and potentially altering answers. Such behavior could be indicative of motivated misreporting, especially when respondents revisit filter or looping questions to modify answers and circumvent subsequent inquiries — a manifestation of satisficing behavior. Conversely, altering answers upon revisiting may also signify optimizing behavior, where respondents strive for utmost accuracy.

This study focuses on the revisiting behavior of web survey respondents, aiming to quantify its frequency, identify associated respondent characteristics, and ascertain who shortens their questionnaire through revisiting.

METHODS & DATA

Using paradata from the probability-based online-administered Generations and Gender Programme (GGP) survey in Estonia (N=8916), we analyze the frequency of revisiting questions, characteristics of these questions, and the ensuing actions. We investigate the connection between revisiting behavior and respondent characteristics using a zero-inflated Poisson regression model and check which respondents' characteristics were connected with a higher proportion of shortening the questionnaire as a result of revisiting questions.

RESULTS

We find a discernible pattern of revisiting questions during the survey, notably prevalent in immediate filter questions, where almost half of respondents go back after a filter question (that can change the routing of the questionnaire).

Contrary to our expectations, the regression analysis did not conclusively support revisiting as a sole indicator of satisficing behavior. The questionnaire size emerged as the most influential factor in revisiting behavior, suggesting that larger questionnaires may burden respondents and potentially lead to motivated misreporting—a form of strong satisficing behavior.

The revisiting observed may reflect respondents' strategies to optimize responses or alleviate survey burden. The complexity of the questionnaire, coupled with respondent motivation and cognitive ability, plays pivotal roles in shaping revisiting behavior, particularly in the case of immediate filters where revisiting may lead to questionnaire shortening.

ADDED VALUE

This study contributes a nuanced understanding of respondents' behavior during web survey self-completion. Utilizing paradata enhances insights into respondents' survey completion patterns and various behavioral types, providing valuable insights for survey design and data quality management.

SOCIALLY DESIRABLE RESPONDING IN PANEL STUDIES – DOES REPEATED INTERVIEWING AFFECT ANSWERS TO SENSITIVE BEHAVIORAL QUESTIONS?

FABIENNE KRAEMER

GESIS – Leibniz Institute for the Social Sciences

RELEVANCE AND RESEARCH QUESTION

Social desirability (SD-) bias (i.e., the tendency to report socially desirable opinions and behaviors instead of revealing true ones) is a widely known threat to response quality and the validity of self-reports. Previous studies investigating socially desirable responding 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 underlying mechanisms of observed changes in SD-bias over time.

METHODS AND DATA

This study investigates socially desirable responding in panel studies using a longitudinal survey experiment comprising six panel waves. The experiment manipulated the frequency of receiving identical sensitive questions (target questions) and assigned respondents to one of three groups: One group received the target questions in each wave (fully conditioned), the second group received the target questions in the last three waves (medium conditioned), and the control group received the target questions only in the last wave of the study (unconditioned). The experiment was conducted within a German non-probability (n = 1,946) and a probability-based panel study (n = 4,660), resulting in 2x3 experimental groups in total. The analysis focusses on between-group and within-group comparisons of different sensitive behavioral measures.

It further includes measures on the questions' degree of sensitivity as a moderating variable. These measures result from an additional survey (n = 237) in which respondents were asked to rate the sensitivity of multiple attitudinal and behavioral questions. To further examine the underlying mechanisms of change, I use a measure on respondents' trust towards the survey (sponsor) and the scores of an established SD-scale.

RESULTS

Results will be presented at the conference in February.

ADDED VALUE

Altogether, this study provides experimental evidence on the impact of repeated interviewing on changes in social desirability bias. It further contributes to the understanding of what causes these changes by examining different levels of exposure to identical sensitive questions and including measures on respondents' trust towards the survey (sponsor) and their scores on a SD-scale.

DISTINGUISHING SATISFICING AND OPTIMISING WEB SURVEY RESPONDENTS USING PARADATA

DANIIL LEBEDEV

GESIS – Leibniz-Institut für Sozialwissenschaften in Mannheim, Germany

RELEVANCE & RESEARCH QUESTION

Web surveys encounter a critical challenge related to measurement error and diminishing data quality, primarily stemming from respondents' engagement in satisficing behavior. Satisficing reflects suboptimal execution of cognitive steps in the answering process.

Paradata, encompassing completion time, mouse movements, and revisiting survey sections, among other metrics, serve to assess respondents' cognitive effort, multitasking tendencies, and motivated misreporting. Despite their individual usage, a comprehensive examination combining various paradata types to discern patterns of satisficing and optimizing behavior has been lacking.

This study seeks to investigate the interplay between different paradata types and data quality indicators derived from survey data, aiming to identify distinct patterns characterizing respondents' satisficing and optimizing behaviors.

METHODS & DATA

Employing a laboratory two-wave experiment with a crossover design involving 93 students, participants were randomly assigned to either satisficing or optimizing conditions in the first wave, with groups reversed in the second. Participants were asked to complete a web survey in either satidficing or in optimising manner. Manipulation checks were used to ensure participants' compliance with a condition.

The survey encompassed open-ended, factual, and matrix questions, coupled with reliable scales gauging trust, values, and other sociological and psychological measures. Paradata, such as completion time, mouse movements, browser focus, reaction to warnings, scrolling, and resizing, were collected using the One Click Survey (1ka.si) online software.

RESULTS

The results revealed that respondents in the optimizing condition exhibited higher data quality compared to those in the satisficing condition, as evidenced by test-retest reliability, completion time, straightlining, and subjective cognitive load. Exploratory factor analysis was employed to scrutinize patterns of advanced paradata values in tandem, shedding light on disparities in survey completion strategies between optimizing and satisficing conditions. The study elucidates the connections between satisficing or optimizing behavior and data quality indicators derived from paradata and survey responses.

ADDED VALUE

This research advances the understanding of satisficing behavior in web surveys by analysing diverse paradata types and uncovering distinctive patterns in respondents' behavior. The findings emphasize the potential of utilizing combined paradata to gain nuanced insights into the survey completion process, thereby enhancing overall data quality.

B5: TO TRACE OR TO DONATE, THAT'S THE QUESTION

SESSION CHAIR: ALEXANDER WENZ University of Mannheim, Germany

EXPLORING THE VIABILITY OF DATA Donations for whatsapp chat logs

JULIAN KOHNE (1,2), CHRISTIAN MONTAG (2)

1: GESIS – Leibniz Institute for the Social Sciences 2: Ulm University

RELEVANCE & RESEARCH QUESTION

Data donations are a new tool for collecting research data. They can ensure informed consent, highly granular, retrospective, and potentially less biased behavioral traces, and are independent from APIs or webscraping pipelines. We thus seek to explore the viability of data donations for a type of highly personal data: WhatsApp chat logs. Specifically, we are exploring a wide range of demographic, psychological, and relational charactersitics and how they relate to peoples donation willingness, censoring, and actual data donation behavior.

METHODS & DATA

We used an opt-in survey assessing demographics, personality, relationship characteristics of a self-selected social relationship, and

concerns for privacy. Participants were also asked whether they are willing to donate a WhatsApp chat from a 1:1 chat from the respective relationship. If they agreed, participants were forwarded to an online platform where they could securely upload, review, self-censor, and donate the chat log. Donated chats were anonimized automatically by first extracting variables of interest (e.g. number of words per message, emoji, smilies, sent domains, response time) and then deleting the raw message content. In a second step, participants selected which parts of the anonymized data should be included in the donations. The study was reviewed and accepted by the ethics committee of Ulm University. So far, 244 people participated in the survey and 140 chat log files with over 1 million messages in total were donated.

PRELIMINARY RESULTS

Preliminary results (based on 198 ppts.) show that participants were mostly university students. Self-indicated willingness to donate a chat was surprisingly high (73%), with a sizable gap to actual donations (39.4%). Interestingly participants rarely excluded any data manually after the automatic anonimization step. Furthermore, we did not find any meaningful differences in data donation willingness and behavior with respect to demographics, personality, privacy concerns, or relationship characteristics.

ADDED VALUE

Our preliminary results highlight, that opt-in data donations can be a viable method to collect even highly sensitive digital trace data if sufficient measures are taken to ensure anonimization, transparancey, and ease-of-use. We will discuss further implications for study design and participant incentivation based on the larger dataset.

THE MIX MAKES THE DIFFERENCE: USING MOBILE SENSING DATA TO FOSTER THE UNDERSTANDING OF NON-COMPLIANCE IN EXPERIENCE SAMPLING STUDIES

RAMONA SCHOEDEL (1,2), THOMAS REITER (2)

1: Charlotte Fresenius Hochschule, University of Psychology, Germany 2: LMU Munich, Department of Psychology

RELEVANCE & RESEARCH QUESTION

For decades, social sciences have focused on broad one-time assessments and neglected the role of momentary experiences and behaviors. Now, novel digital tools facilitate the ambulatory collection of data on a moment-to-moment basis via experience sampling methods. But the compliance to answer short questionnaires in daily life varies considerably between and within participants. Compliance and consequently mechanisms leading to missing data in experience sampling studies, however, still remain in the dark today. In our study we therefore explored person-, context- and behavior-related patterns associated with participants' compliance in experience sampling studies.

METHODS & DATA

We used a data set part (N = 592) of the Smartphone Sensing Panel Study recruited according to quotas representing the German population. We extracted over 400 different person-, context-, and behavior-related variables by combining assessments from traditional surveys (e.g., personality traits), experience sampling (e.g., mood), and passively collected mobile sensing data (e.g., smartphone usage, GPS). Based on more than 25,000 observations, we predicted participants' compliance to answer experience sampling questionnaires. For this purpose, we used a machine learning based modeling approach and benchmarked different classification algorithms using 10-fold cross-validation. In addition, we applied methods from interpretable machine learning to better understand the importance of single variables and constellations of variable groups.

RESULTS

We found that compliance to experience sampling questionnaires could be successfully predicted above chance and that among the compared algorithms the linear elastic net model performed best (MAUC = 0.723). Our follow-up analysis showed that study-related past behaviors such as the average response rate to previous experience sampling questionnaires were the most informative, followed by location information such "at home" or "at work".

ADDED VALUE

Our study shows that compliance in experience sampling studies is related to participants' behavioral and situational context. Accordingly, we illustrate systematic patterns associated with missing data. Our study is an empirical starting point for discussing the design of experience sampling studies in social sciences and for pointing out future directions in research addressing experience sampling methodology and missing data.

C5: POLITICS, MEDIA, TRUST

SESSION CHAIR: FELIX GAISBAUER Weizenbaum-Institut e.V., Germany

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

Konstanz University, 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.

ABSTRACTS 64

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 objective 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 objective 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.

SHARING IS CARING! YOUTH POLITICAL PARTICIPATION IN THE DIGITAL AGE

JULIA SUSANNE WEISS, FRAUKE RIEBE

GESIS, Germany

RELEVANCE & RESEARCH QUESTION

This study addresses a pressing concern in the digital age: the evolution of (online) political participation among young adults. As digital platforms reshape how society engages with politics, traditional definitions and measurements of political involvement require reassessment. The research seeks to unravel the perceived dichotomy between declining conventional political activities and the burgeoning new forms of engagement in digital spaces.

Specifically, our research questions aim to identify and comprehend the spectrum of political participation online and offline among young adults, understand topic-centric engagements, and analyze how participation behaviors differ based on factors like education and digital service utilization. Ultimately, by gauging the behaviors of young adults in the realm of political engagement, this research contributes to both the refinement of existing definitions of political participation and the debate on youth's political engagement trajectory in contemporary settings.

METHODS & DATA

We will conduct an online survey of 16–29-year-olds in December 2023. The respondents for this survey will be recruited via Meta advertisements.

RESULTS

Since the survey will take place in December 2023, nothing can be said about the results at this point. The results will be available in early February 2024.

ADDED VALUE

This study delves into the evolving definitions of political participation and offers methodological insights. It, therefore, explores what can be seen as political participation from the new possibilities digital space offers. Using both closed and open-ended survey questions, we aim to capture a broader spectrum of (online) political participation, potentially filling some gaps in conventional survey techniques.

This approach allows us a more comprehensive understanding of the subject. Additionally, we are working to adjust and propose survey items that reflect current (online) political participation patterns. Through this, our research provides a clearer picture of young adults' political engagement and suggests ways to improve data collection for future research. Finally, our study also provides insight into the extent to which Meta advertisements are suitable for recruiting young people into surveys.

NAVIGATING POLITICAL TURBULENCE: A STUDY OF TRUST AND ONLINE / OFFLINE ENGAGEMENT IN UNSTABLE POLITICAL CONTEXTS

YARON ARIEL, DANA WEIMANN SAKS, VERED ELISHAR

The Max Stern Yezreel Valley College, Israel

RELEVANCE & RESEARCH QUESTION

Within the backdrop of Israel's turbulent 2022 elections, the fifth round of elections within three years, This study delves into the complex interplay between political trust, efficacy, and engagement. It seeks to unravel how individuals' trust in politicians and the political system, coupled with their sense of political efficacy, influences their online and offline engagement in the political process. The research question focuses on identifying the specific predictors of political engagement in a context characterized by political unpredictability and frequent elections.

METHODS & DATA

The study analyzes a representative survey of 530 Israeli respondents during the 2022 Israeli election period. The research evaluates the influence of various variables. These include trust in politicians, the political system, and political efficacy in online and offline political engagement. The analysis focuses on the differentiation between online engagement, such as social media activity, and offline engagement, like attending rallies or voting.

RESULTS

Statistical analysis reveals a robust correlation between political efficacy and both forms of political engagement (r = .62 for online, r = .57 for offline, p < .01). Trust in the political system emerges as a significant predictor of offline engagement (= .36, p < .01), while trust in politicians is more strongly associated with online engagement (= .41, p < .01). Notably, a mediation analysis indicates that political efficacy serves as a mediator in the relationship between trust in politicians and online engagement (indirect effect = 0.15, 95% CI [0.07, 0.24], p < .01). In contrast, such mediating effects between system trust and offline engagement are not observed.

ADDED VALUE

By examining the nuanced factors influencing political engagement during political uncertainty, this study offers new insights into the differentiated impact of trust in politicians and the political system. It underscores the distinct psychological pathways that drive online and offline political engagement, enhancing our understanding of citizen behavior in democracies facing political instability.

These findings have critical implications for political strategists, policymakers, and scholars seeking to foster civic engagement in similar contexts. für die Regulierung von KI schaffen. Doch kann der Al Act tatsächlich zum Innovationsmotor für vertrauenswürdige KI werden oder wird er zum wirtschaftlichen Hemmschuh?

DAS POTENTIAL VON FOUNDATION MODELS UND GENERATIVER KI – EIN BLICK IN DIE ZUKUNFT

SVEN GIESSELBACH IAIS, Germany

A6.1: QUESTIONNAIRE DESIGN CHOICES

SESSION CHAIR: JULIAN B. AXENFELD DIW Berlin, Germany

D5: KI FORUM: IMPULS SESSION – CHANCEN UND REGULIERUNGEN

SESSION MODERATORS: OLIVER TABINO Q Agentur für Forschung YANNICK RIEDER Janssen-Cilag GmbH GEORG WITTENBURG Inspirient

This session is in German.

EU AI ACT: INNOVATIONSMOTOR Oder Innovationsbremse?

ALESSANDRO BLANK

KI Bundesverband, Germany

Der Artificial Intelligence Act (Al Act) der EU ist das erste Regelwerk, das sich mit der Regulierung von Künstlicher Intelligenz (KI) befasst. Mit dem Al Act will die EU einen weltweiten Goldstandard und eine Blaupause

GRID DESIGN IN MIXED DEVICE SURVEYS: AN EXPERIMENT COMPARING FOUR GRID DESIGNS IN A GENERAL DUTCH POPULATION SURVEY.

DEIRDRE GIESEN, MAAIKE KOMPIER, JAN VAN DEN BRAKEL Statistics Netherlands, Netherlands, The

RELEVANCE & RESEARCH QUESTION

Nowadays, designing online surveys means designing for mixed device surveys. One of the challenges in designing mixed device surveys is the presentation of grid questions. In this experiment we compare various design options for grid questions. Our main research questions are: 1) To what extent do these different grid designs differ with respect to response quality and respondent satisfaction? 2) Does this differ for respondents on PCs and respondents on smartphones?

METHODS & DATA

In 2023 an experiment was conducted with a sample of 12060 persons of the general Dutch population aged 16 and older. Sample units were randomly assigned to an online survey in either the standard stylesheet as currently used by Statistics Netherlands (n=2824, 40% of the sample) or an experimental stylesheet (n=7236, 60% of the sample).

Within the current stylesheet, half of the sample units were randomly assigned to the standard grid design as currently used (a table format for large screens and a stem-fixed vertical scrollable format for small screens) and the other half to a general stem-fixed grid design (stemfixed design for both the large and the small screen). Within the experimental stylesheet, one third of the sample was randomly assigned to either the general stem-fix grid design, a carrousel grid design (in which only one item is displayed at the time and after answering one item, the next item automatically 'flies in') or an accordion grid design (all items are presented vertically on one page, and answer options are automatically closed and unfolded after an item is answered).

Various indicators are used to assess response quality, e.g. break-off, item non response, straightlining, mid-point reporting. Respondent satisfaction is assessed with a set of evaluation questions at the end of the questionnaire.

RESULTS

Data are currently being analyzed.

ADDED VALUE

This experiment with a general population sample adds to the knowledge of previous studies on grids. which have mainly been conducted with (access) panels.

TOWARDS A MOBILE WEB QUESTIONNAIRE For the vacation survey: UX design Challenges

VIVIAN MEERTENS, MAAIKE KOMPIER

Statistics Netherlands, Netherlands, The

Key words: Mobile Web Questionnaire Design, Smartphone First Design, Vacation Survey, Statistics Netherlands, UX testing, Qualitative Approach, Mixed Device Surveys

RELEVANCE & RESEARCH QUESTION

Despite the fact that online surveys are not always fit for small screens and mobile device navigation, the number of respondents that start online surveys on mobile devices instead of PC or laptop device, is still growing. Statistics Netherlands (CBS) has responded to this trend by developing and designing mixed device surveys. This study focuses on the redesign of the Vacation Survey, applying a smartphone first approach.

The Vacation Survey is a web only panel survey, that could only be completed on a PC or laptop. The layered design with a master detail approach was formatted in such a way that a large screen was needed to be able to complete the questionnaire. Despite a warning in the invitation letter that a PC or laptop should be used to complete the questionnaire, 14.5% of first-time logins in 2023 were via smartphones, resulting in a redesign with a smartphone first approach. The study examines the applicability and understandability of the Vacation Survey's layered design, specifically its master-detail approach, from a user experience (UX) design perspective.

RESULTS

This study shares key findings of the qualitative UX test conducted at the CBS Userlab. It will explore how visual design aspects influence

respondent behaviour on mobile devices, stressing the importance of observing human interaction when filling in a questionnaire on a mobile phone. The results emphasize the need for thoughtful UX design in mobile web questionnaires to enhanced user engagement and response accuracy.

ABSTRACTS 67

ADDED VALUE

The study provides valuable insights into challenges and implications of transitioning social surveys to mobile devices. By discussing the necessary adaptations for a functional, user-friendly mobile questionnaire, this research contributes to the broader field of survey methodology, offering guidance for future survey designs that accommodate the growing trend of mobile device usage.

OPTIMISING RECALL-BASED TRAVEL DIARIES: LESSONS FROM THE DESIGN OF THE WALES NATIONAL TRAVEL SURVEY

EVA AIZPURUA, PETER CORNICK, SHANE HOWE

National Centre for Social Research, United Kingdom

RELEVANCE & RESEARCH QUESTION

Recall-based travel diaries require respondents to report their travel behaviour over a period ranging from one to seven days. During this period, they are asked to indicate the start and end times and locations, modes of transport, distances, and the number of people on each trip. Depending on the mode, additional questions are asked to gather information on ticket types and costs or fuel types. Due to the specificity of the requested information and its non-centrality for most respondents, travel diaries pose a substantial burden, increasing the risk of satisficing behaviours and trip underreporting.

METHODS & DATA

In this presentation, we describe key decisions made during the design of the Wales National Travel Survey. This push-to-web project includes a questionnaire and a 2-day travel diary programmed into the survey.

RESULTS

Critical aspects of these decisions include the focus of the recall (trip, activity, or location based) and the sequence of follow-up questions (interleaved vs. roster approach). Recent literature suggests that location-based diaries align better with respondents' cognitive processes than trip-based diaries and help reduce underreporting. Therefore, a location-based travel diary was proposed with an auto-complete field to match inputs with known addresses or postcodes. Interactive maps were also proposed for user testing. While they can be particularly useful when respondents have difficulty describing locations or when places lack formal addresses, previous research warns that advanced diary features can increase drop-off rates.

Regarding the follow-up sequence, due to mixed findings in the literature and limited information on the performance of these approaches in web-based travel diaries, experimentation is planned to understand how each approach performs in terms of the accuracy of the filter questions and the follow-up questions. Additionally, this presentation discusses the challenges and options for gathering distance data in recall-based travel diaries, along with learnings from the early phases of diary testing based on the application of a Questionnaire Appraisal System and cognitive/usability interviews.

ABSTRACTS 68

ADDED VALUE

These findings offer valuable insights into the design of complex webbased surveys with multiple loops and non-standard features, extending beyond travel diaries.



SESSION CHAIR: FABIENNE KRAEMER

GESIS — Leibniz Institute for the Social Sciences, Germany

CAN WE IDENTIFY AND PREVENT CHEAT-ING IN ONLINE SURVEYS? EVIDENCE FROM A WEB TRACKING EXPERIMENT.

ORIOL J. BOSCH (1,2,3), MELANIE REVILLA (4)

1: University of Oxford, United Kingdom

2: The London School of Economics, United Kingdom

3: Universitat Pompeu Fabra, Spain

4: Institut Barcelona Estudis Internacionals (IBEI), Spain

RELEVANCE & RESEARCH QUESTION

Survey measures of political knowledge, widely used in political science research, face challenges in online administration due to potential cheating. Previous research reveals a significant proportion of participants resort to online searches when answering political knowledge questions, casting doubt on measurement quality. Existing studies testing potential interventions to curb cheating have relied on indirect measures of cheating, such as catch questions. This study introduces a novel approach, employing direct observations of participants' Internet browsing via web trackers, combined with an experimental design testing two strategies to prevent cheating (instructions and time limit). The paper explores three research questions: what proportion of participants looks up information when posed political knowledge questions (RQ.1)? What is the impact of the interventions on the likelihood of individuals looking up information (RQ.2)? How do estimates from direct observations differ from indirect proxies (e.g., self-reports, paradata) (RQ.3)?

METHODS & DATA

A web survey experiment (N = 1,200) in Spain was deployed within an opt-in access online panel. Cross quotas for age and gender, and quotas for educational level, and region were used to ensure a sample matching on these variables to the Internet adult population. Participants an-

swered six knowledge questions on political facts and current events. Cheating was identified by analysing URLs from web tracking data, and alternative indirect measures were applied, including catch questions, self-reports, and paradata.

RESULTS

Two noteworthy patterns emerge. Firstly, cheating prevalence from web tracking data is below 5%, markedly smaller than levels estimated by indirect measures (2 to 7 times larger). Secondly, based on web tracking data the anti-cheating interventions have no effect. Nonetheless, using indirect measures of cheating we find that both interventions significantly reduce the likelihood of cheating.

ADDED VALUE

This study pioneers the integration of web tracking data and experimental design to examine cheating in online political knowledge assessments. Despite requiring further validation, the substantial differences between web tracking data and indirect approaches suggest two competing conclusion: either cheating in online surveys is substantially lower than first thought, or web tracking data may not be suitable for identifying cheating in online surveys.

THE QUALITY OF SURVEY ITEMS AND THE INTEGRATION OF THE SURVEY QUALITY PREDICTOR 3.0 INTO THE QUESTIONNAIRE DEVELOPMENT PROCESS

LYDIA REPKE

GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION

Designing high-quality survey questions is crucial for reliable and valid research outcomes. However, this process often relies on subjective expertise. In response to this challenge, Saris and colleagues developed the Survey Quality Predictor (SQP), a web-based tool to predict the quality of survey items for continuous latent variables. The research questions driving this presentation are: How can the quality of survey items be predicted? How can SQP 3.0 be effectively integrated into the questionnaire development process?

METHODS & DATA

The quality prediction algorithm (i.e., random forest) of the latest SQP version (3.0) is grounded in a comprehensive analysis involving more than 6,000 survey questions that were part of multitrait-multimethod (MTMM) experiments in 28 languages and 33 countries. The quality prediction of new survey items is based on their linguistic and formal characteristics (e.g., layout and polarity of the answer scale). It is important to note that SQP is not designed to replace traditional methods like cognitive pretesting but serves as a complementary tool in the development phase of questionnaires.

RESULTS

This presentation showcases practical applications of SQP 3.0 in the questionnaire development process. The audience will gain insights into how SQP predicts the quality of survey items. Also, researchers will get to know how they can leverage SQP to identify survey items, enhance item quality before data collection, and detect discrepancies between source and translated versions of survey items.

ADDED VALUE

By incorporating SQP into the questionnaire development toolkit, researchers can enhance the efficiency and objectivity of their survey design processes, ultimately contributing to the advancement of survey research methodologies. In addition, I will highlight the collaborative nature of SQP as an ongoing and evolving research project on survey data quality, emphasizing avenues for potential collaboration among researchers.

PROBABILITY-BASED ONLINE AND MIXED-METHOD PANELS FROM A DATA QUALITY PERSPECTIVE

BLANKA SZEITL (1,2), GERGELY HORZSA (1,2)

1: HUN-REN Centre for Social Sciences, Hungary 2: Panelstory Opinion Polls, Hungary

RELEVANCE & RESEARCH QUESTION

Probability-based online and mixed-method panels are widely used in scientific research, but not as much for market research or political opinion polling. This presentation will explore the case of "Panelstory", the first Hungarian probability-based mixed-method panel, which was established in 2022 with the purpose of utilizing scientific methods to address market research and political opinion polling issues.

METHODS & DATA

We will provide a thorough assessment of panel data based on the total survey error framework to evaluate the quality of indicators such as financial situation, alcohol consumption, interest in politics, health, marital status and media use. Additionally, we will examine the panel composition, response rates, dropout, and recruitment statistics. Non-probability online data collections, face-to-face surveys, and administrative data will be used as reference points. We also relate this to the characteristics of Internet penetration.

RESULTS

The research conducted thus far has revealed that Hungary's Internet penetration rate [82 percent] necessitates a mixed-method design. This is due to the fact that a clear pattern of Internet penetration has been identified in correlation with the indicators being studied. Based on the characteristics of internet penetration in Hungary, in 67 percent of the estimates were biased. For relevant research dimensions such as interest in politics, religiosity, health and marital status, the online data collection significantly under- or overestimates the likely real population proportions.The results of single-mode and mixed-method are notably different in terms of all of the indicators tested.

ADDED VALUE

It is especially important to assess how surveys from probability-based online and mixed-method panels compare to traditional methods such as face-to-face and single-mode designs. This presentation will provide a discussion of a new panel, highlighting both the advantages and potential issues of using scientific results in terms of data quality.

B6.1: AUTOMATIC ANALYSIS OF ANSWERS TO OPEN-ENDED QUESTIONS IN SURVEYS

SESSION CHAIR: BARBARA FELDERER GESIS – Leibniz Institute for the Social Sciences, Germany

USING THE LARGE LANGUAGE MODEL BERT TO CATEGORIZE OPEN-ENDED RESPONSES TO THE "MOST IMPORTANT POLITICAL PROBLEM" IN THE GERMAN LONGITUDINAL ELECTION STUDY (GLES)

JULIA SUSANNE WEISS, JAN MARQUARDT GESIS, Germany

RELEVANCE & RESEARCH QUESTION

Open-ended survey questions are crucial e.g., for capturing unpredictable trends, but the resulting unstructured text data poses challenges. Quantitative usability requires categorization, a labor-intensive process in terms of costs and time, especially with large datasets. In the case of the German Longitudinal Election Study (GLES) spanning from 2018 to 2022, with nearly 400,000 uncoded mentions, it prompted us to explore new ways of coding. Our objective was to test various machine learning approaches to determine the most efficient and cost-effective method for creating a long-term solution for coding responses, ensuring high quality simultaneously. Which approach is best suited for the longterm coding of open-ended mentions regarding the "most important political problem" in the GLES?

METHODS & DATA

Pre-2018, GLES data was manually coded. Shifting to a (partially) automated process involved revising the codebook. Subsequently, the extensive dataset comprising nearly 400,000 open responses to the question regarding the "most important political problem" in the GLES surveys conducted between 2018 and 2022 was employed. The coding process was facilitated using the Large Language Model BERT (Bidirectional Encoder Representations from Transformers). During the entire

..........

process, we tested a whole host of important aspects (hyperparameter finetuning, downsizing of the "other" category, simulations of different amounts of training data, quality control of different survey modes, using training data from 2017) before arriving at the final implementation.

RESULTS

The "new" codebook already demonstrates high quality and consistency, evident from its Fleiss Kappa value of 0.90 for the matching of individual codes. Utilizing this refined codebook as a foundation, 43,000 mentions were manually coded, serving as the training dataset for BERT. The final implementation of coding for the extensive dataset of almost 400,000 mentions using BERT yields excellent results, with a 0/1 loss of 0.069, a Micro F1 score of 0.946 and a Macro F1 score of 0.878.

ADDED VALUE

The outcomes highlight the efficacy of the (partially) automated coding approach, emphasizing accuracy with the refined codebook and BERT's robust performance. This strategic shift towards advanced language models signifies an innovative departure from traditional manual methods, emphasizing efficiency in the coding process.

THE GENESIS OF SYSTEMATIC ANALYSIS METHODS USING AI: AN EXPLORATIVE CASE STUDY

STEPHANIE GAAW, CATHLEEN M. STUETZER, MAZNEV PETKO

TU Dresden, Germany

RELEVANCE & RESEARCH QUESTION

The analysis of open-ended questions in large-scale surveys can provide detailed insights into respondents' views that often can't be assessed with closed-ended questions.

However, due to the large number of respondents, it takes a lot of resources to review the answers within open-ended questions and thus provide them as research results. This contribution aims to show the potential benefits and limitations of using Al-based tools (e.g. ChatGPT), for analyzing open-ended questions in large-scaled surveys. It therefore also aims to highlight the challenge of conducting systematic analysis methods with Al.

METHODS & DATA

As part of a large-scale survey on the use of Al in higher education at a major German university, open-ended questions were included to provide insight into the perceived benefits and challenges for students and lecturers of using Al in higher education. The open-ended responses were then analyzed using a qualitative content analysis. In order to verify whether ChatGPT could be used to analyze the open-ended questions in a faster manner, while maintaining the same quality of results, we asked ChatGPT to analyze the responses in a way similar to our analytical process.

RESULTS

The results show a roadmap of letting ChatGPT analyze our open-ended data. In our case study it obtained categories and descriptions similar to those we obtained by qualitatively analyzing the data ourselves. However, 9 out of 10 times we had to re-prompt ChatGPT to specify the context for the analysis to get the appropriate results. In addition, there were some minor differences in how items were sorted into their respective

categories. Yet, despite these limitations, it became clear that 80% of cases, Chat GPT assigned the responses to the derived categories more accurately than our research team did in the qualitative analysis.

ADDED VALUE

This paper provides insight into how ChatGPT can be used to simplify and accelerate the standard process of qualitative analysis under certain circumstances. We will give insights into our prompts for ChatGPT, detailed findings from comparing its results with our own, and its limitations to contribute to the further development of systematic analysis methods using Al.

INSIGHTS FROM THE HYPERSPHERE – Embedding Analytics in Market Research

LARS SCHMEDEKE, TAMARA KESSLER

SPLENDID Research, Germany

RELEVANCE & RESEARCH QUESTION

In the intersection of qualitative and quantitative research, analyzing open-ended questions remains a significant challenge for data analysts. The incorporation of Al language models introduces the complex embedding space: a realm where semantics intertwine with mathematical principles. This paper explores how Embedding Analytics, a subset of explainable Al, can be utilized to decode and analyze open-ended questions effectively.

METHODS & DATA

Our approach utilized the ada_V2 encoder to transform market research responses into spatial representations on the surface of a 1,536-dimensional hypersphere. This process enabled us to analyze semantic similarities using traditional statistics as well as advanced machine learning techniques. We employed K-Means Clustering for text grouping and respondent segmentation, and Gaussian Mixture Models for overarching topic analysis across numerous responses. Dimensional reduction through t-SNE facilitated the transformation of these complex data sets into more comprehensible 2D or 3D visual representations.

RESULTS

Utilizing OpenAl's ada_V2 encoder, we successfully generated text embeddings that can be plausibly clustered based on semantic content, transcending barriers of language and text length. These clusters, formed via K-Means and Gaussian Mixture Models, effectively yield insightful and automated analyses from qualitative data. The two-dimensional "cognitive constellations" created through t-SNE offer clear and accessible visualizations of intricate knowledge domains, such as brand perception or public opinion.

ADDED VALUE

This methodology allows for a precise numerical analysis of verbatim responses without the need for labor-intensive manual coding. It facilitates automated segmentation, simplification of complex data, and even enables qualitative data to drive prediction tasks. The rich, nuanced datasets derived from semantic complexity are suitable for robust analysis using a wide range of statistical methos, thereby enhancing the efficacy and depth of market rsearch analysis.

B6.2: AI TOOLS FOR SURVEY RESEARCH 2

SESSION CHAIR: FLORIAN KEUSCH University of Mannheim, Germany

VOX POPULI, VOX AI? ESTIMATING German Public Opinion Through Language Models

LEAH VON DER HEYDE (1), ANNA-CAROLINA HAENSCH (1), Alexander Wenz (2)

1: LMU Munich, Germany
 2: University of Mannheim, Germany

RELEVANCE & RESEARCH QUESTION

The recent development of large language models (LLMs) has spurred discussions about whether these models might provide a novel method of collecting public opinion data. As LLMs are trained on large amounts of internet data, potentially reflecting attitudes and behaviors prevalent in the population, LLM-generated "synthetic samples" could complement or replace traditional surveys. Several mostly US-based studies have prompted LLMs to mimic survey respondents, finding that the responses closely match the survey data. However, the prevalence of native-language training data, structural differences between the population reflected therein and the general population, and the relationship between a country's socio-political structure and public opinion, might affect the generalizability of such findings. Therefore, we ask: To what extent can LLMs estimate public opinion in Germany?

METHODS & DATA

We use the example of vote choice as an outcome of interest in public opinion. To generate a "synthetic sample" of the voting-eligible population in Germany, we create personas matching the individual characteristics of the 2017 German Longitudinal Election Study respondents. Prompting GPT-3.5 with each persona, we ask the LLM to predict each respondents' vote choice. We examine how the average party vote shares obtained through GPT-3.5 compare to the survey-based estimates, assess whether GPT-3.5 is able to make accurate estimates for different population subgroups, and compare the determinants of voting behavior between the two data sources.

RESULTS

Based on our prompt design and model configuration, we find that GPT-3.5 does not accurately predict citizens' vote choice, exhibiting a bias towards the Left and Green parties on aggregate, and making better predictions for more "typical" voter subgroups, such as political partisans. Regarding the determinants of its predictions, it tends to miss out on the multifaceted factors that sway individual voter choices.

ADDED VALUE

By examining the prediction of voting behavior using LLMS in a new context, our study contributes to the growing body of research about the conditions under which LLMs can be leveraged for studying public opinion. The findings underscore the limitation of applying LLMs for public opinion estimation without accounting for the biases and potential limitations in their training data.

INTEGRATING LLMS INTO COGNITIVE PRETESTING PROCEDURES: A CASE STUDY USING CHATGPT

TIMO LENZNER, PATRICIA HADLER

GESIS - Leibniz Institute for the Social Sciences, Germany;

RELEVANCE & RESEARCH QUESTION

Since the launch of ChatGPT in November 2022, large language models (LLMs) have been the talk of the town. LLMs are artificial intelligence systems that are trained to understand and generate human language based on huge data sets. In all areas where language data play a central role, they have great potential to become part of a researcher's methodological toolbox. One of these areas is the cognitive pretesting of questionnaires. We identify three tasks where LLMs can augment current cognitive pretesting procedures and potential problems of draft survey questions prior to cognitive testing, (2) suggesting cognitive probes to test draft survey questions, and (3) simulating or predicting respondents' answers to these probes (i.e., generating 'synthetic samples'). In this case study, we examine how well ChatGPT performs these tasks and to what extent it can improve current pretesting procedures.

METHODS & DATA

We conducted a cognitive interviewing study with 24 respondents, testing four versions of a survey question on children's activity levels. Half of the respondents were parents of children aged 3 to 15 years, the other half were adolescents aged 11 to 17 years. In parallel to applying our common pretesting procedures, we prompted ChatGPT 3.5 to perform the three tasks above and analyzed similarities and differences in the outcomes of the LLM and humans.

RESULTS

With respect to tasks (1) and (2), ChatGPT identified some question problems and probes that were not anticipated by humans, but it also missed important problems and probes identified by human experts. With respect to task (3), the answers generated by ChatGPT were characterized by a relatively low variation between individuals with very different characteristics (i.e., gender, age, education) and the reproduction of gender stereotypes regarding the activities of boys and girls. All in all, they only marginally matched the answers of the actual respondents.

ADDED VALUE

To our knowledge, this is one of the first studies examining how LLMs can be incorporated into the toolkit of survey methodologists, particularly in the area of cognitive pretesting.

USING LARGE LANGUAGE MODELS FOR EVALUATING AND IMPROVING SURVEY QUESTIONS

ALEXANDER WENZ (1), ANNA-CAROLINA HAENSCH (2)

1; University of Mannheim, Germany 2: LMU Munich, Germany

RELEVANCE & RESEARCH QUESTION

The recent advances and availability of large language models (LLMs), such as OpenAl's GPT, have created new opportunities for research in the social and behavioral sciences. Questionnaire development and evaluation is a potential area where researchers can benefit from LLMs: Trained on large amounts of text data, LLMs might serve as an easy-to-implement and inexpensive method for both assessing and improving the design of survey questions, by detecting problems in question wordings and suggesting alternative versions. In this paper, we examine to what extent GPT-4 can be leveraged for questionnaire design and evaluation by addressing the following research questions: (1) How accurately can GPT-4 detect problematic linguistic features in survey questions compared to existing computer-based evaluation methods? (2) To what extent can GPT-4 improve the design of survey questions?

METHODS & DATA

We prompt GPT-4 with a set of survey questions and ask to identify features in the question stem or the response options that can potentially cause comprehension problems, such as vague terms or a complex syntax. For each survey question, we also ask the LLM to suggest an improved version. To compare the LLM-based results with an existing computer-based survey evaluation method, we use the Question Understanding Aid (QUAID; Graesser et al. 2006) that rates survey questions on different categories of comprehension problems. Based on an expert review among researchers with a PhD in survey methodology, we assess the accuracy of the GPT-4- and QUAID-based evaluation methods in identifying problematic features in the survey questions. We also ask the expert reviewers to evaluate the quality of the new question versions developed by GPT-4 compared to their original versions.

RESULTS

We compare both evaluation methods with regard to the number of problematic question features identified, building upon the five categories used in QUAID: (1) unfamiliar technical terms, (2) vague or imprecise relative terms, (3) vague or ambiguous noun phrases, (4) complex syntax, and (5) working memory overload.

ADDED VALUE

The results from this paper provide novel evidence on the usefulness of LLMs for facilitating survey data collection.

D6: KI FORUM: KI CAFÉ

SESSION MODERATORS:

OLIVER TABINO Q Agentur für Forschung YANNICK RIEDER Janssen-Cilag GmbH GEORG WITTENBURG Inspirient

This session is in German.

MODERIERTER AUSTAUSCH ZU FOLGENDEN THEMEN:

- Messbare Qualität von KI-Tools ist Grundlage für Vertrauen und Voraussetzung für den betrieblichen Einsatz, aber welche Qualitätskriterien haben sich bewährt? Wie können sie erfasst und verglichen werden?
- Wie implementiert man KI-Anwendungen in Prozesse? Wobei ist die Nutzung bereits etabliert? Was gibt es dabei zu beachten?
- KI und Ethik: Was geht und was nicht?

A7.1: SURVEY METHODS INTER-VENTIONS 2

SESSION CHAIR: JOSS ROSSMANN

GESIS – Leibniz Institute for the S ocial Sciences, Germany

PROMPTS & PRAISES: DYNAMIC FEEDBACK TO REDUCE ROUNDING IN OPEN-ENDED NUMERIC LIST-STYLE QUESTIONS

HANNAH JOSEFINE SCHWÄRZEL, ANKE METZLER, MAREK FUCHS Technische Universität Darmstadt, Germany

RELEVANCE & RESEARCH QUESTION

In Web surveys, open-ended frequency questions are often prone to rounding (also heaping or bunching). Interactive feedback has proven effective in reducing rounding in numeric open-ended questions (Schnell et al., 2022). Particularly, conditional prompts proofed to be

ABSTRACTS 73

more effective in reducing rounding than unconditional prompts. In this study, we examine the role of dynamic feedback (text bubbles) to reduce rounding in numeric list-style open-ended questions. We assess whether appreciative dynamic feedback for precise answers and prompts asking for more precise answers when rounding was identified has the potential to improve data quality.

METHODS & DATA

In 2023, 2,739 respondents from a German online access panel participated in a field-experimental Web survey. Respondents were asked to report amounts spent for either their monthly food expenditures (eight items) or their vacation expenses (six items). Respondents were randomly assigned to either a control group providing no feedback or the experimental group providing two types of dynamic feedback.

A text bubble requesting more precise answers appeared when a reported answer was prone to rounding whereas precise answers triggered an appreciative feedback. An answer was considered rounded if it was a multiple of ten (food expenditures) or a multiple of hundred (vacation expenses).

RESULTS

Respondents exposed to the dynamic feedback provided significantly fewer rounded answers compared to the respondents assigned to the control group. They significantly more often stated precise numbers in every item. Also, the proportion of respondents that answered all items precisely was significantly higher in the experimental group. In addition, the mean number of precise answers was significantly higher in the experimental group than in the control group.

ADDED VALUE

The use of dynamic feedback in numeric list-style open-ended questions has a positive effect on survey responses (less rounding) and the potential to improve data quality of Web surveys. Whether the provision of more precise values (less rounding) reflects more valid responses or is just the consequence of random answering needs further assessment.

PUSHING OLDER TARGET PERSONS TO THE WEB: DO WE STILL NEED A PAPER QUESTIONNAIRE?

JAN-LUCAS SCHANZE, CAROLINE HAHN, OSHRAT HOCHMAN

GESIS – Leibniz-Institut für Sozialwissenschaften, Germany

RELEVANCE & RESEARCH QUESTION

While a sequential, push-to-web mode sequence is very well established in survey research and commonly used in survey practice, many large-scale social surveys still prefer to contact older target persons with a concurrent design, offering a paper questionnaire alongside a web-based questionnaire from the first letter onwards.

In this presentation, we compare the performance of a sequential design with a concurrent design for target persons older than 60 years. We analyse response rates and compare the sample compositions and distributions in key items within resulting net samples. Ultimately, we aim to investigate whether we can push older respondents to the web and whether a paper questionnaire is still required for this age group.

METHODS & DATA

Data stems from the 10th Round of the European Social Survey (ESS) carried out in self-completion modes (CAWI/PAPI) in 2021. In Germany, a mode choice sequence experiment was implemented for all target persons older than 60 years. 50% of this group was invited with a push-to-web approach, offering a paper questionnaire in the third mailing. The control group was invited with a concurrent mode sequence, offering both modes from the beginning on.

RESULTS

Results shows similar response rates for the concurrent design and the sequential design (AAPOR RR2: 38.4% vs. 37.3%). This difference is not statistically significant. In the concurrent group, 21% of the respondents answered the questionnaire online, while in the sequential group this was the case for 50% of all respondents. The resulting net samples are very comparable. Looking at various demographic, socio-economic, attitudinal, and behavioural items, no significant differences were found. In contrast, elderly respondents answering online are younger, more often male, much better educated, economically better off, more politically interested, or more liberal towards immigrants than their peers answering the paper questionnaire.

ADDED VALUE

Online questionnaires are considered as not fully appropriate for surveying the older population. This research shows that a higher share of this group can be pushed to the web without negative effects for response rate or sample composition. However, a paper questionnaire is still required for improving the sample composition.

CLARIFICATION FEATURES IN WEB SURVEYS: USAGE AND IMPACT OF "ON-DEMAND" INSTRUCTIONS

PATRICIA HADLER, TIMO LENZNER, RANJIT K. SINGH, LUKAS SCHICK GESIS – Leibniz Institute for the Social Sciences, Germany;

RELEVANCE & RESEARCH QUESTION

Web surveys offer the possibility to include additional clarifications to a survey question via info buttons that can be placed directly beside a word in the question text or next to the question. Previous research on the use of these clarifications and their impact on survey response is scarce.

METHODS & DATA

Using the non-probability Bilendi panel, we randomly assigned 2,000 respondents to a condition in which they A) were presented clarifications as directly visible instructions under the question texts, B) could click / tip on clarifications via an info button next to the word the respective clarification pertained to, C) could click / tip on clarifications via an info button to the right of the respective question text or D) received no clarifications at all. All questions used an open-ended numeric answer format and respondents were likely to give a smaller number as a response if they read the clarification.

RESULTS

Following the last survey question that contained a clarification, we asked respondents in conditions A) through C) whether they had clicked / tipped on or read the clarification. In addition, we measured the use of the on-demand clarifications using a client-side paradata script. Results

showed that while 24% (B) and 15% (C) of respondents claimed to have clicked on the last-shown on-demand clarification, only 14% (B) and 6% (C) actually did so for at least one question with clarification. Moreover, the responses to the survey question did not differ significantly between the conditions with on-demand instructions (B and C) and the condition with no clarifications (D). Thus, the only way to ensure that respondents adhere to a clarification is to present it as an always visible instruction as in condition A.

ADDED VALUE

The results demonstrate that presenting complex survey questions remains challenging. Even if additional clarification is needed for some respondents only, this clarification should be presented to all respondents; however, with the potential disadvantage of increasing response burden. To learn more about how respondents process clarification features, we are currently carrying out a qualitative follow-up study applying cognitive interviewing.

A7.2: SOCIAL MEDIA RECRUITED SURVEYS

SESSION CHAIR: TOBIAS RETTIG University of Mannheim, Germany

ASSESSING THE IMPACT OF ADVERTISEMENT DESIGN ON RESPONSE QUALITY IN SURVEYS USING SOCIAL MEDIA RECRUITMENT

JESSICA DONZOWA (1,2), SIMON KÜHNE (2), ZAZA ZINDEL (2)

1: Max Planck Institut for Demographic Research, Germany 2: Bielefeld University, Germany

RELEVANCE & RESEARCH QUESTION

Researchers are increasingly using social media platforms for survey recruitment. Typically, advertisements are distributed through these platforms to motivate users to participate in an online survey. To date, there is little empirical evidence on how the content and design characteristics of advertisements can affect response quality in surveys based on social media recruitment. This project is the first comprehensive study of the effects of ad design on response quality in surveys recruited via social media.

METHODS AND DATA

We use data from the SoMeRec survey, which was conducted via Facebook ads in Germany and the United States in June 2023 and focused primarily on climate change and migration. The survey ad campaign featured 15 images with different thematic associations to climate change and migration, including strong and loose associations and neutral images. A commercial access panel company was contracted to include identical survey questions serving as benchmark comparison. The Facebook sample consisted of 7,139 respondents in Germany and 13,022 in the US, while the access panel consisted of 1,555 surveys in Germany and 1,576 surveys in the US. In our analyses, we compare common data quality indicators, including completion time, straightlining, item non-response, and follow-up availability, across different ad features.

RESULTS

First analysis show that survey completion time is higher for thematic ad designs compared to neutral ads and the reference sample. There are differences in the overall item non-response rate, with higher item non-response for the immigration-themed ad designs. There are no significant differences in straightlining between samples and ad designs. Finally, respondents recruited through neutral ads were more likely to be available for follow up surveys than those recruited through themed ads.

ADDED VALUE

Our study advances the literature by studying the general population in Germany and the US, by testing various indicators of survey data quality, and by including a benchmark survey of respondents not recruited through social media. The results clearly indicate an effect of ad design on survey data quality and highlight the importance of sample and recruitment design for estimates based on social media recruitment and online surveys.

DO EXPENSIVE SOCIAL MEDIA AD GROUPS PAY OFF IN THE RECRUITMENT OF A NON-PROBABILISTIC PANEL? AN INSPECTION ON COVERAGE AND COST STRUCTURE

JESSICA DAIKELER, JOACHIM PIEPENBURG, BERND WEISS

GESIS - Leibniz Institute for the Social Sciences, Germany

RELEVANCE & RESEARCH QUESTION

Social media advertisement is becoming an increasingly popular method of recruiting participants for studies in the social sciences. Recently, more and more participants of surveys are recruited via social media. This method of recruitment has been particularly prominent for recruiting special populations for surveys, such as migrants or LGBT persons, but recently meta has significantly reduced these selection criteria.

However, meta still allows the selection of common socio-demographic characteristics, such as age and gender, when placing an ad. Meta estimates these socio-demographic characteristics based on the user's data. With this information, we took an non-probabilistic quota-sampling-like approach by specifying to meta the desired peoples' proportions for socio-demographic characteristics which should click on the ad and be directed to the recruitment survey of our nonprobabilistic panel.

However, the volatile and hard to control nature of social media recruitment opens it up to scrutiny and demands evaluation. In this study we assess coverage issues and cost effectiveness of utilizing Meta advertisement in recruiting respondents for a non-probabilistic online panel, we consider three aspects in detail. First, we evaluate the extent to which the targeting criteria, namely age and gender achieve a balanced sample at different stages of the registration process into the panel and give recommendations for adjustments.

Furthermore, we validate whether these social media targeting criteria are reliable and agree with the survey answers. Third, we assess the cost structure in the light of the response propensities at the different stages of the recruitment process and investigate whether expensive social media ad groups pay off in the long-term.

METHODS & DATA

We are using data from the recruiment of the new GESIS Panel Plus. The recuitmenr process includes several steps and we sill consider each step individually using multivariate analysis methods.

RESULTS

First results suggest that expensive recruitment groups do not pay off in the long term.

ADDED VALUE

These research will open up the black box of cost structure in relation to socio - demographic attributes when using Meta as recruitment frame for cross-sectional and longitudinal surveys.

B7: MOBILE APPS AND SENSORS

SESSION CHAIR: RAMONA SCHOEDEL

Charlotte Fresenius Hochschule, University of Psychology, Germany

MECHANISMS OF PARTICIPATION IN Smartphone app data collection: A research synthesis

WAI TAK TUNG, ALEXANDER WENZ

University of Mannheim

RELEVANCE & RESEARCH QUESTION

Smartphone app data collection has recently gained increasing attention in the social and behavioral sciences, allowing researchers to integrate surveys with sensor data, such as GPS to measure location and movement. Similar to other forms of surveys, participation rates of such studies in general population samples are generally low. Previous research has identified several study- and participant-level determinants of willingness to participate in smartphone app data collection. However, a comprehensive overview of which factors are predictors of willingness and a theoretical framework are currently lacking and some of the effects are inconsistent. To guide future app-based studies, we address the following research questions:

ABSTRACTS 75

(1) Which study- and participant-level characteristics affect the willingness to participate in smartphone app data collection?

(2) Which theoretical frameworks can be used to understand participation decisions in smartphone app data collection?

METHODS & DATA

We conduct a systemic review and a meta-analysis on existing studies with app-based data collection guided by the Preferred Reporting Items for Systematic reviews and Meta-analysis (PRISMA) framework (Moher et al. 2009). We compile a list of keywords to search for relevant literature in bibliographic databases. We focus on peer-reviewed articles published in English. We also perform double coding to ensure a reliable selection of literature for the analysis. Finally, we map the identified determinants of willingness to potential theoretical frameworks that can explain participation behavior.

RESULTS

In the systematic review, we summarize findings about study-level characteristics that are under the researchers' control, such as monetary incentives or invitation mode, and participant-level characteristics, such as privacy concerns and socio-demographics. Meanwhile, the meta-analysis focuses on selected characteristics, which have been most often covered in previous research.

ADDED VALUE

This study will provide a holistic understanding of the current state of research on participation decisions in app-based studies. The findings will also help researchers to design effective invitation strategies for future studies.

"THE VALUE OF PRIVACY IS NOT AS HIGH AS FINDING MY PERSON": SELF-DISCLOSURE PRACTICES ON DATING APPS ILLUSTRATE AN EXISTENTIAL DILEMMA FOR DATA PROTECTION

LUSINE PETROSYAN, GRANT BLANK

University of Oxford, United Kingdom

RELEVANCE & RESEARCH QUESTION

Dating apps create a unique digital sphere where people must disclose sensitive personal information about their demographics, location, values and lifestyle. Because of these intimate disclosures, dating apps constitute a strategic research site to explore how privacy concerns influence personal information disclosure. We use construal-level theory to understand how context influences a decision to disclose. Construal-level theory refers to the influence of psychological distance: the more psychologically distant an event the more mental effort required to understand it. When people have no direct experience in a context they rely on conventional stereotypes and quick generalizations. Using this theory we ask the research question: Why do people choose to disclose or not disclose personal Information on their dating app profile?

METHODS & DATA

We use in-depth, key-informant interviews with 27 active male and female users of the dating site Hinge. Interviews were transcribed and assigned descriptive, process-oriented and interpretative codes using Atlas.ti software.

RESULTS

Dating site users distinguish two kinds of privacy risks. One class of threats is other dating app users who may misuse their information for embarrassment, harassment or stalking, particularly if it could identify the user. These are contexts where users have personal experience. People consider very carefully what information to disclose or hide at the user-level. The second class is the platform-level: app providers who use or sell their information for targeted advertisements. In this context users have no direct experience. Platform-level use is abstract and requires serious mental effort to understand it. Hence it is seen as not threatening and it is ignored. These results confirm construal-level theory.

ADDED VALUE

This research uncovers a previously unnoticed mechanism that governs privacy awareness. It provides clear policy guidelines for enhancing privacy awareness on social media and the Internet in general. Specifically, to encourage people to protect their personal information psychological distance has to be reduced. This can be done by explicit warnings about data use, or explicit statements about data sale and what third parties may do with the information. Warnings should be easily visible on the home page or other prominent locations.

MONEY OR MOTIVATION? DECISION CRITERIA TO PARTICIPATE IN SMART SURVEYS

JOHANNES VOLK, LASSE HÄUFGLÖCKNER

Destatis - Federal Statistical Office Germany, Germany

RELEVANCE & RESEARCH QUESTION

The German Federal Statistical Office (Destatis) is continuing to develop its data collection instruments and is working on smart surveys in this context. By smart surveys we mean the combination of traditional question-based survey data collection and digital trace data collection by accessing device sensor data via an application (GPS, camera, microphone, accelerometer, ...).

Unlike traditional surveys, smart surveys not only ask respondents for information but also require them to download an app and allow access to sensor data. Destatis conducted focus groups to learn more about the attitudes, motives and obstacles regarding the willingness to participate in smart surveys. This was done as part of the European Union's Smart Survey Implementation (SSI) project, in which Destatis is participating alongside other project partners.

METHODS & DATA

Three focus groups with a total of 16 participants were conducted at the end of October 2023. The group discussions were led by a moderator

using a guideline. The discussions lasted around two hours each and were video-recorded.

RESULTS

Overall, it became clear that participants are more willing to take part in a survey, to download an app and to grant access to sensor data if they see a purpose in doing so on the one hand and have trust on the other. In order to motivate people to participate, it seems particularly important against this background to provide transparent information explaining why to conduct the survey, why they should participate, why access to the sensor data is desired as well as what is being done to ensure a high level of data protection and data security.

ADDED VALUE

In official statistics, the development of new survey methods is seen as an important step towards modern data collection. However, modern survey methods can only make a positive contribution if they are used by respondents. The results are intended to provide information on how potential respondents can best be addressed to participate. In the further course of the SSI project, a quantitative field test for recruitment is planned. The results of the focus groups will also be used to prepare this test.