

Bots in web survey interviews: A showcase

Höhne^{1,2}, Claassen^{1,2}, Shahania¹, & Broneske¹

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

² Leibniz University Hannover

Fall Conference of the DGS Method Section on “Computational Sociology”

Bremen (Germany) – October 9 to 11, 2024

This research is funded by the
German Society for Online Research



Introduction I

- **Web surveys struggle with increasingly low response rates** (Daikeler et al. 2020)
- **Social media platforms, such as Facebook and Instagram, offer “sophisticated” advertisement and targeting systems** (Kühne & Zindel 2020; Pötzschke et al. 2023; Zindel 2022)
 - *Quick and easy access to unprecedented and diverse respondent pool*
 - *Supports recruitment of (some) hard-to-reach populations*
- **However, data quality and integrity are potentially threatened by bots** (Griffin et al. 2022; Storozuk et al. 2020; Xu et al. 2022; Yarrish et al. 2019; Zhang et al. 2022)
 - *Programs that autonomously interact with systems, such as web surveys*
 - *Bots may change survey outcomes and thus political and social decision-making* (Xu et al. 2022)
- **Bots were already used to manipulate public opinion through social media**
 - *For example, during Brexit-Referendum in 2016* (Gorodnichenko et al. 2021)

Introduction II

- There is ample literature on how bots infiltrate social media, distribute fake news, and skew public opinion (Howard et al. 2018; Ross et al. 2019; Shi et al. 2020)
- Consequences of bots for web surveys can be severe
 - *Bot-based responses may differ from human responses introducing measurement error*
 - *Bots completing web surveys undermine public trust in social research* (Xu et al. 2022)
 - *Bots can lead to (in-)direct financial damages* (Storozuk et al. 2020; Xu et al. 2022)
- Research on how to prevent bots from infiltrating web surveys is scarce (Griffin et al. 2022; Storozuk et al. 2020; Xu et al. 2022; Yarrish et al. 2019; Zhang et al. 2022)
 - *Methods preventing bots from entering web surveys (e.g., CAPTCHAs)*
 - *Analyzing answer behavior (e.g., open answers)*
 - *Analyzing completion behavior (e.g., response times)*

Limitations and Research Question

- Existing studies have methodological drawbacks
 - *Bot prevalence and consequences are investigated on an observational level*
 - *No probability-based detection models and no simultaneous analysis of multiple indicators*
- No distinction between rule-based and AI-based bots (Naga 2021; Shrivastav 2023)
 - *Most studies only consider rule-based bots*
 - *Existing knowledge about rule-based bots may not hold for AI-based bots*
- AI-based bots might be able to ...
 - *... tackle CAPTCHAs (i.e., challenge-response tests)*
 - *... mimic completion behavior (e.g., mouse movements)*
 - *... respond to question repetitions consistently (e.g., test-retest)*
 - *... respond to questions meaningfully (e.g., open questions)*

?? Do bots varying in sophistication show different completion characteristics ??

Method: Bot Development

- A programmer was asked to program four bots with increasing capabilities
- More sophisticated bots inherit the skills of less sophisticated bots
 - *Cumulative skill sets*

Rule-based bots	AI-based bots
Basic skills bot <ul style="list-style-type: none">+ Randomly answers one question per page (per question type)+ Randomly answers open text fields based on predefined strings+ Tackles invisible honey pot questions	Medium skills bot II (inherits Medium-I bot skills) <ul style="list-style-type: none">+ Classifies web survey content into opinion-based, emails, and attention checks using LLM (Gemini Pro)+ Uses LLM to understand and answer questions meaningfully+ Reads questions and mimics human time delay
Medium skills bot I (inherits Basic bot skills) <ul style="list-style-type: none">+ Handles multiple questions per page and type+ Handles CAPTCHAs with text, objects, or numbers embedded in a picture+ Generates random email addresses with valid domains	Advanced skills bot (inherits Medium-II bot skills) <ul style="list-style-type: none">+ Remembers previous answers (memory)+ Answers based on respondent characteristics (personas)+ Handles questions with audio-visual content (speech-to-text)+ Simulates paradata (mouse movements and clicks, scrolling, and keystrokes)

Method: Bot Showcase

The image shows a browser window on the left and an IPython console on the right. The browser window displays a survey page from DZHW (Deutsches Zentrum für Hochschul- und Wissenschaftsforschung) with the following text:

Nun geht es um die Situation von gleichgeschlechtlichen Paaren in Deutschland.

Seit 2017 können gleichgeschlechtliche Paare in Deutschland heiraten. Dadurch können sie als Ehepaar gemeinsam ein Kind adoptieren.

Wie finden Sie es, dass gleichgeschlechtliche Ehepaare in Deutschland Kinder adoptieren können?

Sehr gut

Eher gut

Eher nicht gut

The IPython console on the right shows the bot's interaction with the survey:

```
Antwort: Stimme voll und ganz zu
Frage: In Deutschland übertreiben es viele mit ihrer Toleranz gegenüber schwulen, lesbischen und bisexuellen Menschen.
Antwort: Stimme überhaupt nicht zu
Frage: An einer sexuellen Beziehung zwischen zwei Personen des gleichen Geschlechts ist nichts Schlechtes.
Antwort: Stimme voll und ganz zu

=====
Selected answer by LLM to radio button: Stimme überhaupt nicht zu
Selected radio button: v_83x4-label
Bot succesfully clicked the radio button
We found a radio_button question on page 3
No visible question found, the whole content of the page is: skip Tivian Survey Questionnaire Umfrage Es gibt in Deutschland unterschiedliche Meinungen zu gleichgeschlechtlichen Partnerschaften. Wie ist das bei Ihnen: Inwieweit stimmen Sie den folgenden Aussagen zu oder nicht zu? Schwule, lesbische und bisexuelle Menschen sollten ihr Leben so führen dürfen, wie sie es wollen. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu Es ist ekelhaft, wenn schwule und lesbische Menschen sich in der Öffentlichkeit küssen. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu Schwule, lesbische und bisexuelle Menschen sollten dieselben Rechte wie heterosexuelle Menschen haben. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu In Deutschland übertreiben es viele mit ihrer Toleranz gegenüber schwulen, lesbischen und bisexuellen Menschen. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu An einer sexuellen Beziehung zwischen zwei Personen des gleichen Geschlechts ist nichts Schlechtes. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu Schwule, lesbische und bisexuelle Menschen sollen aufhören, so einen Wirbel um ihre Sexualität zu machen. Stimme voll und ganz zu Stimme eher zu Stimme eher nicht zu Stimme überhaupt nicht zu Sind Sie ein Bot? Ja Nein Weiter
All radio buttons in this section are not clickable, skipping this section.
Total time spent in delays: 65.04284763336182 seconds
Currently on page 3, clicking Next at time: 2024-10-01 14:56:50.703774
We found a radio_button question on page 4
```

Method: Web Survey Design and Trials

- Web survey on same-gender partnerships was programmed with Unipark
 - *Each of the four bots took the web survey 100 times (N = 400) in August 2024*
 - *Starting with the Advanced bot, we ran the bots one-by-one through the web survey*
- The web survey included ...
 - *... 3 open narrative questions*
 - *... 26 closed questions*
 - *... 1 picture CAPTCHA (counting cars)*
 - *... 2 honey pot questions*
 - *... 1 instructional manipulation check (IMC)*
 - *... 1 check-all-that-apply question (CATA)*
 - *... paradata in the form of completion times*
- The web survey included 43 questions, tasks, and instructions on 28 pages

Method: Web Survey Screenshots I

DZHW
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung

Wir freuen uns, dass Sie die Umfrage gestartet haben. Durch Ihre Teilnahme unterstützen Sie ein Forschungsprojekt zur „Wahrnehmung gleichgeschlechtlicher Partnerschaften in Deutschland“. Das Projekt wird durch das Deutsche Zentrum für Hochschul- und Wissenschaftsforschung (DZHW) durchgeführt und ist durch die Deutsche Gesellschaft für Onlineforschung (DGOF) gefördert. Neben Ihren Antworten auf unsere Fragen erheben wir zusätzliche Prozessdaten (z. B. Antwortzeiten), die es uns erlauben, die Qualität unserer Umfrage besser zu beurteilen.

Die Teilnahme an der Umfrage wird etwa **5 Minuten** in Anspruch nehmen.


Durch Ihre Teilnahme können Sie **5€ gewinnen**. Bitte teilen Sie uns daher am Ende der Umfrage die gültige E-Mail-Adresse Ihres PayPal-Accounts mit, um an der Verlosung teilzunehmen.

Alle Ihre Informationen werden anonymisiert. Das heißt, alle Antworten werden getrennt von Ihren persönlichen Informationen gespeichert und verarbeitet. Es können keine Rückschlüsse auf Sie selbst gezogen werden. Ihre Teilnahme ist selbstverständlich freiwillig.

Das DZHW nimmt den Datenschutz sehr ernst. Wir halten uns an alle gesetzlichen Vorgaben der Datenschutzgrundverordnung (DSGVO) und der entsprechenden Bundes- und Landesdatenschutzgesetze.

Bevor Sie mit der Umfrage starten, möchten wir Sie noch darauf hinweisen, dass die Zunahme an sogenannten Bots (Programme, die menschliche Aktivitäten durchführen) inzwischen auch eine Bedrohung für die Qualität von Umfragen ist. Aus diesem Grund haben wir in dieser Umfrage verschiedene Maßnahmen ergriffen, die uns dabei helfen, Bots und Menschen zu unterscheiden.

Geben Sie bitte die Anzahl an Autos in das offene Feld ein, die Sie im Bild sehen.



Geben Sie bitte die Anzahl der Autos hier ein:

CAPTCHA

DZHW
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung

Nun geht es um die Situation von gleichgeschlechtlichen Paaren in Deutschland.

Seit 2017 können gleichgeschlechtliche Paare in Deutschland heiraten. Dadurch können sie als Ehepaar gemeinsam ein Kind adoptieren.

Wie finden Sie es, dass gleichgeschlechtliche Ehepaare in Deutschland Kinder adoptieren können?

Sehr gut

Eher gut

Eher nicht gut

Überhaupt nicht gut

WEITER

Single question

DZHW
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung

Nun eine Frage zum Thema Diskriminierung. Mit Diskriminierung ist gemeint, dass eine Person oder Gruppe aufgrund von persönlichen Merkmalen schlechter als eine andere Person oder Gruppe behandelt wird.

Inwiefern ist Ihrer Meinung nach die Diskriminierung schwuler, lesbischer und bisexueller Menschen ein Problem oder kein Problem in Deutschland?

Bitte schreiben Sie Ihre Antwort in das offene Feld.

WEITER

Open question

DZHW
Deutsches Zentrum für Hochschul- und Wissenschaftsforschung

Es gibt in Deutschland unterschiedliche Meinungen zu gleichgeschlechtlichen Partnerschaften. Wie ist das bei Ihnen: Inwieweit stimmen Sie den folgenden Aussagen zu oder nicht zu?

Schwule, lesbische und bisexuelle Menschen sollten ihr Leben so führen dürfen, wie sie es wollen.

Stimme voll und ganz zu

Stimme eher zu

Stimme eher nicht zu

Stimme überhaupt nicht zu

Es ist ekelhaft, wenn schwule und lesbische Menschen sich in der Öffentlichkeit küssen.

Stimme voll und ganz zu

Stimme eher zu

Stimme eher nicht zu

Stimme überhaupt nicht zu

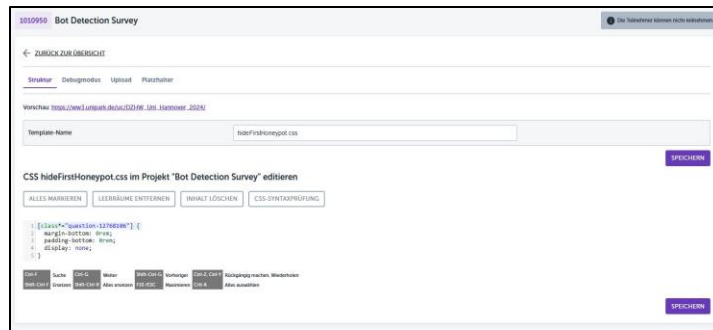
Schwule, lesbische und bisexuelle Menschen sollten dieselben Rechte wie heterosexuelle Menschen haben.

Stimme voll und ganz zu

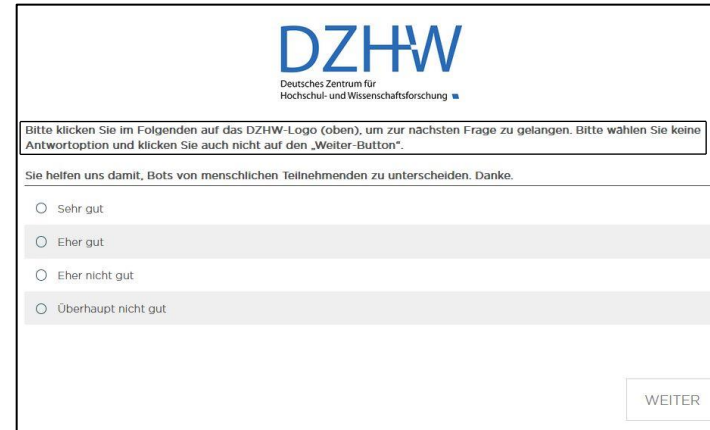
Stimme eher zu

Multiple questions

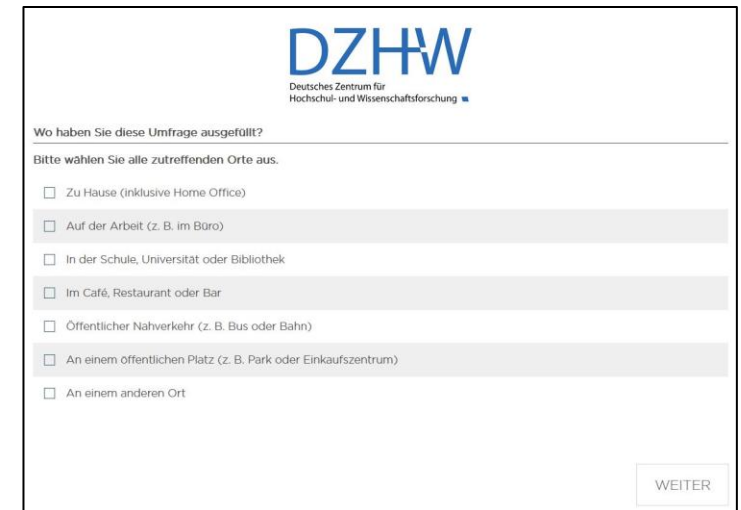
Method: Web Survey Screenshots II



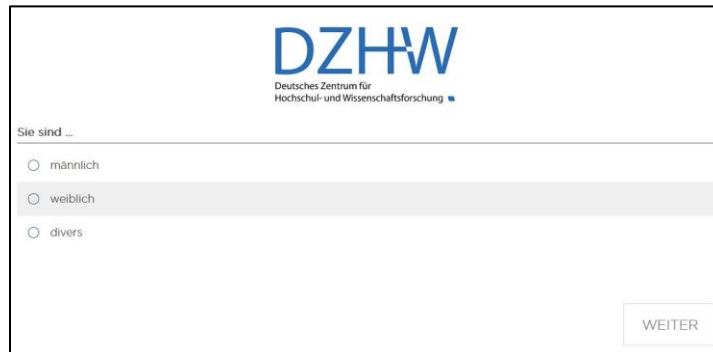
Honey pot



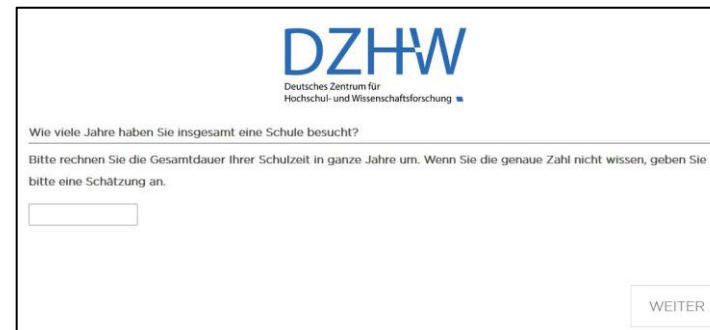
IMC



CATA

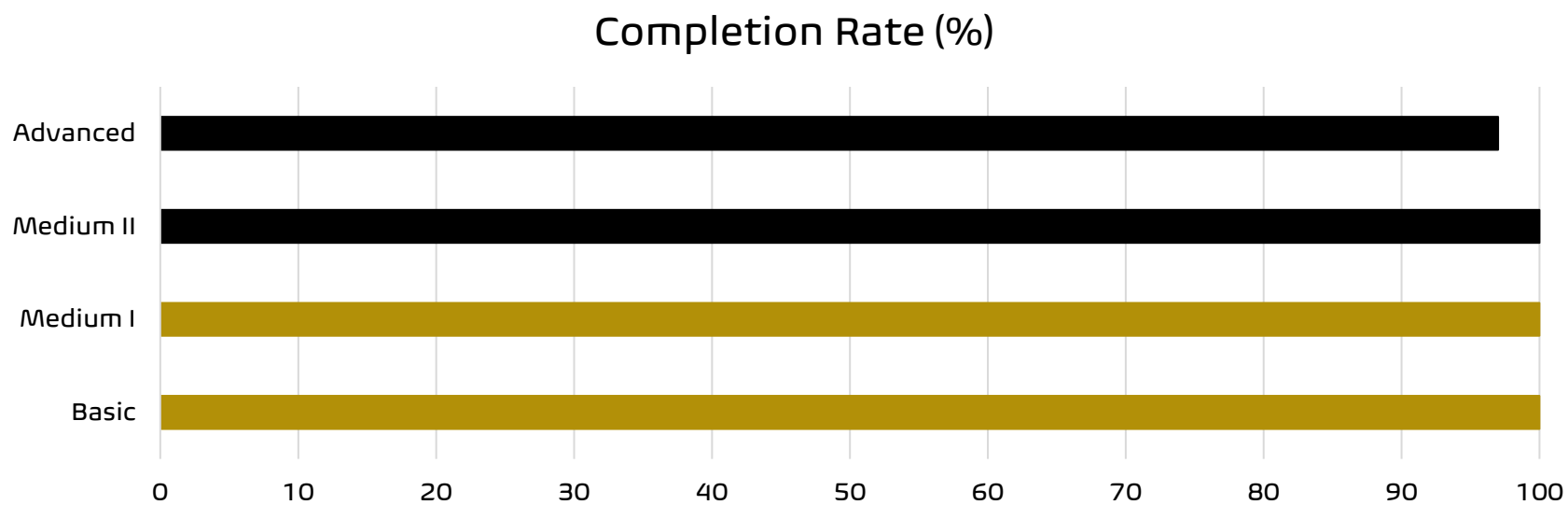


Gender (Demo)



Education (Demo)

Results: Web Survey Completion

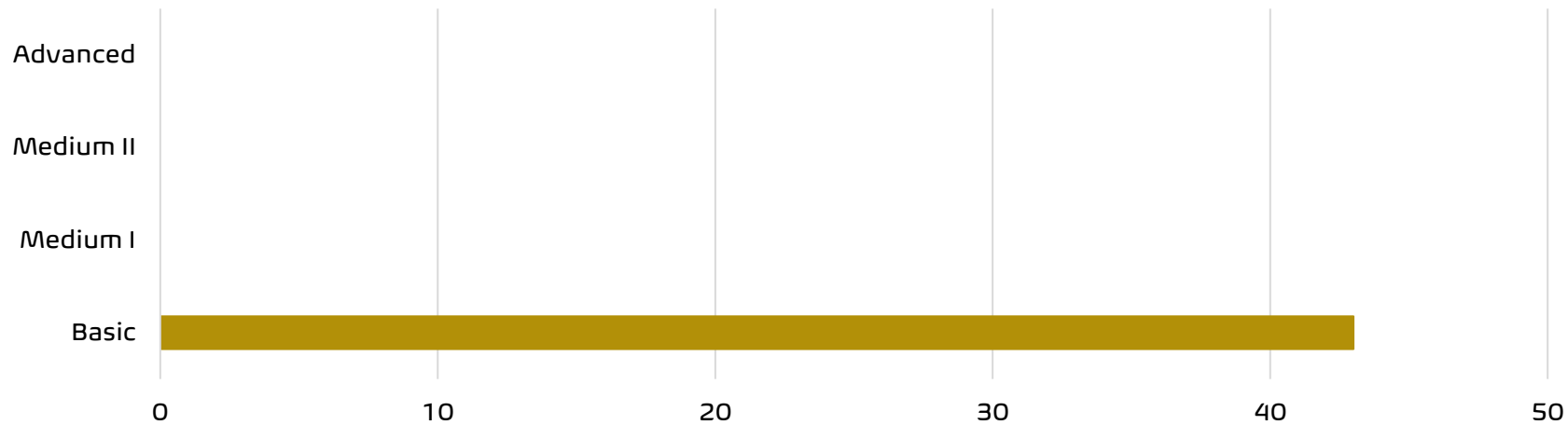


All bots complete the web survey somehow. Break-offs are very limited

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on all 43 questions, tasks, and instructions placed on 28 web survey pages.

Results: Item-nonresponse

Item-nonresponse Rate (%) – Single Questions

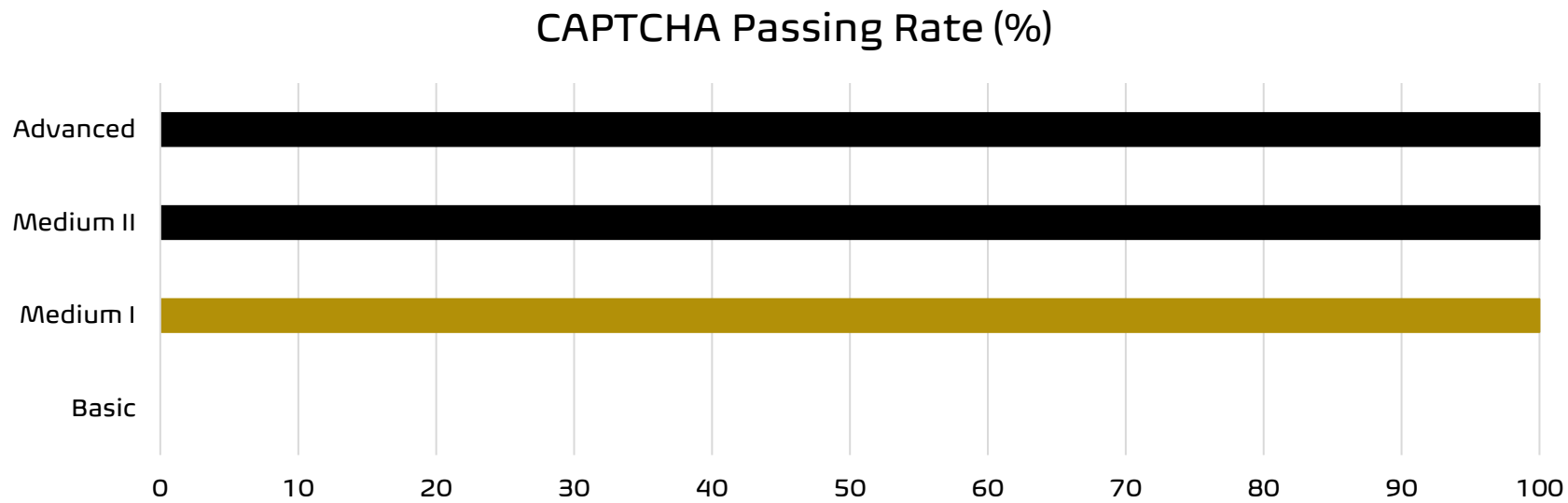


No differences between closed and open questions

Item-nonresponse is comparatively high for the basic bot

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on 26 closed questions and three open narrative questions.

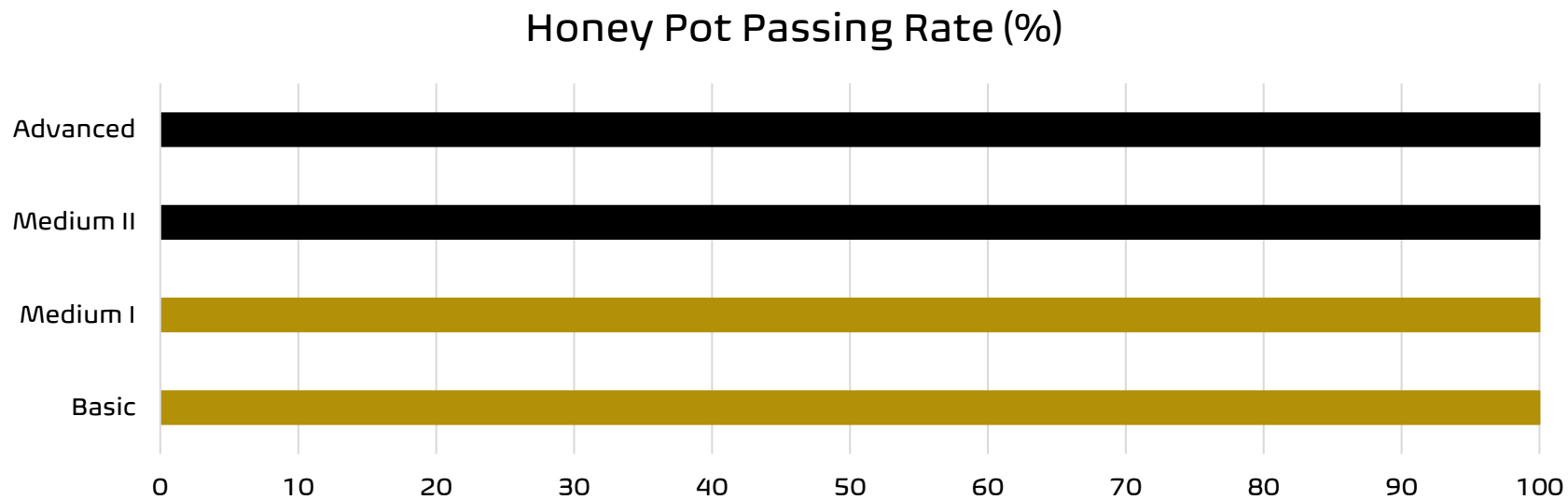
Results: CAPTCHA



As programmed, CAPTCHAs do not constitute a problem for the bots, except for the basic one

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on one CAPTCHA placed on the welcome page.

Results: Honey Pot Questions

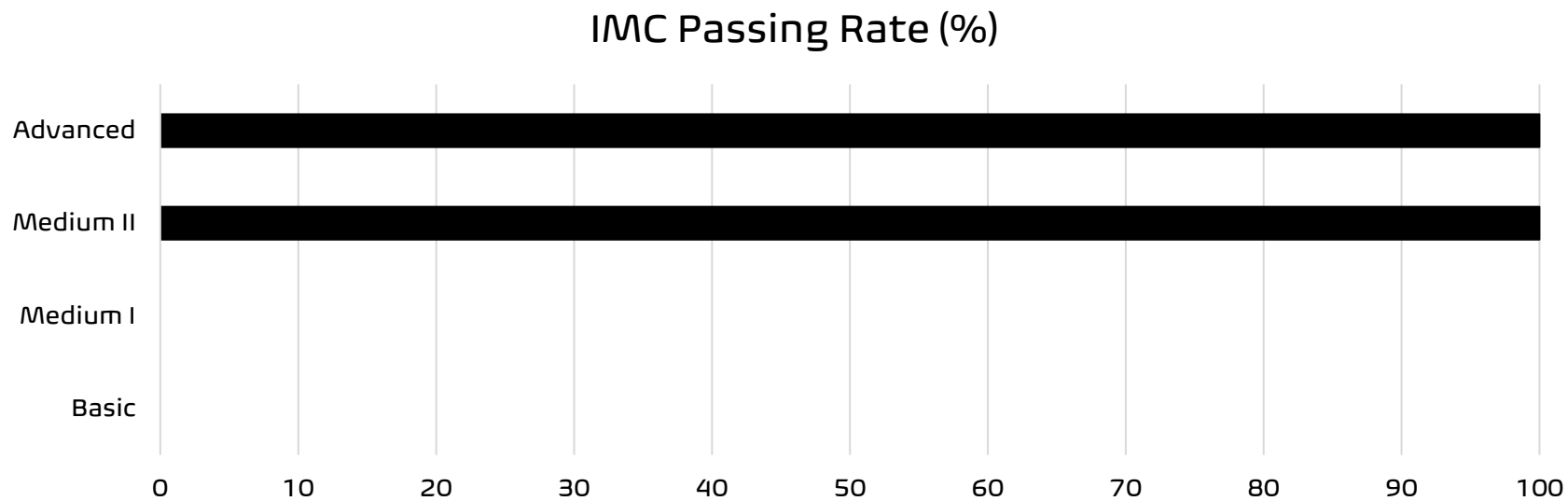


All bots conquer invisible honey pot questions, although they are not explicitly programmed to do so

→ Selenium WebDriver

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on two honey pot questions implemented in the source code of two web survey pages.

Results: IMC

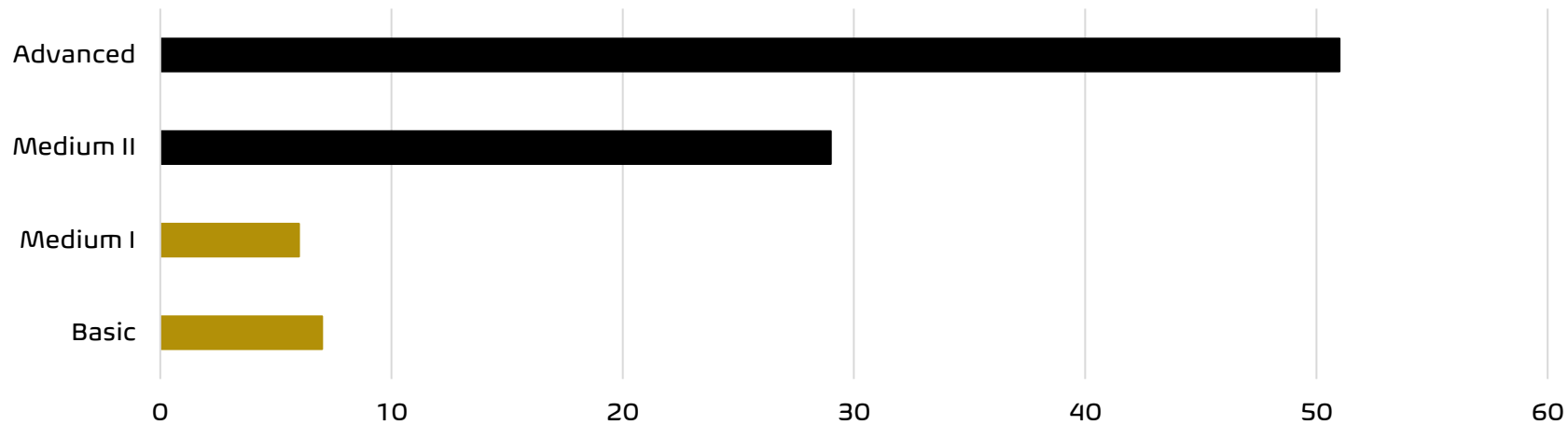


IMCs challenge all bots. However, the advanced bot shows a “learning effect” getting better over trials

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on one IMC placed on one web survey page. IMC = Instructional Manipulation Check.

Results: Answer Length

Answer Length (Words) – Single Open Questions

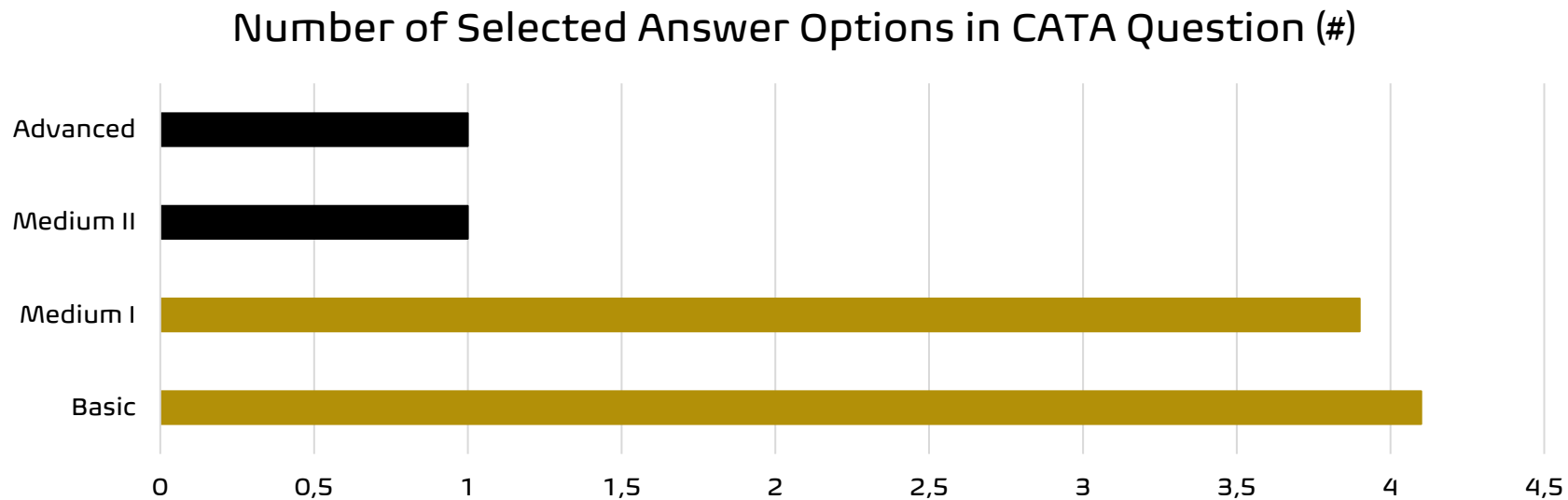


Answer length increases with bot sophistication

It appears that the Advanced bot gets "chatty"

Note. Rule-based bots (Basic and Medium I) and AI-based bots (Medium II and Advanced). Based on three narrative open questions placed on three web survey pages.

Results: CATA Question

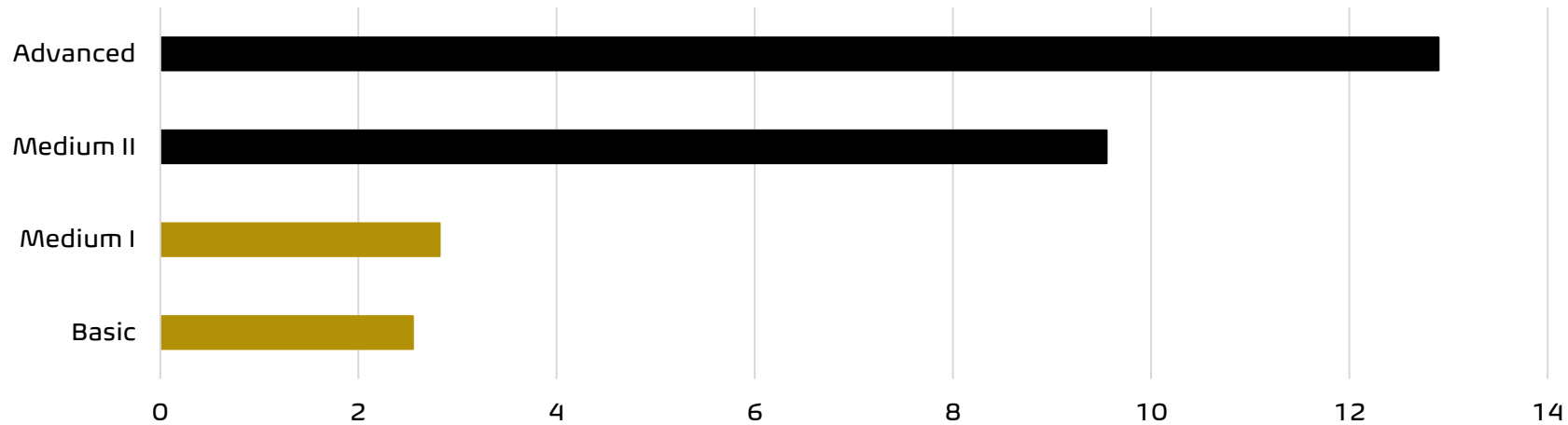


Rule-based bots select a high number of survey locations (e.g., home, public transport, work)

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on one CATA question on survey location placed on one web survey page. CATA = Check-All-That-Apply.

Results: Completion Times

Completion Times in Minutes (min)



There are clear completion time differences between rule- and AI-based bots

Note. Rule-based bots (gold lines) and AI-based bots (black lines). Based on all 43 questions, tasks, and instructions placed on 28 web survey pages. We used the open-source “Embedded Client Side Paradata” tool (Schlosser & Höhne 2018).

Discussion and Conclusion

- Common whisper about bot behavior is only partially true
 - *For example, CAPTCHAs and honey pot questions do not pose a great challenge*
- There are some clear differences between rule- and AI-based bots
 - *AI-based bots provide comparatively long, tailored open answers*
 - *Rule-based bots select a very high (impossible) number of answer options*
 - *AI-based bots produce similar completion times as humans*
- In a next step, we look at completion behavior using paradata
- We then run machine learning algorithms in an unsupervised setting
 - *Extracting features from closed (e.g., inconsistency) and open answers (e.g., Type Token Ratio)*
 - *Applying NLP to detect robotic language for AI-based bots*
- Prediction models plus bot-based pretest application will be accessible soon

CS3 meeting: Saijal Shahania (DZHW)

October 24, 2024, from 3:15 to 4:00 PM

The bot that lived: SurveyBot's role in automated web survey pretesting

In the last decade, web surveys have replaced traditional survey data collection methods, such as in-person interviews, offering substantial cost reduction and paradata collection but presenting complexities due to various question formats and dynamic features. To test the proper implementation of web surveys, researchers and practitioners still rely on costly and time-consuming pretests that are conducted manually by humans. We present SurveyBot, an open-source tool, for automating web survey pretesting through rule-based bots as well as AI-based bots that use Large Language Models (LLMs) and Automatic Speech Recognition (ASR). Tested on two web surveys, SurveyBot effectively handles diverse question types and dynamic features, substantially reducing pretesting costs and human effort.

Many thanks for your attention!

hoehne@dzhw.eu
www.jkhoehne.eu
[@jkhoehne](https://twitter.com/jkhoehne)

Literature I

- Daikeler, J., Bosnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8, 513-539. <https://doi.org/10.1093/jssam/smz008>
- Griffin, M., Martino, R. J., LoSchiavo, C., Comer-Carruthers, C., Krause, K. D., Stults, C. B., & Halkitis, P. N. (2022). Ensuring survey research data integrity in the era of internet bots. *Quality & Quantity*, 56, 2841-2851. <https://doi.org/10.1007/s11135-021-01252-1>
- Gorodnichenko, Y., Pham, T., & Talavera, O. (2021). Social media, sentiment and public opinions: Evidence from #Brexit and #USElection. *European Economic Review*, 136, 103772. <https://doi.org/10.1016/j.euroecorev.2021.103772>
- Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration. *Journal of Information Technology & Politics*, 15, 81-93. <https://doi.org/10.1080/19331681.2018.1448735>
- Kühne, S., & Zindel, Z. (2020). Using Facebook and Instagram to recruit web survey participants: A step-by-step guide and applications. *Survey Methods: Insights from the Field*. <https://surveyinsights.org/?p=13558>
- Naga, K. (2021). How chatbots are enabling a paradigm shift for organisations. <https://www.electronicsforu.com/technology-trends/tech-focus/chatbots-enabling-paradigm-shift-organisations>
- Pöttschke, S., Weiß, B., Daikeler, J., Silber, H., & Beuthner, C. (2023). A guideline on how to recruit respondents for online surveys using Facebook and Instagram: Using hard-to-reach health workers as an example. *GESIS Survey Guidelines*. https://doi.org/10.15465/gesis-sg_en_045
- Ross, B., Pilz, L., Cabrera, B., Brachten, F., Neubaum, G., & Stieglitz, S. (2019). Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. *European Journal of Information Systems*, 28, 394-412. <https://doi.org/10.1080/0960085X.2018.1560920>
- Schlosser, S., & Höhne, J.K. (2018). ECSP – Embedded Client Side Paradata. Zenodo. <https://doi.org/10.5281/zenodo.1218941>
- Shi, W., Liu, D., Yang, J., Zhang, J., Wen, S., & Su, J. (2020). Social bots' sentiment engagement in health emergencies: A topic-based analysis of the covid-19 pandemic discussion on Twitter. *International Journal of Environmental Research and Public Health*, 17, 8701. <https://doi.org/10.3390/ijerph17228701>

Literature II

- Shrivastav, A. (2023). Generative AI chatbots: Gamechanger or doomsayer to intelligent conversations. <https://www.kellton.com/kellton-tech-blog/generative-ai-chatbots-gamechanger-or-doomslayer-to-intelligent-conversations>
- Storozuk, A., Ashley, M., Delage, V., & Maloney, E. A. (2020). Got bots? Practical recommendations to protect online survey data from bot attacks. *The Quantitative Methods for Psychology*, 16, 472-481. <https://doi.org/10.20982/tqmp.16.5.p472>
- Xu, Y., Pace, S., Kim, J., et al. (2022). Threats to online surveys: Recognizing, detecting, and preventing survey bots. *Social Work Research*, 46, 343-350. <https://doi.org/10.1093/swr/svac023>
- Yarrish, C., Groshon, L., Mitchell, J. D., Appelbaum, A., Klock, S., Winternitz, T., & Friedman-Wheeler, D. G. (2019). Finding the signal in the noise: Minimizing responses from bots and inattentive humans in online research. *The Behavior Therapist*, 42, 235-242.
- Zhang, Z., Zhu, S., Mink, J., Xiong, A., Song, L., & Wang, G. (2022). Beyond bot detection: Combating fraudulent online survey takers. In F. Laforest, R. Troncy, E. Simperl, D. Agarwal, A. Gionis, I. Herman, & L. Médini (Eds.), *WWW '22: Proceedings of the ACM Web Conference 2022* (pp. 699-709). Association for Computing Machinery. <https://doi.org/10.1145/3485447.3512230>
- Zindel, Z. (2022). Social media recruitment in online survey research: A systematic literature review. *Methods, Data, Analysis*. <https://doi.org/10.12758/mda.2022.15>