Do People Really Know Their Fertility Intentions? Analyzing Correspondence between Self-Reported Fertility Intentions and Narratives

Xiao Xu[1,2], Anne H. Gauthier[1,2], Gert Stulp[2], and Antal van den Bosch[3]

[1]Netherlands Interdisciplinary Demographic Institute (NIDI)
[2]University of Groningen, The Netherlands
[3]Utrecht University, The Netherlands

Abstract

Fertility intention data from surveys often serve as a crucial component in modeling fertility behaviors. Yet, the persistent gap between stated intentions and actual fertility decisions, coupled with the prevalence of uncertain responses, has cast doubt on the overall utility of intentions and sparked controversies about their origin and nature. In this study, we use survey data from a representative sample of Dutch women. With the help of open-ended questions (OEQs) on fertility and Natural Language Processing (NLP) methods, we are able to conduct an in-depth analysis of fertility narratives. Specifically, we annotated the (expert) perceived fertility intentions of respondents and compared them to their self-reported intentions from the survey. Through this analysis, we aim to reveal the disparities between self-reported intentions and the narratives. Furthermore, by applying neural topic modeling methods, we could uncover which topics and characteristics are more prevalent among respondents who exhibit a significant discrepancy between their stated intentions and their probable future behavior, as reflected in their narratives.

Keywords— fertility intentions, constructive preferences, narrative framework, natural language processing, topic modeling

1 Introduction

Fertility intentions, measured by close-ended questions, are presented in many large-scale demographic surveys (Corsi et al. 2012; Hough and Mayhew 1983; Vikat et al. 2007) and play a key role in fertility research. Responses to fertility intentions have been frequently used by researchers as a predictive factor for whether and when people have children (Liefbroer 2009; Morgan and Rackin 2010; Freitas and Testa 2017; Trinitapoli and Yeatman 2018), but the predictive validity is not always ideal (Morgan and Bachrach 2011; Aiken et al. 2016). Therefore, the utility of measuring fertility intentions in predicting individual (or couples') fertility outcomes has been controversial because intentions often deviate from outcomes (Morgan and Bachrach 2011; Sennott and Yeatman 2018; Müller et al. 2022). In the European context, this is often presented in the format of underachieved fertility or "fertility gap" (Philipov 2009).

As a response, researchers have been asking more fundamental questions on the origin and nature of fertility intentions themselves in recent years. Bhrolcháin and Beaujouan (2019) investigated the prevalent uncertainty in intentions and argued that fertility intentions are constructed over time, which was supported by the empirical study from Müller et al. (2022). Vignoli et al. (2020) proclaimed that fertility decisions are based on a combination of "shadows of the past" and "narratives of the future".

However, while the theoretical advances call for more empirical support, the data collection method in this field has rarely been questioned. Although it's increasingly becoming a consensus in population studies that fertility intentions are contextual and unstable, they are still measured by close-ended questions with pre-defined options in most surveys. Therefore, we still know very little about the "construction" process or factors of fertility intention, nor do we know the effectiveness of the current format of questions.

In this study, we use online open-ended questions (OEQs) in a survey to collect the narratives behind the fertility intentions of Dutch women. We first annotate the responses to OEQ by expert readers, to obtain "our" versions of fertility intentions from a new and neutral perspective. Then, we make a comparison between the two versions of intentions and model the discrepancy from the characteristics of the respondents. Finally, we use Topic Modeling, a Natural Language Processing (NLP) method to automatically analyze the open-ended data, which allows us to interpret large-scale survey qualitative responses. Thus, we expect to identify major narratives on people with different fertility intentions and different levels of discrepancy between self-reported and annotated intentions. With the qualitative insights from topic models, we get access to the narratives that were understudied or overlooked in previous fertility studies.

2 Data

The data used in this study is collected through LISS (Longitudinal Internet Studies for the Social sciences) panel administered by CentERdata at Tilburg University, The Netherlands[1]. The panel covers a representative sample of Dutch individuals who participate in monthly Internet surveys. Only households in which at least one household member spoke Dutch are included.

The study is based on the second wave of the module *Social networks and fertility (in Dutch: Sociale relaties en kinderkeuzes onderzoek)* which was included in the LISS panel in 2021 (as a follow-up to a similar module fielded in 2018). The module's objective was to investigate fertility intentions and attitudes in relation to people's personal networks [2]. For this second wave, 596 female participants were invited and 464 women between the ages of 21 and 44 completed the questionnaire. The survey was conducted in Dutch. Each respondent received €7.50 for completing the questionnaire, which on average takes 15 minutes to finish. This compensation rate is twice as high as the standard rate of LISS panel (∼€2.50 per 10 minutes). The OEQ regarding fertility intentions is presented to respondents who are not currently pregnant ($N = 433$). After removing six answers that were without information (e.g. "niets", nothing, or not in Dutch), there were in total 423 responses available.

2.1 Open-ended question

We used an open-ended question (OEQ) to capture narratives on fertility intentions and uncertainty. Since the usage of online OEQs and automatic analysis of responses is innovative in demographic research, we conducted an experiment on evaluating data quality before this study, which proved that the quality of responses was satisfactory.

The OEQ was placed directly after the standard closed questions on fertility intention from the Generations & Gender Surveys (GGS) (Gauthier, Cabaço, and Emery 2018). The OEQ asked, "Can you tell us more about what makes you (un)certain about whether or not to have children?"[3] The answers contain 31 words on average.

3 Methods

3.1 Annotation

The open-ended responses were annotated by three native Dutch speakers. Two of whom are authors of this paper and experts in the field of the demographic study of fertility intentions, and the third one was a university student without demographic background. The annotators were asked to choose the fertility outcome of the respondent from "definitely yes", "probably yes", "do not know", "definitely no" and "probably no", based on the responses they read. After each annotator finishes the coding task individually, they will meet up to settle the disagreements and finalize the annotation.

3.2 Modeling discrepancy

We assess the effects of respondent characteristics on the discrepancy between the self-reported and annotated fertility intentions, measured by the distance between them. Here, we define the five categories ("definitely yes", "probably yes", "do not know", "definitely no" and "probably no") as number 1 to 5, and use the absolute value of the difference between the two numbers as the distance. For respondent characteristics, we included age and education level as indicators of the ability of respondents, where the education levels were classified into low/medium/high according to the International Standard Classification of Education (ISCED) based on mapping to Dutch education system by the United Nations Educational, Scientific and Cultural Organization (UNESCO). Whether respondents used a PC (N = 235) or mobile devices (smartphone or tablet) (N = 198) will also be included.

[1] `https://www.lissdata.nl/`

[2] See `https://doi.org/10.34894/EZCDOA` for full details on the study, as well as Stulp 2021 and Buijs and Stulp 2022.

[3] Two versions of questions with slight differences were tested in a previous study. Since answers to these two questions were similar on a suite of characteristics (e.g., sentence length, number of nouns), we did not differentiate between these two questions in the subsequent analyses. Original question in Dutch.

3.3 Topic modeling

The advance of natural language processing (NLP) techniques has made large-scale automatic analysis of responses to OEQs possible (Züll 2016; Schonlau and Couper 2016; He and Schonlau 2020). Topic Modeling (Blei and Jordan 2003) is a popular method to extract narrative themes from textual data and has been widely applied in processing responses to OEQs in other disciplines such as social media study and marketing (Roberts, Stewart, Tingley, Airoldi, et al. 2013; Roberts, Stewart, Tingley, Lucas, et al. 2014; Pietsch and Lessmann 2018). The model extracts frequently mentioned topics from the answers of respondents who varied in their certainty of fertility preferences.

In conventional topic modeling algorithms, such as the Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003), each response to an open question is seen as being generated by a distribution of latent topics. In turn, topics are probabilistic distributions of words that co-occur across responses in the data set, and are therefore considered to be generated from a single latent topic. For example, a set of responses that involved discussions on housing concerns may contain terms such as *rent*, *mortgage* and *apartment*. Their common co-occurrences would be recognized by the model and the words would be identified to appear in the same topic with high probability. The results can then be interpreted by researchers and identified as relevant to housing.

A limitation of LDA is that it does not make use of word ordering in the documents. This can be problematic when using short texts such as survey responses (Phan et al. 2010), because they ignore the relation between words and lead to the problem of data sparsity (Rao et al. 2016). In recent years, pre-trained neural language models have been increasingly used in NLP tasks. Instead of representing documents as bags of words (BoW), they treat documents as a continuous sequence of tokens, and produce contextualized representations that contain information on semantic and order relationships among words, which leads them to be increasingly used in topic modeling tasks. Promising novel methods that integrated deep neural networks and embeddings, include Top2Vec (Angelov 2020), SCHOLAR (Card, Tan, and Smith 2017) and Contextualized Topic Models (Bianchi, Terragni, and Hovy 2020), all of which show superior performance on various quantitative metrics compared to LDA (Zhao et al. 2021). In this study, we used Contextualized Topic Models (CTM) in this study. To incorporate prior knowledge on the Dutch language, we use RobBERT (Delobelle, Winters, and Berendt 2020), a Dutch BERT model pre-trained on over 126 million lines of Dutch text.

4 Results

Results will be included in the future version of this paper.

References

Aiken, Abigail RA et al. (2016). "Rethinking the pregnancy planning paradigm: unintended conceptions or unrepresentative concepts?" In: *Perspectives on sexual and reproductive health* 48.3, p. 147.

Angelov, Dimo (2020). "Top2Vec: Distributed Representations of Topics". In: *CoRR* abs/2008.09470. arXiv: 2008.09470. URL: https://arxiv.org/abs/2008.09470.

Bhrolcháin, Máire Ní and Éva Beaujouan (2019). "Do people have reproductive goals? Constructive preferences and the discovery of desired family size". In: *Analytical family demography*. Springer, pp. 27–56.

Bianchi, Federico, Silvia Terragni, and Dirk Hovy (2020). "Pre-training is a hot topic: Contextualized document embeddings improve topic coherence". In: *arXiv preprint arXiv:2004.03974*.

Blei, David M and Michael I Jordan (2003). "Modeling annotated data". In: *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pp. 127–134.

Blei, David M, Andrew Y Ng, and Michael I Jordan (2003). "Latent dirichlet allocation". In: *the Journal of machine Learning research* 3, pp. 993–1022.

Buijs, Vera L and Gert Stulp (2022). "Friends, family, and family friends: Predicting friendships of Dutch women". In: *Social Networks* 70, pp. 25–35.

Card, Dallas, Chenhao Tan, and Noah A Smith (2017). "Neural models for documents with metadata". In: *arXiv preprint arXiv:1705.09296*.

Corsi, Daniel J et al. (2012). "Demographic and health surveys: a profile". In: *International journal of epidemiology* 41.6, pp. 1602–1613.

Delobelle, Pieter, Thomas Winters, and Bettina Berendt (Nov. 2020). "RobBERT: a Dutch RoBERTa-based Language Model". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics, pp. 3255–3265. DOI: 10.18653/v1/2020.findings-emnlp.292. URL: https://www.aclweb.org/anthology/2020.findings-emnlp.292.

Freitas, Rita and Maria Rita Testa (2017). *Fertility desires, intentions and behaviour: A comparative analysis of their consistency*. Tech. rep. Vienna Institute of Demography Working Papers.

Gauthier, Anne H, Susana Laia Farinha Cabaço, and Tom Emery (2018). "Generations and Gender Survey study profile". In: *Longitudinal and Life course studies* 9.4, pp. 456–465.

He, Zhoushanyue and Matthias Schonlau (2020). "Automatic Coding of Open-ended Questions into Multiple Classes: Whether and How to Use Double Coded Data". In: *Survey Research Methods*. Vol. 14. 3, pp. 267–287.

Hough, Mike and Patricia Mayhew (1983). *The British crime survey: first report*. Vol. 76. HM Stationery Office Ixindon.

Liefbroer, Aart C (2009). "Changes in family size intentions across young adulthood: A life-course perspective". In: *European Journal of Population/Revue européenne de Démographie* 25.4, pp. 363–386.

Morgan, S Philip and Christine A Bachrach (2011). "Is the Theory of Planned Behaviour an appropriate model for human fertility?" In: *Vienna yearbook of population research* 9, pp. 11–18.

Morgan, S Philip and Heather Rackin (2010). "The correspondence between fertility intentions and behavior in the United States". In: *Population and development review* 36.1, pp. 91–118.

Müller, Maximilian W et al. (2022). "The illusion of stable fertility preferences". In: *Population Studies* 76.2, pp. 169–189.

Phan, Xuan-Hieu et al. (2010). "A hidden topic-based framework toward building applications with short web documents". In: *IEEE Transactions on Knowledge and Data Engineering* 23.7, pp. 961–976.

Philipov, Dimiter (2009). "Fertility Intentions and Outcomes: The Role of Policies to Close the Gap/Intentions de fécondité et fécondité observée: Rôle des politiques publiques dans la réduction du décalage". In: *European Journal of Population/Revue Europénne de Démographie*, pp. 355–361.

Pietsch, Andra-Selina and Stefan Lessmann (2018). "Topic modeling for analyzing open-ended survey responses". In: *Journal of Business Analytics* 1.2, pp. 93–116.

Rao, Yanghui et al. (2016). "Social emotion classification of short text via topic-level maximum entropy model". In: *Information & Management* 53.8, pp. 978–986.

Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Edoardo M Airoldi, et al. (2013). "The structural topic model and applied social science". In: *Advances in neural information processing systems workshop*

on topic models: computation, application, and evaluation. Vol. 4. Harrahs and Harveys, Lake Tahoe, pp. 1–20.
Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, et al. (2014). "Structural topic models for open-ended survey responses". In: *American journal of political science* 58.4, pp. 1064–1082.
Schonlau, Matthias and Mick P Couper (2016). "Semi-automated categorization of open-ended questions". In: *Survey Research Methods*. Vol. 10. 2, pp. 143–152.
Sennott, Christie and Sara Yeatman (2018). "Conceptualizing childbearing ambivalence: A social and dynamic perspective". In: *Journal of Marriage and Family* 80.4, pp. 888–901.
Stulp, Gert (2021). "Collecting large personal networks in a representative sample of Dutch women". In: *Social Networks* 64, pp. 63–71.
Trinitapoli, Jenny and Sara Yeatman (2018). "The flexibility of fertility preferences in a context of uncertainty". In: *Population and Development Review* 44.1, p. 87.
Vignoli, Daniele et al. (2020). "Uncertainty and narratives of the future: a theoretical framework for contemporary fertility". In: *Analyzing contemporary fertility*. Springer, pp. 25–47.
Vikat, Andres et al. (2007). "Generations and Gender Survey (GGS): Towards a better understanding of relationships and processes in the life course". In: *Demographic Research* 17.14, pp. 389–440. DOI: `10.4054/DemRes.2007.17.14`. eprint: `https://www.demographic-research.org/volumes/vol17/14/17-14.pdf`. URL: `https://www.demographic-research.org/volumes/vol17/14/`.
Zhao, He et al. (2021). "Topic modelling meets deep neural networks: A survey". In: *arXiv preprint arXiv:2103.00498*.
Züll, Cornelia (2016). "Open-ended questions". In: *GESIS Survey Guidelines* 3.