WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

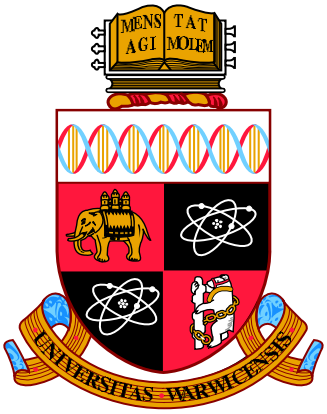
Permanent WRAP URL:

http://wrap.warwick.ac.uk/161247

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications



Essays in Behavioural Economics and Language

Neha Bose

Thesis submitted to the **University of Warwick**
for the degree of **Doctor of Philosophy**

Department of Economics
June, 2021

Contents

List of Tables

List of Figures

Acknowledgements

This thesis would not have been possible without the constant support of my supervisors, Prof Daniel Sgroi, Prof Thomas Hills and Dr Clement Imbert. I cannot begin to express how grateful I am to Daniel for his continued guidance, patience and kind words of encouragement throughout my PhD journey. I am deeply indebted to Thomas for his incredibly helpful feedback and for always inspiring me to do better. I would also like to extend my sincere thanks to Clement for his timely support and advice.

I am grateful to the University of Warwick, especially to the Department of Economics, for giving me this opportunity and for being my home for the better part of a decade. I also want to thank my PhD cohort, fondly (or pretentiously?) referred to as the 'Warwick hipsters', who helped make the journey enjoyable.

Special thanks to my family, especially my parents and grandparents, for their immeasurable support. Many thanks to my sister, Megha, for being the most amazing sister, for always encouraging me and for being my much needed 'comic relief'. Last but not the least, I am extremely grateful to Yash for being my biggest supporter, for always believing in me and cheering me on - the PhD would have been a thousand times more difficult without your support!

I dedicate this thesis to my grandmother who taught me the value of education - Mummum, although you are not here to see me complete this journey, I am sure you would have been the happiest.

Declaration

This thesis is submitted to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy in Economics. I declare that it has not been submitted for a degree at another university. Chapter 1 is co-authored with Prof Daniel Sgroi (University of Warwick) and Chapter 2 is co-authored with Prof Thomas Hills (University of Warwick) and Prof Daniel Sgroi. I am the sole author for Chapter 3. I confirm that all experiments reported in the thesis received ethical approval from research ethics committee and were pre-registered (with further details of ethical approval and pre-registration provided in the chapters).

Neha Bose
June 2021

Abstract

Continuing with the growing trend of language analysis within Economics, this thesis utilises text analysis tools in a novel evaluation of 3 distinct communication settings - small talk, climate change communication and political speeches.

The first chapter studies the implications of *small talk*, the most ubiquitous form of communication, for strategic decision making and the mechanism driving this. In a laboratory setting, after only 4-minutes of small talk, subjects developed impressions or beliefs about their partners' personalities, particularly *extraversion*, which affected behaviour in subsequent strategic interactions. Subjects were more inclined to cooperate in a public goods game when they believed their partners to be extraverted and found it harder to out-reason partners perceived as similar in type to themselves in a level-k reasoning task. Analysing the text used during the small talk chat revealed that talking more is associated with trait extraversion, which indeed provides an accurate forecast of type.

Next, the second chapter assesses climate change communication strategies to reduce planned meat intake. A pre-registered online randomised control trial with 1220 subjects, revealed shockingly low awareness about the environmental impact of meat. An evaluation of 6 information interventions showed that the most effective messages were based on scientific knowledge and *efficacy salience* i.e. concrete information about the consequences of a dietary shift. The study found support for a *targeted messaging approach*, by highlighting the health benefits of a plant-based diet for subjects with health concerns, and evidence of *motivated reasoning* related to meat consumption among frequent meat-eaters. Examining donations to a climate change charity and the text in the information recalled at the end of the study provided further insight into the interventions.

Lastly, the third chapter analyses U.S. Congressional speeches to build linguistic measures of the speakers' attitude towards immigrants. An "immigration corpus" containing 24,351 immigration policy-related speeches between 1990 and 2015 was compiled. The corpus was used to build two distinct measures of attitude towards immigrants, *sentiment* (or valence) and *concreteness*. The measures, particularly sentiment, displayed strict partisan polarisation and variation over time and across states in a manner consistent with the history of immigrant outcomes. A speaker-specific measure of sentiment was a significant predictor of voting behaviour on immigration bills. Applying a Latent Dirichlet Allocation (LDA) topic modelling algorithm displayed trends in the diverse topics discussed in immigration speeches, such as rise in concern over national security post 9/11.

Introduction

Language being a key tool for imparting cultural knowledge and gaining insight into others' minds, can help explain a variety of outcomes including attitude change, attribution, social perception, and intergroup bias and stereotyping (Krauss and Chiu, 1998). Within Economic research, there has been a recent increase in the use of *natural language processing* tools to explain diverse strategic behaviours and build measures of inherent attitudes or bias. For example, language analysis has been used to explain the effect of positive mood on cooperation (Proto, Sgroi, et al., 2019), build a historical measure of national subjective wellbeing (Hills et al., 2019), to measure trends in the partisanship of U.S. Congressional speeches (Gentzkow et al., 2019) and build a measure of gender attitude of judges based on authored opinions (Ash et al., 2021).

Continuing with the trend of incorporating language analysis within the Economics discipline, this thesis employs aspects of text analysis to evaluate behaviour in three distinct communications settings - 1. small talk communication which helps develop personality impressions, which in turn affect strategic behaviour 2. climate change communication interventions to encourage a reduced meat diet and 3. analysis of U.S. Congressional speeches to build measures of inherent attitude towards immigrants. The list of chapters in the thesis, along with co-author names and author contributions, are provided in Table 1.

Table 1: Chapter names and author contributions

No.	Chapter Name	Working paper series	Authors	Author contributions
1	The Role of Personality Beliefs and "Small Talk" in Strategic Behaviour	CAGE Working Paper No. 409	Neha Bose & Daniel Sgroi	NB contributed to experiment design, conducting experiments, data analysis, and manuscript writing and revisions. DS contributed to experiment design and manuscript revisions.
2	Climate Change and Diet	IZA Discussion Paper No. 13426	Neha Bose, Thomas Hills & Daniel Sgroi	NB contributed to experiment design, conducting experiments, data analysis, and manuscript writing and revisions. TH contributed to experiment design and manuscript revisions. DS contributed to experiment design and manuscript revisions.
3	Attitude towards Immigrants: Evidence from U.S. Congressional Speeches	CAGE Working Paper No. 464	Neha Bose	—

While the relationship between personality and strategic decision making has been widely studied in Economics (Proto, Rustichini, and Sofianos, 2019; Proto and Rustichini, 2014; Rustichini et al., 2016; Johnson et al., 2009; Hirsh and Peterson, 2009), what remains unexplored is the impact of *personality impressions* on strategic behaviour. Chapter 1 presents the results from a laboratory experiment, which paired subjects with a stranger with whom they could engage in "small talk" (or *seemingly* non-strategic chatter) for 4 minutes. The chat was conducted through a chat box on their screens, without knowing the nature of any future dealings. Following the chat, subjects developed beliefs about each others' personalities, particularly trait extraversion. These beliefs affected decision making in two well-known strategic games - the *Public Goods Game*, which is a game of cooperation, and the *11-20 money request game* (Arad and Rubinstein, 2012), which examines level-k reasoning. Note that, extraversion, as a fundamental and pervasive personality trait (Costa and McCrae, 1980), was the only trait which could be reasonably detected after a brief bout of communication.

Having engaged in small talk, subjects were observed to cooperate more in the Public Goods Game when they believed their partner to be extraverted and found it harder to out-reason opponents who they perceived as similar in type to themselves. Analysis of the language used during the small talk communication revealed that the more talkative partners were believed to be extraverted. However, with regards to personality beliefs, there remained a persistent own-type bias as extraverts, prone to *complementary self projection bias*, are likely to overstate the extravertion in others.

Aside from the novel use of personality beliefs to explain strategic behaviour, this chapter contributes to the literature exploring accuracy of *personality attribution* (Eaton and Funder, 2003; Little and Perrett, 2007; Naumann et al., 2009). This study also contributes to the modest literature on implications of small talk (Pullin, 2010; Das and Chen, 2007; Ragan, 2014). Lastly, the study adds to literature on *strategic sophistication* (Fe et al., 2019; Gill and Prowse, 2016; Georganas et al., 2015) by showing that subjects adjust strategies based on *endogenous* belief formation about the opponent's type.

Next, Chapter 2 addresses climate change communication strategies. Despite scientific consensus regarding the environmental impact of meat (Schiermeier, 2019; Tilman and Clark, 2015; Stehfest et al., 2009), there is shockingly low awareness among the general public (Sanchez-Sabate and Sabaté, 2019; de Boer et al., 2016; Macdiarmid et al., 2016; Vanhonacker et al., 2013). Interventions aiming to increase public awareness about the environmental benefits of a reduced

meat diet have been scattered and inconclusive, with most studies focussing on a single behavioural barrier (see Harguess et al., 2020 for a systematic literature review). This study makes a novel attempt to evaluate the effectiveness of 6 information interventions, framed using 6 supporting theories, against a control group with baseline scientific information, in reducing planned meat intake. The list of 6 interventions was compiled to address all major barriers to reducing meat consumption: scientific knowledge gap, perceived inefficacy of action (with 2 different efficacy frames), lack of personal relevance, cognitive dissonance and social norms.

In an online pre-registered randomised controlled trial, involving 1220 participants, only 9% reported diet-related actions as effective against climate change. Compared to the control group, the most effective interventions were based on additional scientific knowledge and *efficacy salience* i.e. providing information about a reduced meat diet in an easily understandable unit. The study also found support for a *targeted messaging approach* as providing information about the health benefits of a plant-based diet proved effective for subjects with health concerns. Further insight into the information interventions was acquired by evaluating donation to a climate change charity and developing a text-based measure of relative memorability of the interventions. Interestingly, frequent meat eaters displayed *motivated reasoning* (Kunda, 1990), as they were less inclined to acknowledge the environmental impact of meat and were morally offended at being informed.

Moreover, the results of this chapter can offer guidance to future researchers and policy-makers about effective information strategies for large-scale communication campaigns to promote a reduced meat diet. The study also contributes to the existing branch of literature designing interventions to promote reduced meat intake (Bertolotti et al., 2016; Graham and Abrahamse, 2017; Vainio et al., 2018). More broadly, the study adds to research exploring interventions to 'nudge' sustainable behaviour change (Allcott and Rogers, 2014; Allcott, 2011; Cialdini et al., 1990; Hafner, Elmes, and Read, 2019), specifically information interventions (Goldstein et al., 2008; Hafner, Elmes, Read, and White, 2019). Lastly, this work contributes to studies exploring narrative approaches to presenting information (Shiller, 2017; Slater et al., 2003; McQuiggan et al., 2008).

Finally, Chapter 3 conducts a novel analysis of U.S. Congressional speeches to develop linguistic measures of attitude towards immigrants. Despite contributing to the workforce and tax revenue of the host country, immigrants continue to be subject to constant discrimination, making immigration and attitude towards

immigration a critical matter for public policy. In this chapter, an "immigration corpus" containing 24,351 immigration policy-related speeches, delivered on the floor of the U.S. Congress between 1990 to 2015, was compiled. This corpus was used to build two distinct measures of attitude towards immigrants, based on *sentiment* (or valence) and *concreteness* (as a proxy for social proximity). This approach is inspired from Li and Hills, 2020 who built an immigration corpus from New York Times articles.

Of the two linguistic measures, sentiment displayed systematic variation over time and across states in a manner consistent with the history of immigrant outcomes and political ideology in the U.S. Examination of the voting patterns of the speakers on seminal immigration bills revealed that a speaker-specific sentiment measure is a significant positive predictor of pro-immigration voting behaviour, even after controlling for speaker characteristics and district level socio-economic variables. Furthermore, applying a Latent Dirichlet Allocation or LDA topic modelling algorithm (Blei et al., 2003) to the immigration corpus demonstrated variation in concern over different immigrant-related topics discussed over time, such as a spike in concern related to national security post the 9/11 attacks.

A study of Congressional speeches related to immigrant issues can provide crucial information about the *intensity* of the speaker's attitude toward immigrants, which could subsequently have an impact on shaping immigration legislation. The approach proposed in the paper can be adopted by future researchers to study politicians' attitudes towards other key issues such as gender, race and climate change to predict voting behaviour on relevant bills. This study also contributes to the modest branch of literature which develops linguistic measures of bias including prejudice towards immigrants in news coverage (Li and Hills, 2020; Mastro et al., 2014), and racial (Rice et al., 2019) and gender (Ash et al., 2021) bias in judges' authored opinions. Further, the study contributes to past work examining roll call voting behaviour of U.S. Congress members on immigration bills (Facchini and Steinhardt, 2011; Milner and Tingley, 2011; Fetzer, 2006; Gonzalez and Kamdar, 2000) which had till now relied on economic, demographic and political determinants of voting behaviour. Lastly, this chapter adds to literature using Natural Language Processing tools to study U.S. Congressional speeches (Gentzkow et al., 2019; Jensen et al., 2012).

References

Allcott, H. (2011). Social norms and energy conservation [Special Issue: The Role of Firms in Tax Systems]. *Journal of Public Economics*, *95*(9), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, *104*(10), 3003–37. https://doi.org/10.1257/aer.104.10.3003

Arad, A., & Rubinstein, A. (2012). The 11-20 money request game: A level-k reasoning study. *American Economic Review*, *102*(7), 3561–73. https://doi.org/10.1257/aer.102.7.3561

Ash, E., Chen, D. L., & Ornaghi, A. (2021). Gender attitudes in the judiciary: Evidence from us circuit courts. *Center for Law & Economics Working Paper Series*, *2019*(02). https://doi.org/10.3929/ethz-b-000472797

Bertolotti, M., Chirchiglia, G., & Catellani, P. (2016). Promoting change in meat consumption among the elderly: Factual and prefactual framing of health and well-being [Special Issue: Consumer behaviour in a changing world - Selected papers from the AAEA/EAAE joint seminar in Naples, March 25-27, 2015]. *Appetite*, *106*, 37–47. https://doi.org/10.1016/j.appet.2016.02.150

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of machine Learning research*, *3*, 993–1022.

Cialdini, R., Reno, R., & Kallgren, C. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, *58*, 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678. https://doi.org/10.1037//0022-3514.38.4.668

Das, S. R., & Chen, M. Y. (2007). Yahoo! for amazon: Sentiment extraction from small talk on the web. *Management science*, *53*(9), 1375–1388. http://www.jstor.org/stable/20122297

de Boer, J., de Witt, A., & Aiking, H. (2016). Help the climate, change your diet: A cross-sectional study on how to involve consumers in a transition to a

low-carbon society. *Appetite*, *98*, 19–27. https://doi.org/10.1016/j.appet.2015.12.001

Eaton, L. G., & Funder, D. C. (2003). The creation and consequences of the social world: An interactional analysis of extraversion. *European Journal of Personality*, *17*(5), 375–395. https://doi.org/10.1002/per.477

Facchini, G., & Steinhardt, M. F. (2011). What drives U.S. immigration policy? Evidence from congressional roll call votes. *Journal of Public Economics*, *95*(7-8), 734–743. https://doi.org/10.1016/j.jpubeco.2011.02.008

Fe, E., Gill, D., & Prowse, V. (2019). *Cognitive Skills and the Development of Strategic Sophistication* (CAGE Online Working Paper Series No. 448). Competitive Advantage in the Global Economy (CAGE). https://ideas.repec.org/p/iza/izadps/dp11326.html

Fetzer, J. S. (2006). Why Did House Members Vote for H.R. 4437? *International Migration Review*, *40*(3), 698–706. https://doi.org/10.1111/j.1747-7379.2006.00041.x

Gentzkow, M., Shapiro, J. M., & Taddy, M. (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*, *87*(4), 1307–1340. https://doi.org/10.3982/ECTA16566

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory*, *159*, 369–400. https://doi.org/10.1016/j.jet.2015.07.012

Gill, D., & Prowse, V. (2016). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy*, *124*(6), 1619–1676. https://doi.org/10.1086/688849

Goldstein, N., Cialdini, R., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482. https://doi.org/10.1086/586910

Gonzalez, J. G., & Kamdar, N. (2000). Do Not Give Me Your Tired, Your Poor! Determinants of Legislator Voting on Immigration Issues. *Eastern Economic Journal*, *26*(2), 127–143.

Graham, T., & Abrahamse, W. (2017). Communicating the climate impacts of meat consumption: The effect of values and message framing. *Global Environmental Change*, *44*, 98–108. https://doi.org/10.1016/j.gloenvcha.2017.03.004

Hafner, R., Elmes, D., Read, D., & White, M. P. (2019). Exploring the role of normative, financial and environmental information in promoting uptake

of energy efficient technologies. *Journal of Environmental Psychology*, *63*, 26–35. https://doi.org/10.1016/j.jenvp.2019.03.004

Hafner, R. J., Elmes, D., & Read, D. (2019). Promoting behavioural change to reduce thermal energy demand in households: A review. *Renewable and Sustainable Energy Reviews*, *102*, 205–214. https://doi.org/10.1016/j.rser.2018.12.004

Harguess, J. M., Crespo, N. C., & Hong, M. Y. (2020). Strategies to reduce meat consumption: A systematic literature review of experimental studies. *Appetite*, *144*, 104478. https://doi.org/10.1016/j.appet.2019.104478

Hills, T. T., Proto, E., Sgroi, D., & Seresinhe, C. I. (2019). Historical analysis of national subjective wellbeing using millions of digitized books. *Nature Human Behaviour*, *3*(12), 1271–1275. https://doi.org/10.1038/s41562-019-0750-z

Hirsh, J. B., & Peterson, J. B. (2009). Extraversion, neuroticism, and the prisoner's dilemma. *Personality and Individual Differences*, *46*(2), 254–256. https://doi.org/10.1016/j.paid.2008.10.006

Jensen, J., Naidu, S., Kaplan, E., Wilse-Samson, L., Gergen, D., Zuckerman, M., & Spirling, A. (2012). Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech. *Brookings Papers on Economic Activity*, 1–81.

Johnson, M. K., Rustichini, A., & MacDonald, A. W. (2009). Suspicious personality predicts behavior on a social decision-making task. *Personality and Individual Differences*, *47*(1), 30–35. https://doi.org/10.1016/j.paid.2009.01.050

Krauss, R. M., & Chiu, C.-Y. (1998). Language and social behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., pp. 41–88). McGraw-Hill.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480.

Li, Y., & Hills, T. (2020). Language of Outgroup Prejudice: Revealing Perceptions towards 60 U.S Immigrant Groups from Language Patterns. https://doi.org/10.31219/osf.io/zxhde

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology*, *98*(1), 111–126. https://doi.org/10.1348/000712606X109648

Macdiarmid, J. I., Douglas, F., & Campbell, J. (2016). Eating like there's no tomorrow: Public awareness of the environmental impact of food and reluctance to eat less meat as part of a sustainable diet. *Appetite*, *96*, 487–493. https://doi.org/10.1016/j.appet.2015.10.011

Mastro, D., Tukachinsky, R., Behm-Morawitz, E., & Blecha, E. (2014). News Coverage of Immigration: The Influence of Exposure to Linguistic Bias in the News on Consumer's Racial/Ethnic Cognitions. *Communication Quarterly*, *62*(2), 135–154. https://doi.org/10.1080/01463373.2014.890115

McQuiggan, S. W., Rowe, J. P., Lee, S., & Lester, J. C. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems* (pp. 530–539). Springer Berlin Heidelberg.

Milner, H., & Tingley, D. (2011). The economic and political influences on different dimensions of united states immigration policy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2182086

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance [PMID: 19762717]. *Personality and Social Psychology Bulletin*, *35*(12), 1661–1671. https://doi.org/10.1177/0146167209346309

Proto, E., & Rustichini, A. (2014). Cooperation and Personality. *The Warwick Economics Research Paper Series (TWERPS)*, (1045). https://ideas.repec.org/p/wrk/warwec/1045.html

Proto, E., Rustichini, A., & Sofianos, A. (2019). Intelligence, Personality, and Gains from Cooperation in Repeated Interactions. *Journal of Political Economy*, *127*(3), 1351–1390. https://doi.org/10.1086/701355

Proto, E., Sgroi, D., & Nazneen, M. (2019). Happiness, cooperation and language. *Journal of Economic Behavior & Organization*, *168*, 209–228. https://doi.org/10.1016/j.jebo.2019.10.006

Pullin, P. (2010). Small talk, rapport, and international communicative competence: Lessons to learn from belf. *The Journal of Business Communication (1973)*, *47*(4), 455–476. https://doi.org/10.1177/0021943610377307

Ragan, S. (2014). Sociable talk in women's health care contexts: two forms of non-medical talk. In J. Coupland (Ed.), *Small talk* (pp. 289–307). Routledge. https://doi.org/10.4324/9781315838328-15

Rice, D., Rhodes, J. H., & Nteta, T. (2019). Racial bias in legal language. *Research & Politics*, *6*(2). https://doi.org/10.1177/2053168019848930

Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, *64*, 122–137. https://doi.org/10.1016/j.socec.2016.04.019

Sanchez-Sabate, R., & Sabaté, J. (2019). Consumer Attitudes Towards Environmental Concerns of Meat Consumption: A Systematic Review. *International Journal of Environmental Research and Public Health*, *16*(7). https://doi.org/10.3390/ijerph16071220

Schiermeier, Q. (2019). Eat less meat: UN climate-change report calls for change to human diet. *Nature*, *572*(7769), 291–292. https://doi.org/10.1038/d41586-019-02409-7

Shiller, R. J. (2017). Narrative Economics. *The American Economic Review*, *107*(4), 967–1004. http://www.jstor.org/stable/44251584

Slater, M. D., Buller, D. B., Waters, E., Archibeque, M., & LeBlanc, M. (2003). A test of conversational and testimonial messages versus didactic presentations of nutrition information. *Journal of nutrition education and behavior*, *35*(5), 255–259. https://doi.org/10.1016/s1499-4046(06)60056-0

Stehfest, E., Bouwman, L., Van Vuuren, D. P., Den Elzen, M. G., Eickhout, B., & Kabat, P. (2009). Climate benefits of changing diet. *Climatic Change*, *95*(1-2), 83–102. https://doi.org/10.1007/s10584-008-9534-6

Tilman, D., & Clark, M. (2015). Food, agriculture & the environment: Can we feed the world & save the earth? *Daedalus*, *144*(4), 8–23. https://doi.org/10.1162/DAED_a_00350

Vainio, A., Irz, X., & Hartikainen, H. (2018). How effective are messages and their characteristics in changing behavioural intentions to substitute plant-based foods for red meat? The mediating role of prior beliefs. *Appetite*, *125*, 217–224. https://doi.org/10.1016/j.appet.2018.02.002

Vanhonacker, F., Loo, E. J. V., Gellynck, X., & Verbeke, W. (2013). Flemish consumer attitudes towards more sustainable food choices. *Appetite*, *62*, 7–16. https://doi.org/10.1016/j.appet.2012.11.003

1 The Role of Personality Beliefs and "Small Talk" in Strategic Behaviour [1]

WITH DANIEL SGROI

Humans are predisposed to forming "first impressions" about the people we encounter including impressions about their personality traits. While the relationship between personality and strategic decision-making has been widely explored, we examine the role of personality impressions in predicting strategic behaviour and devising behavioural responses. In a laboratory setting, after only 4-minutes of "small talk", subjects developed a sense of the personality of their partners, particularly extraversion, which consequently changed their behaviour in future interactions. Subjects cooperated more in public goods games when they believed their partner to be extraverted and found it more difficult to out-guess opponents they perceived as similar to themselves in a level-k reasoning task, having engaged in conversation with them. We trace how language can generate these effects using text analysis, showing that talking more makes individuals appear extraverted and pro-social which in turn engenders pro-social behaviour in others.

1.1 Introduction

It is human nature to form "first impressions" or perceptions about the people we meet based on observable verbal and non-verbal behaviours. Social psychologists suggest that the central unit used to understand the behaviour of those around us is closely bound to our perceptions about personality traits (Moskowitz and Olcaysoy Okten, 2016). Information about others' traits plays an integral role when inferring their behaviour in a new setting (Hoffman et al., 1981), which in turn can help us prepare our own behavioural response when we interact with them. The implication is that anything that helps us learn about the personality of others can and will change our behaviour towards them in the future.

[1]Funding for this project was provided by the ESRC CAGE Centre (Grant Ref RES-626-28-0001). University of Warwick Departmental IRB approval obtained (12-03-2018). The experiment for the study is registered at the AEA RCT Registry (RCT ID AEARCTR-0002903) https://www.socialscienceregistry.org/trials/2903. The authors are grateful to Thomas Hills, Kirill Pogorelskiy, Anu Realo, Gordon Brown and Sharun Mukand from Warwick University for helpful discussions related to the paper and to John Taylor from Warwick Business School for help with the experimental sessions. The authors would also like to thank Andis Sofianos for providing the Raven's test matrices used in Proto, Rustichini, and Sofianos, 2019. The authors declare no conflicts of interest.

Personality impressions can be based on a wide variety of elements, such as conversations, manner of speaking, non-verbal actions and physical appearance. Much of the prior literature has focused on personality beliefs formulated from observed physical appearance (Naumann et al., 2009), recorded expressions or behaviour (Hall et al., 2008) and face-to-face interactions (Eaton and Funder, 2003). In our experimental study, we focus on personality beliefs formed in a brief (4-minute) period of "small talk" communication conducted using instant messaging software, together with the ensuing impact of such beliefs on behaviour in later strategic interactions in the laboratory. The emphasis on small talk follows from its ubiquitous role in any social interaction. In a period of negotiation there is often an initial burst of small talk, during a typical working day office workers might chat next to the water cooler or in the office corridor, and appointments with a doctor or financial adviser might begin with pleasantries and a mention of the weather. Opting for instant messaging in the laboratory allows us to omit any confounding effects originating from visual and auditory stimuli. Also, by allowing communication only before the nature of future interactions is known we avoid discussions about future strategies.

Personality theory has become a useful tool in Economics to explain strategic behaviour (Proto, Rustichini, and Sofianos, 2019; Proto and Rustichini, 2014; Rustichini et al., 2016; Johnson et al., 2009; Hirsh and Peterson, 2009). We expand on this by exploring the impact of impressions about another individual's personality on subsequent strategic interactions with them. Given our controlled laboratory setting and the brevity of the communication, our analysis focuses on the two broadest and most fundamental personality traits, extraversion and neuroticism (Costa and McCrae, 1980). Extraversion and neuroticism, which are associated with positive and negative affect, respectively (Costa and McCrae, 1980; Canli, 2004; Watson, Wiese, et al., 1999; Watson and Clark, 1992) are most likely to be detected in a short bout of interaction due to their pervasive nature. Extraverts, characterised by sociability, warmth, gregariousness and positive emotions (McCrae and Costa, 1999), stand out in most social settings. On the other hand, the temperamental traits of high emotions, fear, anger and poor inhibition of impulse, associated with neuroticism (Costa and McCrae, 1980), could also be distinctive in a brief interaction.[2]

Our research strategy is to consider free-form communication: subjects in our laboratory experiment were not aware that they would eventually face each other

[2]In line with this literature, our results confirm that beliefs about the other three "Big Five" traits, openness, conscientiousness and agreeableness cannot be accurately detected in our experiment. In fact our results suggest that, of the two fundamental traits, subjects could only form reasonably accurate beliefs about extraversion after a short conversation.

in strategic settings, but even if they realised that it was likely they had no inkling of the rules of the games to follow. Nevertheless, in the treatment setting they were given the opportunity to communicate with each other: an opportunity not made available to those in the control setting, who instead produced text in an unrelated placebo task. The advantage of this setting is of course that any variation in behaviour between treatment and control groups must be linked causally to the treatment.

Subjects were asked to complete both a standard personality test (the Big Five Inventory[3] or BFI (John and Srivastava, 1999)) and attempt to guess how their partner might have answered the questions in this test. This enabled us to measure the role of communication in developing a cohesive set of beliefs about the personality of their partner. Similarly, subjects were asked to take an IQ test and try to predict how their partners might have performed in the same test. Subjects were also asked to take the "Eyes Test" (Baron-Cohen et al., 2001), which served as a measure of the mental modelling of others, otherwise known as "Theory of Mind" (Coricelli and Nagel, 2009), which could potentially affect the accuracy of belief formation. The Eyes Test and belief elicitation are incentivised as there are measurable correct answers.

Following belief elicitation, subjects engaged in two archetypal and well-understood games: the two-person public goods game and the 11-20 money request game. The public goods game examines social preferences and free-riding and can also be seen as the simplest possible setting in which there is tension between team-work and individual rationality. The 11-20 money request game (Arad and Rubinstein, 2012), on the other hand, is a simple two player game which triggers level-k reasoning (Costa-Gomes et al., 2001) and tests cognitive ability in a competitive environment (Fe et al., 2019). The public goods game requires players to specify how much they are willing to contribute to a communal pot (Fehr and Gaechter, 2002; Herrmann et al., 2008). While both players benefit from contributions, the individually rational choice is to contribute nothing, hoping to free-ride on the other player's contributions. The 11-20 game grants players payment equal to their numerical choice but with a high bonus if they pick a number one below that of their rival. The game is normally modelled using level-k reasoning: if level 0 (L0) involves the non-strategic choice of 20, then L1 (defined as the best response to L0) would be to pick 19. More generally LK, best responding to LK-1 necessitates a choice of 20-K, enabling us to infer the cognitive level of a player through their numerical choice. To omit learning effects the experiment

[3]The Big Five personality traits are extraversion, neuroticism, agreeableness, conscientiousness and openness.

is restricted to one-shot games. Just prior to playing these games, players were asked to predict how their partners might play which was again incentivised: giving us an insight into belief formation. In this way we form a direct link from communication to belief formation to behaviour in two distinct settings.

Our results indicate that beliefs about others' personalities, formed after engaging in small talk with them, can influence decisions made in outcome interdependent games[4]. However, the manner in which personality beliefs influence decision-making depends on the nature of the game. In the level-k reasoning task, where the objective is to out-think the partner, what matters is the perceived difference between the player and their partner's personalities, which may be due to the human tendency of anchoring to self-knowledge when inferring the choices of similar others (Tamir and Mitchell, 2013). In particular, the level chosen in the 11-20 money request game is influenced by the perceived similarity (or difference) between the player and their partner's extraversion. The smaller the perceived difference, the higher the level chosen. This result is consistent with the *perceived similarity hypothesis* (Thomas et al., 2014). The hypothesis posits that individuals believe that those perceived as similar to themselves will think and act like them when faced with the same situation. When the perceived difference between the player and the partner's personality is small, the player chooses a higher level, suspecting that the partner will reason likewise and choose a higher level themselves. This makes it harder for a player to best respond to the distribution of level-k beliefs when the perceived difference between the player and the partner is small, as it becomes harder to out-think the opponent[5].

In contrast, choices in the social preferences game are influenced by the absolute value of the partner's perceived type. We find that, for players who engage in small talk with their partner, cooperation in the public goods game increases when the partner is believed to be extraverted. This is in line with the known association of trait extraversion with pro-social behaviours like cooperation (Carlo et al., 2005; Burke and Hall, 1986). Moreover, *beliefs* about partner's extraversion has a greater effect on cooperation relative to *own* extraversion, a finding robust to whether we use Ordinary Least Squares (OLS) or 2-stage least squares (2SLS) instrumental variable regression specification.

[4]The impact of personality beliefs on strategic behaviour was significantly more pronounced among the treated subjects who engaged in small talk, compared to the control, who had no information upon which to base predictions about their partner's personality.

[5]In the paper we use the terms *opponent* and *partner* interchangeably to refer to the individual the subject was randomly matched with, as the study involved both competitive and cooperative tasks. However, to keep the language neutral, during the experiment the partner or opponent was referred to as 'the other player' (see Experiment script in Appendix 1.C).

Since small talk communication is the only means that players have to develop personality beliefs in the study, and the opportunity to communicate is the only difference between the control and treatment groups, we conducted a direct examination of the text used during small talk. We observed that the more talkative partners are believed to be extraverted, consistent with Mehl et al., 2006, who found that personality judges rated talkative individuals higher on extraversion. While the number of words used is especially helpful as a mechanism for detecting extraverts, providing a reasonably accurate forecast of type, there remains a persistent own-type bias: particularly, extraverts are prone to *complementary self projection bias* making them likely to overstate the extraversion in their partners.

Extraversion is particularly relevant when examining the role of personality beliefs in influencing strategic behaviour. This is because, of all the personality traits, subjects could only form reasonably accurate beliefs about a stranger's extraversion, after engaging in small talk with them for a brief period. Extraverts, due to their sociability, vigour and outgoing friendliness, are distinctive by nature, making them the most detectable in a brief interaction. Accurate impressions about the other personality dimensions might require future research involving longer interaction times in real-world settings.

Alongside our main contribution on the role of personality beliefs on strategic behaviour, we contribute to research exploring *personality attribution*, by focusing on impressions formed from instant messaging rather than physical appearance or face-to-face interaction (Eaton and Funder, 2003; Little and Perrett, 2007; Naumann et al., 2009; Albright et al., 1988). We also add to the existing modest research on the role of small talk which has focused on topics such as building solidarity in work places (Pullin, 2010), examining investor sentiment using discussions on stock message boards (Das and Chen, 2007) and improving medical outcomes (Ragan, 2014).[6] Our study instead focuses on the role of small talk on unknown future strategic settings and in particular on the relationship with personality theory which in turn feeds into belief formation. Our focus is therefore on the mechanism that allows unstructured communication to alter behaviour and outcomes that are unknown at the time of communication. Lastly, our study contributes to the literature on strategic sophistication which finds that individuals adjust strategies given the information they have about the opponents (Fe

[6]We should also contrast the literature on "small talk" with the the large literature on communication with *prior knowledge* of what is to follow (Charness and Dufwenberg, 2006; Bochet et al., 2006; Cooper et al., 1992; Dawes et al., 1977) in which individuals can send messages that relate to future decision-making. In contrast to this "cheap talk" literature, our paper studies how communication between players can affect behaviour when the nature of any future interaction ("rules of the game") is unknown to the players which makes it harder to incorporate strategic content into communication, forcing our subjects to engage in small talk.

et al., 2019; Georganas et al., 2015; Gill and Prowse, 2016). Existing work finds that people adjust strategies based on *exogenous* information provided such as information about the opponent's cognitive ability (Fe et al., 2019). We add to this literature through a novel examination of how individuals adjust their behaviour in the light of *endogenous* belief formation about the opponent's personality.

The rest of the paper is structured as follows. Section 1.2 details the experimental design and the core hypotheses. Section 1.3 presents the results from the experiment. As the very first study of the interaction between personality beliefs, small talk and strategic behaviour, our work will be necessarily exploratory. Thus, the study can act as a first step before further research: we discuss this in section 1.4, along with a discussion of the key findings. Section 1.5 concludes.

1.2 Methodology

1.2.1 Experimental Design

The experiment was conducted in a laboratory setting. At the onset of the experiment each subject was asked to take the 44-item Big Five Inventory personality test or BFI (John and Srivastava, 1999). The answers to the BFI questionnaire were used to compute an average score for each of the 5 personality traits and the trait scores were then standardised (so each trait distribution had mean 0 and standard deviation 1). This was followed by an incentivised cognitive ability test, taken from the Raven's Progressive Matrices test (Raven, 2003), in which subjects were asked to attempt 30 visual puzzles (adapted from Proto, Rustichini, and Sofianos, 2019). The test was incentivised to motivate cognitive effort required in the task, as is the standard approach within Economics (Proto, Rustichini, and Sofianos, 2019; Proto, Sgroi, et al., 2019). After the Raven's test the subjects were asked their beliefs about their own performance in the test which was also incentivised. Each subject was randomly allocated to one of two groups and randomly paired with a partner from the same group as follows:

Control: Players were not allowed to communicate with their partners in this condition. Subjects were asked to take part in a placebo task for 4 minutes (full experiment instructions are provided in Appendix 1.C). Then the players were asked their beliefs about their partner's personality and cognitive abilities. For the former, beliefs were elicited using an 11-item short version of the BFI questionnaire, adapted from Rammstedt and John, 2007 and modified to allow subjects to indicate how they felt their partners would answer the questions (the

personality belief questionnaire is presented in Appendix 1.D).[7] We could then form personality beliefs directly from the answers they provided. For the latter, subjects were asked how they felt their partner's performed in the Raven's task. After answering the questions related to beliefs, subjects were told the rules of the first game. They were asked for their beliefs about their partner's strategy followed by their own decision in the game. After completing the first game they were told the rules of the second game. As with game 1, they were asked their beliefs about the partner's strategy and their own decision in the game. The partner remained the same for both games. The outcomes of both games were announced at the end of the experiment. Beliefs about the partner's cognitive abilities and personality, and beliefs about their strategies were incentivised.

Treatment: The procedure in the treatment group was the same as the control except, instead of the placebo task, subjects were allowed to electronically communicate with their partners through a chat box on their screens. Note that crucially communication occurred before the nature of future decisions were apparent which makes it difficult to incorporate strategic content specific to the game into communication. Communication time was limited to 4 minutes. Following communication, the players were asked to answer the same belief questions as the control group. After answering the questions, the subjects were told the rules of the first game and asked to play the game. The process was repeated with the second game, as with the control condition.

Subjects were asked to play 2 games, the public goods game and the 11-20 money request game. In the *public goods game* each subject was allocated 20 Experimental Pounds (EP) and, along with their partner, were asked to choose (simultaneously) how much to contribute (c_i) to a joint project. c_i was restricted to be an integer between 0 and 20. Payoffs were determined as: $\pi_i = (20 - c_i) + \frac{3}{4}(c_i + c_j)$ where i and j were the two players. Higher contributions while more costly, were more socially beneficial. In the public goods game, the selfish equilibrium is 0 and the mutually cooperative response is 20. In the *11-20 money request game* participants were asked to play the basic version of the game (Arad and Rubinstein, 2012). Each player was randomly matched with another player. They were both asked to request an amount of money, an integer between 11 and 20 EP. Each player received the amount they requested. A player received an additional

[7] In essence, players were asked to retake the BFI, albeit a shorter version, but rather than considering how they would answer each question, they were instead asked how their partner would answer. This allows us to form a belief in much the same way as we formed implied trait values. The 11-item questionnaire consists of 2 items each for the traits extraversion, conscientiousness, openness and neuroticism and 3 items for the agreeableness trait. An average score was computed for each trait and the trait scores were then standardised.

amount of 20 EP if they asked for exactly one less than the other player. This game has been used to study cognitive hierarchy and in particular level-k thinking. In level-k hierarchy models (Nagel, 1995; Stahl and Wilson, 1995; Stahl and Wilson, 1994) players' levels or types are heterogeneous but they are assumed to be drawn from the same distribution. Peoples' beliefs are based on naive initial assessment of others' likely response called level-0 (or L0) and then beliefs are modified via iterated best response. So level 1 (L1) best responds to L0, L2 to L1 and so on. Following Arad and Rubinstein, 2012, 20 is considered the salient and L0 choice since it requires no strategic thinking about the other player's choice. This implies that a choice of 19 is the L1 choice as it best responds to the L0 strategy and in general the level-X choice is to request 20-X. In the level-k model, the level chosen by a subject is a measure of their strategic sophistication or *type* or rather a measure of the player's beliefs about the partner's strategic sophistication or type (Georganas et al., 2015). The game has no pure Nash equilibrium. The order of the 2 games was randomised across sessions.

Following the two games, subjects were asked to take the *Eyes Test* (Baron-Cohen et al., 2001). For this test, subjects were shown 36 close-up photographs of the eyes and surrounding areas of the face of celebrities and were provided with 4 response options (such as playful, terrified, joking etc.) per photograph. The participants were asked to pick the option which most closely described the mental state of the person in the photograph. Subjects were then asked to answer a list of 30 questions about their risk attitude, the Domain Specific Risk Taking Scale or DOSPERT (Blais and Weber, 2006). Each subject was then asked a series of socio-demographic questions including age, gender and native language.

1.2.2 Logistics

The experiment was conducted between May and November 2018. Subjects were recruited through the SONA online recruitment system at the University of Warwick in the UK. The participants were undergraduate, postgraduate and staff members at the University. The experiment was implemented using Z-tree (Fischbacher, 2007) and pre-registered with the AEA RCT registry (Sgroi, 2018). The experiment received ethical approval from Economics Department Internal Ethical Approval Process, University of Warwick. 338 subjects took part in the study, with 170 subjects in the control condition and 168 in the treatment group. Out of the 170 control group subjects, 110 subjects played the public goods game first, followed by the 11-20 money request game, and 60 subjects played the games in reverse order. Out of 168 treatment group subjects, 106 played the public goods game first and 62 played the 11-20 money request game first. There

were 17 sessions conducted, 20 subjects per session on average. An experimental session lasted for approximately 75 minutes.

The final payoff for subjects in the experiment was made up of several components. Firstly, there was a show-up fee of £4. Second, the players received payoffs based on performance in either the public goods game or 11-20 money request game (chosen randomly). The payoffs for the games were in experimental pounds (EP) with the exchange rate as 5 EP = £1. Third, 2 questions out of the 36 questions of the Eyes Test and 2 puzzles of the 30 puzzles of the Raven's test were randomly selected with each correct answer accruing a further £1. Lastly, belief questions (about own-cognitive ability, partner's personality and cognitive ability, and beliefs about partner's decisions in the 2 tasks) were also incentivised. For the personality beliefs, 1 out of 11 questions was randomly picked and if the answer matched that of the partner then the subject was awarded £1. For the other 4 belief questions, subject was awarded £1 for each correct answer. The socio-demographic questions were not incentivised. The average earnings from the study was £13.20 (including the show-up fee of £4), with a minimum earning of £8.35 and maximum of £18.

1.2.3 Hypotheses

Of the "Big Five" personality traits, the scope of our paper is limited to the two broadest, most fundamental and pervasive traits: extraversion and neuroticism (Costa and McCrae, 1980). These two traits were the original "Big Two" personality dimensions (Eysenck, 1947). Extraversion and neuroticism, have garnered much attention in the literature owing to their well-established association with positive and negative affect, respectively (Canli, 2004; Watson, Wiese, et al., 1999; Watson and Clark, 1992; Costa and McCrae, 1980) which gives these two traits the greatest chance to be detected in a short bout of communication.

Extraverts by their nature stand out and even in a few minutes it may become clear that you are dealing with someone who is characterised by sociability, gregariousness, assertiveness, warmth, activity and overall positive emotions (McCrae and Costa, 1999). On the other hand, the temperamental traits of general emotionality, fearfulness, anger and impulsivity, are associated with the neuroticism trait, and are related to high negative affect (Costa and McCrae, 1980), which might also be detectable in a brief conversation. This makes any short communication, such as in our study, more suited to developing reliable beliefs about the partner's (or the opponent's) extraversion and neuroticism traits, which can be interpreted by the perceiver as positive and negative vibe given off by the opponent, respectively. However, a brief small talk conversation seems insuffi-

cient to form beliefs about the partner's remaining three Big Five traits. This is because while a brief chat is sufficient to form an overall positive (*extraversion*) or negative (*neuroticism*) view about someone, it is not adequate to convey any usable information about whether the opponent is trusting (an aspect of trait *agreeableness*) or lazy (an aspect of trait *conscientiousness*) or imaginative (an aspect of trait *openness*). Thus, we will limit our hypotheses to the effect of the fundamental personality traits on belief formation and strategic decision making.[8] Our experimental setup gives us the following testable hypotheses.

Hypothesis 1: *Personality beliefs about the opponent are not only influenced by the opponent's true personality measure, but the beliefs are also influenced by the player's own personality.*

This hypothesis is consistent with the conceptual framework for the impact of social environment on personality proposed by Eaton and Funder, 2003, which posits that perceptions (or predictions) about any individual's personality trait can be influenced by the degree to which the predictor possesses that specific trait themselves. This seems particularly true for extraverts who stimulate a positive social environment around them due to their own positivity, making them prone to projecting their extraversion or sociability onto others (Eaton and Funder, 2003; Thorne, 1987). For our study, we would only expect to see personality projection in the treatment group since any personality beliefs that appear in the control group must be spurious (given the control group have no information whatsoever upon which to base predictions about their partner's personality).

Hypothesis 2: *Strategic decision making in outcome interdependent tasks is affected by the individual's beliefs about the opponent's personality, an effect which is significantly more pronounced among treatment group subjects who engage in small talk communication.*

We also formulate individual hypotheses about the unique way in which personality beliefs can affect the two different tasks.

Hypothesis 2a: *In the 11-20 money request game, rather than one's own personality or beliefs about the opponent's personality, we hypothesise that choices in the game will be influenced by the perceived differences in the pair's personalities.*

Due to the strategic nature of the 11-20 money request game, the objective of

[8]With respect to beliefs about the opponent's IQ, we will refrain from formulating any hypotheses given the lack of available literature and where appropriate we will present our results about IQ beliefs as more speculative.

this level-k reasoning game is to correctly gauge the opponent's choice and then attempt to out-think them. Thus, the game does not solely depend on one's own type, but success in the game is determined by the ability to out-guess the opponent by assessing their type. Despite the well established link between IQ and level-k reasoning (Gill and Prowse, 2016), *beliefs* about opponent's IQ might seem like an unreliable measure of the opponent's strategic sophistication or type in the limited interaction time available. Beliefs about the opponent's fundamental personality traits on the other hand can appear as a more reliable measure of the opponent's type due to the increased likelihood of them being detected through a brief chat. While personality itself lacks any association with level-k reasoning, any difference (or similarity) between the pair's types (which for our study is personality types) can be interpreted by the player as an indicator of the opponent's behaviour and thus, in turn, can act as a determinant of own decision making. Consistent with simulation theories of social cognition, individuals tend to anchor on self-knowledge to form mental images about similar others (Tamir and Mitchell, 2013). The *perceived similarity hypothesis* states that the greater the perceived similarity between the individual and their opponent the more likely it is that the individual will believe their opponent to think and act like themselves (Thomas et al., 2014), making perceived similarity or differences a potential contributor to iterative reasoning processes.

Hypothesis 2b: *Players who believe their opponents (or partners) are extraverted, will believe that their opponents will cooperate more and then they in turn will cooperate more themselves.*

This seems reasonable given the known association between extraversion and pro-social behaviours like cooperation (Carlo et al., 2005; Burke and Hall, 1986). This might encourage the individual to cooperate more, with the hope of mutual cooperation boosting earnings.

Hypothesis 3: *More talkative opponents are believed to be extraverted.*

In this paper, we randomly allocate players either to a treatment in which they engage in small talk with their partners or to a control in which they do not. Since small talk is the only interaction the subjects engage in before eliciting beliefs about the partners' personalities, it must form the basis for these beliefs. From the player's perspective the number of words is relatively simple to calculate, arguably easier than say considering the emotional content of words in a very brief conversation. Thus, it is hypothesised that subjects using more words will be rated higher on the extraversion scale as extraverts are usually characterised

by their sociability and talkativeness (Goldberg, 1990; Costa and McCrae, 1992). Further, in a study of personality traits in its natural habitat, personality judges rated talkative participants as more extraverted (Mehl et al., 2006). We will also evaluate other linguistic features, namely valence, arousal and dominance content of the words spoken by the partner. Valence refers to the pleasantness of a stimulus, arousal is the intensity of emotion provoked by a stimulus, and dominance is the degree of control exerted by a stimulus (Warriner et al., 2013).

Note, while the hypotheses related to personality beliefs (*hypothesis 1*) and the strategic decision making tasks (*hypotheses 2a* and *2b*) were formulated before the experimental trials (based on the pertinent literature cited), the results from the text analysis (*hypothesis 3*) were harder to predict prior to the study owing to the novelty of the setup and were thus more exploratory in nature.

1.3 Results

This section tests our core hypotheses. Section 1.4 offers a more in-depth discussion of the key findings of the paper. All regressions reported were run with standardised variables with standard errors clustered at the pair level. The summary statistics of the variables used in the paper are presented in the table 1.A.1 and the balance tests for the intervention groups are provided in table 1.A.2.

1.3.1 Result 1: Personality projection

We begin by looking at the factors that might affect the beliefs which players develop about their partners' personality traits. The aim is to examine *hypothesis 1* which proposes that beliefs about an individual's personality depend not only on their true personality traits but are also affected by the predictor's personality.

Table 1.1 reports the results of an OLS regression model. The dependent variable is the belief reported by the player about their partner's level of extraversion and neuroticism.[9] The independent variables in columns 1 and 3 are the player's own personality scores, the partner's true personality scores (as reported by the partner using the BFI), and their interactions with the treatment dummy which equals 1 if the player was in the small talk condition and 0 otherwise. Columns 2 and 4 also control for the subject's IQ, Eyes Test score, age, a dummy variable for being female, and risk aversion (along with the interactions of the control variables with the treatment dummy). Column 2 shows that in the treatment group,

[9]Recall that beliefs are formed in much the same way as underlying values: while personality is assessed using the BFI questionnaire, personality beliefs are elicited using a shorter version of the BFI (Rammstedt and John, 2007). For both, average trait scores are calculated and the standardised values are used in the regressions.

Table 1.1: Impact of own personality and partner's true personality on beliefs about partner's personality

	Extraversion Belief		Neuroticism Belief	
	(1)	(2)	(3)	(4)
Own Extraversion × Treatment	0.2139*	0.2962**	-0.1105	-0.1241
	(0.117)	(0.125)	(0.117)	(0.130)
Own Neuroticism × Treatment	0.1484	0.1531	-0.0470	-0.0418
	(0.125)	(0.131)	(0.110)	(0.109)
Partner's Extraversion × Treatment	0.4108***	0.4199***		
	(0.108)	(0.110)		
Partner's Neuroticism × Treatment			0.0269	-0.0005
			(0.103)	(0.102)
Own Extraversion	0.0209	0.0248	-0.0822	-0.0718
	(0.073)	(0.080)	(0.073)	(0.075)
Own Neuroticism	-0.0075	0.0008	0.0462	0.0600
	(0.085)	(0.087)	(0.083)	(0.080)
Partner's Extraversion	-0.1280*	-0.1339*		
	(0.070)	(0.075)		
Partner's Neuroticism			0.0866	0.1069
			(0.071)	(0.070)
Treatment	0.3539***	-0.3127	-0.5100***	-0.1983
	(0.098)	(0.632)	(0.102)	(0.550)
Controls	No	Yes	No	Yes
N	338	338	338	338

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The specification for the OLS regressions is:

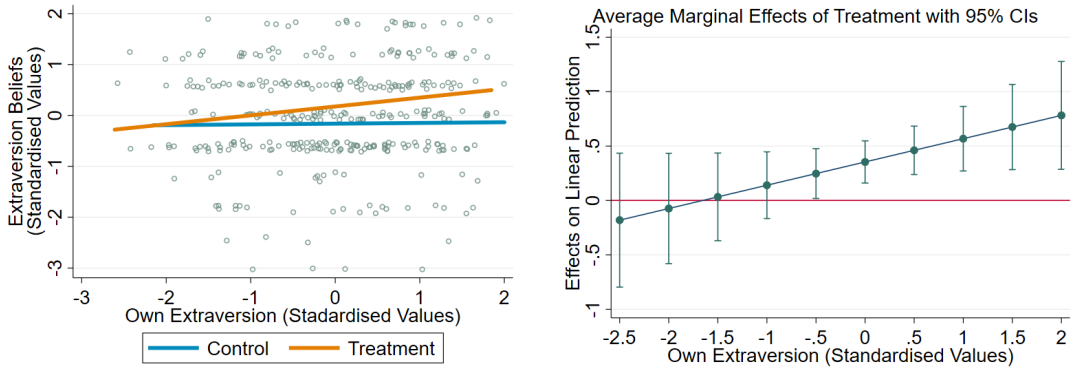
$$E_i(pers_j) = \beta_1 pers_i \times Treat + \beta_2 pers_j \times Treat + \gamma_1 pers_i + \gamma_2 pers_j + \phi z_i \times Treat + \lambda Treat + \omega z_i + \epsilon_i \quad (1.1)$$

$pers_i$ is player i's personality, $E_i(pers_j)$ is player i's beliefs about partner j's personality and $pers_j$ is partner j's true personality. Also, $Treat$ is the treatment dummy which equals 1 if the player is in the small talk group and 0 otherwise, z_i are individual characteristics of i (i.e. the control variables, namely player i's IQ, Eyes Test score, age, a dummy variable for being female, and risk aversion) and ε_i is an idiosyncratic error term.

an increase in the *player's own extraversion* by 1 standard deviation increases the beliefs about *partner's extraversion* by 0.3 standard deviations more than in the control group (p-value < 0.05). Furthermore, an increase in 1 standard deviation in partner's true extraversion increases the player's beliefs about their partner's extraversion by 0.4 standard deviations more in the treatment group than in the control group (p-value < 0.01). Note that the negative coefficient in the control group for Partner's extraversion (in columns 1 and 2) is spurious and a statistical artifact driven by noise, since in the control group subjects had no reliable source of information about their partners' true extraversion. This

biases the coefficient for Partner's Extraversion × Treatment upwards. However, the effect of partner's true extraversion on beliefs developed about the partner's extraversion remains significant when limiting the analysis just to the treatment group, even after adding the control variables, with coefficient .286 and p-value < 0.01. This coefficient reflects the impact of partner's true extraversion on extraversion beliefs, as compared to an 'ideal' control group with a coefficient of 0 (which of course is impossible to replicate using human subjects).

Column 4 shows that in the treatment group, an increase in the player's extraversion by 1 standard deviation decreases the beliefs about partner's neuroticism by 0.1 standard deviations more than in the control group, although the differential effect is statistically insignificant. Column 4 also shows that a partner's true neuroticism has no significant effect on beliefs developed about their neuroticism trait. Thus, we find that a 4-minute small talk chat can lead to reliable beliefs about a partner's extraversion but *not* neuroticism. The relation between own extraversion and beliefs about partner's extraversion is depicted in Figure 1.1. Consistent with *hypothesis 1*, we observe that extraverts project their positive affect onto their partners.



(a) (b)

Figure 1.1: Relationship between the player's beliefs about partner's extraversion and the player's own extraversion score. (a) shows that individuals are more likely to project their own extraversion on to their partners in the Treatment group compared to Control. (b) shows that this difference in extraversion projection between the Treatment and the Control group increases with the value of the predictor's own extraversion.

For the other 3 Big Five Traits, agreeableness, conscientiousness and openness, the Pearson correlation coefficients between beliefs and true values in the treatment group were trivial and statistically insignificant, with coefficients (r) 0.0372 (p-value = 0.6319), 0.0403 (p-value = 0.6044) and -0.0588 (p-value = 0.4491), respectively. Only for extraversion did we observe significant correlation (r =

0.2513, p-value = 0.0010) between beliefs and true scores in the treatment group, while the coefficient for neuroticism was also insignificant (r = 0.1169, p-value = 0.1314).

We also observed that overestimation of partner's extraversion increases with the player's own extraversion (Table 1.A.3). This overestimation is significantly (p-value < 0.05) more pronounced in the treatment group, compared to the control. Further, we found that with increasing performance in the eyes test, the inaccuracy in the player's beliefs about partner's extraversion is significantly (p-value < 0.10) lower in the treatment group compared to the control. This is consistent with the literature on the eyes test (Baron-Cohen et al., 2001), which posits that better performance in the eyes test indicates increased theory of mind ability, which in turn leads to improved understanding of others' mental states. With regards to beliefs about partners' cognitive abilities, it was observed that players project beliefs about their own IQ onto beliefs about partners' IQ, irrespective of whether they are in the control or treatment group (Table 1.A.4).

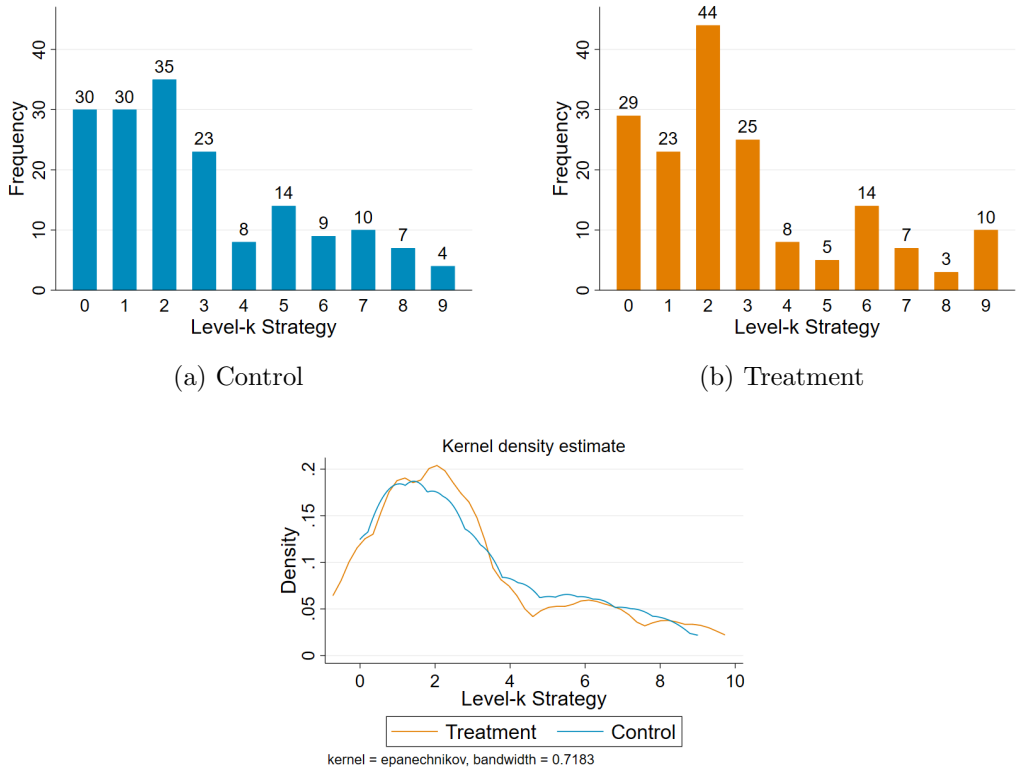
1.3.2 Result 2: Strategic decision-making and personality

Since we divided hypothesis 2 into two parts, each associated with one of our two games, we will also divide our results in the same way.

Result 2a: Level-k reasoning and perceived similarity

Recall that *hypothesis 2a* claims that level-k reasoning is influenced by the perceived differences (or similarities) in the player and their opponent's types (which for our study is personality types). In our data, L2 is the most frequently played strategy in both conditions: where 20.6% players choose L2 in the control condition and over 26% do so in the treatment condition (Figure 1.2). The Kolmogorov-Smirnov test revealed that there is no statistical difference between the distribution of levels of the 2 groups. Further, there is no significant difference between the payoffs earned in the 11-20 game by the control and the treatment group subjects (while the treatment group earns 19.7 EP on average, the control group earns 19.6 EP). Since the level-k game is a competitive game, so long as the communication is two-sided, small talk is unlikely to benefit either player.

Table 1.2 reports the results of OLS regressions. In columns 1-3 the dependent variable is the player's beliefs about the level-k strategy chosen by the partner and in columns 4-6 the dependent variable is the level-k strategy chosen by the player. The independent variables are perceived differences between player's own personality and the partner's personality, and the interaction of perceived differences with the treatment dummy. The perceived differences are computed



(a) Control

(b) Treatment

(c) Kernel Density Plot

Figure 1.2: The distribution of level-k strategy chosen in the 11-20 money request game

by taking the standardised absolute difference between the player's own personality trait scores and the player's beliefs about the partner's personality trait scores. Columns 2 and 4 also include the player's own personality and the personality measures interacted with the treatment dummy as explanatory variables. Columns 3 and 6 include sensible control variables i.e. player's eyes test score, IQ, gender, the player's beliefs about partner's IQ and the order of play of the two games, which is a dummy that equals 1 when the 11-20 game is played first and 0 when the public goods game is played first (along with the variables interacted with the treatment dummy). Columns 3 and 6 also include the control variables - player's age and risk aversion, along with their interactions with the treatment dummy. Column 3 shows that an increase in 1 standard deviation in perceived difference in extraversion decreases the player's beliefs about partner's level choice by 0.5 more in the treatment group than in the control group (p-value < 0.10). Column 6 shows that an increase in 1 standard deviation in perceived difference in extraversion decreases the player's own level-k strategy by 0.6 more in the treatment group than in the control group (p-value < 0.05).

Thus, there is an inverse relationship between the perceived difference in extraver-

Table 1.2: Impact of (absolute) difference between own personality and beliefs about partner's personality on level-k strategy chosen

	Level Belief			Level Chosen		
	(1)	(2)	(3)	(4)	(5)	(6)
DiffExtraversion × Treatment	-0.5302*	-0.5562*	-0.5260*	-0.6597***	-0.7373***	-0.6442**
	(0.269)	(0.283)	(0.289)	(0.237)	(0.242)	(0.254)
DiffNeuroticism × Treatment	0.1879	0.2460	0.3734	-0.0415	0.0235	0.1925
	(0.248)	(0.258)	(0.292)	(0.248)	(0.243)	(0.265)
DiffExtraversion	0.1470	0.1430	0.1036	0.2046	0.1792	0.1345
	(0.198)	(0.194)	(0.197)	(0.177)	(0.172)	(0.175)
DiffNeuroticism	-0.1579	-0.1632	-0.2618	-0.1604	-0.1620	-0.2974
	(0.183)	(0.188)	(0.213)	(0.174)	(0.178)	(0.186)
Treatment	0.1668	0.1515	-2.8375	0.0677	0.0330	-2.2355
	(0.267)	(0.268)	(2.058)	(0.279)	(0.276)	(1.860)
Own Extraversion × Treatment		-0.0312	0.0404		-0.1293	0.0116
		(0.294)	(0.344)		(0.290)	(0.312)
Own Neuroticism × Treatment		-0.2018	-0.1717		-0.4371	-0.4405
		(0.279)	(0.306)		(0.278)	(0.279)
Own Extraversion		-0.0532	-0.1518		-0.1726	-0.2696
		(0.195)	(0.201)		(0.211)	(0.212)
Own Neuroticism		0.0132	-0.1102		0.1998	0.0391
		(0.198)	(0.216)		(0.198)	(0.196)
Eyes Test Score × Treatment			0.5507*			0.6041*
			(0.303)			(0.309)
Own IQ × Treatment			-0.2617			-0.2965
			(0.292)			(0.299)
IQ Belief × Treatment			0.3253			0.1933
			(0.311)			(0.264)
Female × Treatment			-0.7230			-0.8284
			(0.611)			(0.555)
Order × Treatment			1.0992*			1.0541*
			(0.576)			(0.592)
Eyes Test Score			-0.4245*			-0.4401*
			(0.247)			(0.248)
Own IQ			0.1777			0.2357
			(0.200)			(0.210)
IQ Belief			-0.3339			-0.3220*
			(0.204)			(0.192)
Female			1.1333***			1.4426***
			(0.431)			(0.384)
Order			-0.7822**			-1.0035**
			(0.392)			(0.408)
Controls	No	No	Yes	No	No	Yes
N	338	338	338	338	338	338

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The specification for the OLS regressions is:

$$Y_i = \nu Diffpers_i \times Treat + \tau Diffpers_i + \eta Treat + \kappa pers_i \times Treat + \theta pers_i + \rho z_i \times Treat + \psi z_i + \xi_i \quad (1.2)$$

Y_i is player i's beliefs about partner j's level chosen in the 11-20 game in columns 1-3. For columns 4-6 Y_i is the level chosen by player i in the game. $Diffpers_i$ i.e. the absolute difference in i and j's personalities as perceived by i i.e. $|E_i(pers_j) - pers_i|$ where $pers_i$ is player i's personality, $E_i(pers_j)$ is player i's beliefs about partner j's personality and $pers_j$ is partner j's true personality. Also, $Treat$ is the treatment dummy, z_i are individual characteristics of i and ξ_i is an idiosyncratic error term. z_i includes player i's eyes test score, IQ, gender, the i's beliefs about partner j's IQ, the order of play of the two games, which is a dummy that equals 1 when the 11-20 game is played first and 0 when the public goods game is played first and the additional control variables, player i's age and risk aversion.

sion between the players, and the player's level-k strategy, as well as the player's beliefs about their partner's level-k strategy choice. Hence, the smaller the perceived difference between the two players the greater the beliefs about partner's level choice and the greater the level chosen by the player.[10] This result supports *hypothesis 2a* and is consistent with the *perceived similarity hypothesis* which posits that people project their own thinking and decision-making process to predict how their partners might think and act when individuals believe their partners to possess attributes similar to their own (Thomas et al., 2014). Thus, when players believe their partners to be similar to themselves (small perceived difference), they believe their partners will reason more and choose a higher level (i.e. lower number in the 11-20 game). This in turn makes the player choose a higher level. Similar results were not observed for perceived difference between player's own IQ and partner's IQ.

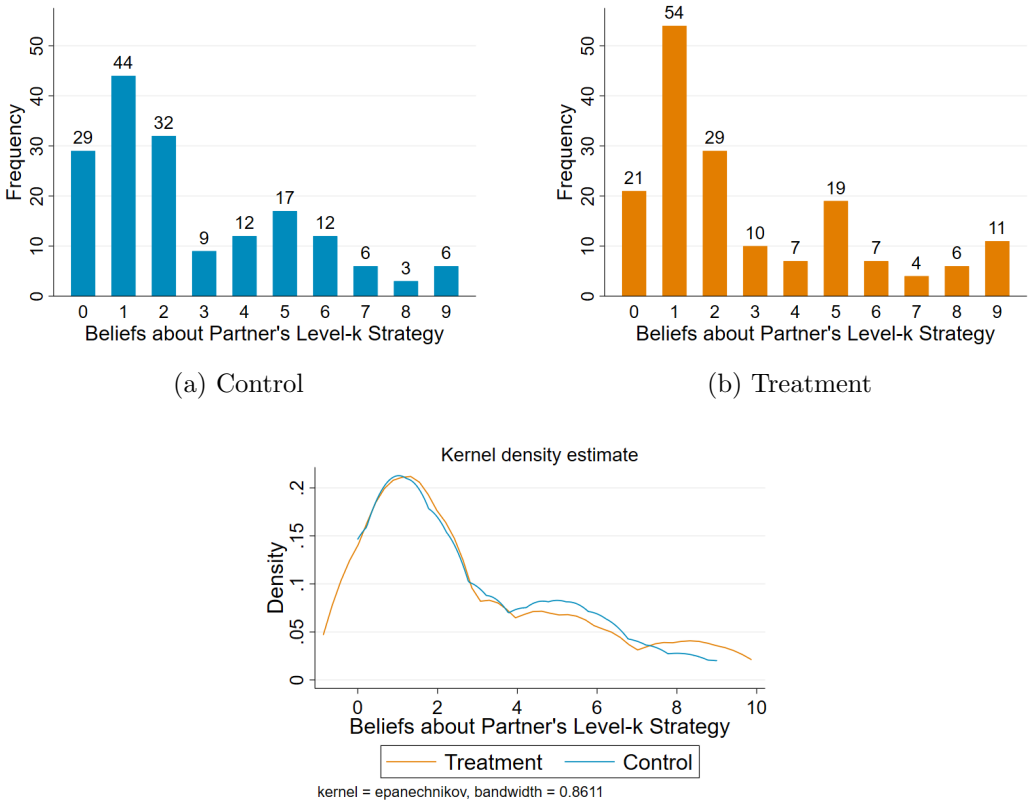
Being female enhances beliefs about partner's level-k choice, as well as player's own level-k choice, although there is no significant differential treatment effect.[11] Further, an increase in the eyes test score by 1 standard deviation increases level belief and level chosen by 0.5 and 0.6 more in the treatment than in the control group, respectively. This supports the finding (Fe et al., 2019; Georganas et al., 2015) that greater engagement in theory of mind is associated with superior level-k reasoning, though in this study the effect is significantly (p-value < 0.10) stronger in the treatment group when the players are able to engage in small talk with their partners, compared to the control group. In the control group, order of the tasks has a negative effect on the level-k belief and their own level-k action, whereas in the treatment group the coefficients are positive.

Next, the paper looks at the distribution of the players' beliefs about the level-k strategy chosen by their partners (Figure 1.3). The distribution is presented in Table 1.3, along with the unique mixed strategy Nash equilibrium distribution for risk-neutral players. The distributions of beliefs observed in both treatment and control groups are different from the equilibrium distribution. In both groups, L1 (i.e choosing 19) is the most frequently believed level-k choice by partners. Table 1.4 calculates the expected payoffs based on the distribution of level-k beliefs observed. For both control and treatment groups, L2 (choosing 18) has

[10]Note that the results remain similar when we control for beliefs about partner's personality. The results are omitted here for parsimony but presented in Table 1.A.5.

[11]Nettle and Liddle, 2008 and Stiller and Dunbar, 2007 have found that women score higher on the social-cognitive element of theory of mind, indicating greater ability to reason about others' mental states. This could explain why women choose higher levels.

the highest associated expected payoffs.[12]



(a) Control

(b) Treatment

(c) Kernel Density Plot

Figure 1.3: The distribution of the player's beliefs about partner's level-k strategy in the 11-20 money request game

Table 1.3: Distribution of Level-k beliefs

Level	0	1	2	3	4	5	6	7	8	9
Equilibrium (%)	5	10	15	20	25	25				
Treatment (%)	12.50	**32.14**	17.26	5.95	4.17	11.31	4.17	2.38	3.57	6.55
Control (%)	17.06	**25.88**	18.82	5.29	7.06	10.00	7.06	3.53	1.76	3.53

Table 1.4: Expected payoffs from the distribution of Level-k beliefs

Level	0	1	2	3	4	5	6	7	8	9
Treatment (EP)	20.00	21.50	**24.43**	20.45	17.19	15.83	16.26	13.83	12.48	11.71
Control (EP)	20.00	22.41	**23.18**	20.76	17.06	16.41	16.00	14.41	12.71	11.35

[12]It should be noted that the number of people who best-responded to their own belief about their partner's level choice i.e. chose to request an amount which was exactly 1 lower than what they believed their partner would chose was 184 out of 334 (94 in the control group and 90 in the treatment group) i.e. 54.4%. The low proportion of people best-responding to their own belief suggests that rather than having an exact belief about their partner's level choice, they may have formed a distribution of beliefs.

Table 1.5: Impact of (absolute) difference between own personality and beliefs about partner's personality on the probability of choosing the best response - Probit Model

	Control		Treatment	
	(1) Pr(Level=2)	(2) Pr(Level=2)	(3) Pr(Level=2)	(4) Pr(Level=2)
DiffExtraversion	-0.0453	-0.0492	0.0846***	0.0945***
	(0.038)	(0.036)	(0.030)	(0.032)
DiffNeuroticism	-0.0008	-0.0078	-0.0459	-0.0362
	(0.031)	(0.031)	(0.032)	(0.034)
Own Extraversion		0.0115		0.0017
		(0.029)		(0.045)
Own Neuroticism		0.0573*		-0.0399
		(0.032)		(0.037)
Own IQ		0.0655*		0.0566
		(0.035)		(0.039)
IQ Belief		-0.0482*		-0.0070
		(0.029)		(0.035)
Eyes Test Score		0.0541		0.0498
		(0.038)		(0.032)
Controls	No	Yes	No	Yes
N	170	170	168	168

Standard errors in parentheses. Statistical significance indicated as follows:

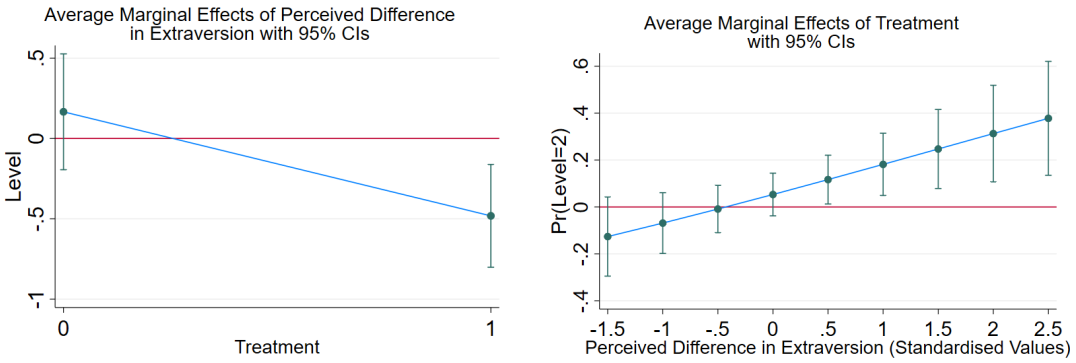
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the average marginal effects from Probit regressions. 'Controls' imply the player's age, gender, risk aversion, and the order of the two games played.

Table 1.5 uses a probit model to examine the effect of perceived differences in the player's and their partner's personalities on the probability of best responding to the distribution of level-k beliefs, in the control and treatment groups separately. The dependent variable is the probability of choosing the best response to the distribution of beliefs which in this case is L2 for both control and treatment groups. Column 4 shows that the probability of best responding increases significantly (p-value < 0.01) by 9 percentage points with a 1 standard deviation increase in the perceived difference in extraversion in the treatment group. The effect is negative and insignificant in the control group. Hence, greater the perceived difference in extraversion, higher the chances of best responding by the player in the treatment group. Alternatively, this implies that greater the *perceived similarity* between the player and their partner, lower are the chances of the player best responding in the treatment group. This result is consistent with

hypothesis 2a which supports the perceived similarity hypothesis. When the perceived difference in extraversion is small, the player believes that their partner will act similar to themselves. This makes it harder to out-think or out-reason the opponent, thus reducing the probability of best responding. This result holds only when the players engage in small talk as otherwise the players have nothing to base their personality beliefs on and so absent small talk, their beliefs are unlikely to affect decision making.

The results hold even after controlling for the player's IQ and eyes test score, the player's beliefs about partner's IQ and other controls - player's age, gender, risk aversion and the order of games played. In the control group, increase in the player's IQ by 1 standard deviation increases the probability of best responding by 6 percentage points where as increase in beliefs about the partner's IQ decreases the probability of best responding by 5 percentage points. The player's own neuroticism measure also has a significantly (p-value < 0.10) positive effect on the probability of best responding in the control group.[13] The relationship between level choice and perceived difference in extraversion is depicted in Figure 1.4.



(a) (b)

Figure 1.4: Perceived differences in the players and their partners' extraversion, and level-k choices made. (a) Effect of perceived difference in extraversion on level choice in control and treatment groups. The figure shows that perceived difference in extraversion has a significant negative effect on the player's level-k choice in the treatment group. (b) shows that the effect of small talk treatment on probability of best responding to the distribution of level beliefs increases as the perceived difference in extraversion increases.

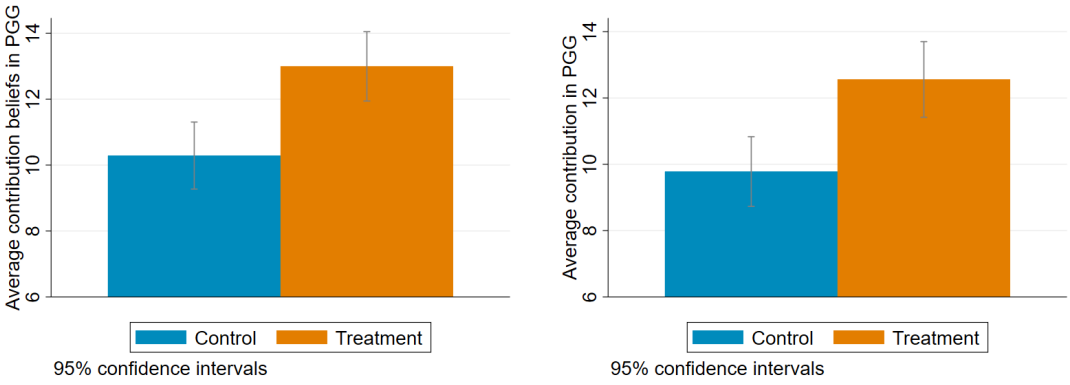
Result 2b: Cooperation and extraversion beliefs

Next, we examine the results of the public goods game to test *hypothesis 2b* which states that a player's cooperation in the game will increase with their beliefs

[13]Note that the results are robust to the inclusion of personality beliefs as control variables, which are omitted here for parsimony, but are presented in Table 1.A.6. The results also remain similar when a logit model is used instead of probit as shown in Table 1.A.7.

about their opponent's extraversion, since the player will expect an extraverted opponent to cooperate more. Of the two fundamental personality traits, we expect extraversion to be especially relevant for the public goods game, since it is extraversion that is most associated with pro-social behaviours (Carlo et al., 2005; Burke and Hall, 1986).[14]

In the public goods game, the average beliefs about partner's contribution in the treatment group was 13 experimental pounds (EP), where as in the control group it was 10.3 EP. This difference is statistically significant with p-value < 0.01 and a t-statistic of -3.640. The average contribution in the treatment group was 12.6 EP, whereas in the control group it was 9.8 EP. This difference is statistically significant with p-value < 0.01 and a t-statistic of -3.525 (Figure 1.5). The Kolmogorov-Smirnov tests for equality of distributions of own contribution as well as beliefs about partner's contribution between the treatment and control groups were rejected with p-value < 0.01 for both. This is consistent with the existing literature which finds that pre-game communication of any form increases cooperation rates (Dawes et al., 1977; Bochet et al., 2006).



(a) Contribution Belief (b) Own-Contribution

Figure 1.5: (a) Average Beliefs about Partner's Contribution and (b) Average Contribution in the Public Goods Game

Our analysis for the public goods game will only consider the observations in which the subjects played the public goods game before the level-k reasoning game. The rationale for this is that playing the level-k game first seems to trigger level-k reasoning (Georganas et al., 2015), thus biasing decision-making in the social preferences task. On the other hand, since the level-k game strictly requires level-k reasoning, without invoking any social preferences (a point made explicitly in Arad and Rubinstein, 2012), the results of the 11-20 game are not biased

[14]We also see from Table 1.A.8 that beliefs about partner's neuroticism has no significant effect on decision making in the public goods game.

by playing the public goods game first. Further, treated subjects contribute significantly more on average compared to control group subjects, only when the public goods game is played first, where as the difference is insignificant when the public goods game is played second (Figure 1.A.2). The results from the public goods game, for those who played the 11-20 game first are presented in Figure 1.A.3 and Table 1.A.11.

We examine *hypothesis 2b* using equation 1.3. $Choice_i$ is player i's choice (or contribution) in the public goods game, $pers_i$ is player i's personality, $E_i(pers_j)$ is player i's beliefs about partner j's personality, z_i are individual characteristics of i and ε_i is an idiosyncratic error term.

$$Choice_i = \beta_1 pers_i + \beta_2 E_i(pers_j) + \gamma z_i + \varepsilon_i \tag{1.3}$$

$$E_i(pers_j) = \lambda_1 pers_j + \lambda_2 pers_i + \rho z_i + \epsilon_i \tag{1.4}$$

Players' tendency to project their own extraversion onto their partners creates an endogeneity issue (result 1), and as such estimation of equation 1.3 requires valid instruments. Beliefs about partner's extraversion depend on two components - the player's own extraversion and the partner's true extraversion, as discussed in section 1.3.1. These two components are independent as the two players are randomly matched. Therefore, beliefs about partner's extraversion can be instrumented with the partner's true extraversion. Equation 1.4 is the first stage. $pers_j$ is the partner j's true personality.

The first stage results presented in Table 1.6 show that partner's true extraversion significantly enhances beliefs about partner's extraversion in the treatment, but not in the control group, since in the control group the player has no interaction with their partner.[15] Table 1.7 presents the results of a two-stage least squares instrumental variable (IV) regression for the treatment group. Since the endogeneity bias only exists for the treatment group, equation 1.3 is estimated without an instrumental variable for the control group, and is presented in columns 1 and 2 of Table 1.7.

Columns 3 and 4 of Table 1.7 show that in the treatment group, an increase in 1 standard deviation in extraversion belief, increases beliefs about partner's con-

[15]To test for weak instruments, a Wald test is conducted, which tests the null that the coefficients of the endogenous regressors are zero. The null for the treatment group, is rejected at the 5% level. This suggests that weak instruments are not an issue here. Further, the F-statistic in the first stage regression (for two-stage least squares) is greater than 10, which indicates that the instruments are strong (Staiger and Stock, 1997) for the treatment group.

Table 1.6: First Stage: Extraversion beliefs and Public goods game

	Control		Treatment	
	(1) Extraversion Belief	(2) Extraversion Belief	(3) Extraversion Belief	(4) Extraversion Belief
Own Extraversion	0.0299	0.0333	0.2147**	0.2614**
	(0.086)	(0.102)	(0.106)	(0.103)
Partner's Extraversion	-0.1015	-0.0977	0.3541***	0.3648***
	(0.081)	(0.092)	(0.093)	(0.094)
Own IQ		-0.1034		0.0121
		(0.103)		(0.102)
IQ Belief		-0.0559		0.0166
		(0.147)		(0.095)
Eyes Test Score		-0.0470		0.1195
		(0.107)		(0.073)
Controls	No	Yes	No	Yes
N	110	110	106	106

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
'Controls' refers to the player's age, gender and risk aversion.

tribution and own-contribution by 0.6 and 0.5 standard deviations, respectively (p-value < 0.05 for both). On the other hand, an increase in 1 standard deviation in own-extraversion decreases beliefs about partner's contribution, as well as the player's own-contribution by 0.3 (p-value < 0.05) and 0.2 (insignificant) standard deviations, respectively. Thus, beliefs about partner's extraversion has a positive and relatively larger effect, compared to own-extraversion, on decision-making in the public goods game in the treatment group. For the control group, column 2 shows that the player's extraversion significantly (p-value < 0.05) and negatively impacts contribution level. Beliefs about partner's extraversion has no significant effect on both beliefs about partner's contribution and own-contribution in the control group (which makes perfect sense since in the control group, where there is no interaction, players have no basis upon which to form sensible beliefs about their partners). Columns 3 and 4 can essentially be summarised as showing that there are two forces at work in determining how the contribution level is affected by extraversion: a direct and negative effect of own-extraversion, and an indirect and positive effect that works through beliefs about the partner's extraversion. Overall, the role of beliefs seems stronger than own-extraversion though both

Table 1.7: Impact of beliefs about partner's personality and own personality on beliefs about partner's contribution and own contribution in Public Goods Game

	Control OLS		Treatment IV	
	(1) Contribution Belief	(2) Own Contribution	(3) Contribution Belief	(4) Own Contribution
Extraversion Belief	0.0601	0.1110	0.6091**	0.5184**
	(0.082)	(0.092)	(0.264)	(0.262)
Own Extraversion	-0.0733	-0.2041**	-0.3074**	-0.2018
	(0.095)	(0.088)	(0.134)	(0.138)
Own IQ	-0.0583	-0.0417	0.0856	0.1548
	(0.096)	(0.084)	(0.094)	(0.103)
IQ Belief	0.1250	0.1140	0.0871	0.2402***
	(0.091)	(0.100)	(0.086)	(0.088)
Eyes Test Score	-0.0431	-0.0015	0.1043	0.1502
	(0.096)	(0.118)	(0.117)	(0.139)
Controls	Yes	Yes	Yes	Yes
N	110	110	106	106

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Controls' refers to the player's age, gender and risk aversion.

are important.[16] Moreover, consistent with *hypothesis 2b*, we find that players cooperate more in the public goods game when they believe their partners to be extraverted.

Following Soto and John, 2009, we divide extraversion of the player into 2 facets, assertiveness and activity.[17] This is done to examine which particular facet of extraversion is responsible for driving cooperation decisions. While assertiveness can be defined as preference for exerting control in a group setting (Soto and John, 2012), activity (or enthusiasm) describes both positive emotions and outgoing friendliness or sociability (DeYoung et al., 2007). The facet analysis (Table 1.A.10) revealed that of the 2 facets of extraversion, it is facet assertiveness which

[16]Estimating equation 1.3 for the treatment group using OLS, and not an IV approach, yields similar results where, in the treatment group, beliefs about partner's extraversion has a significant positive effect on both beliefs about partner's contribution as well as own contribution in the public goods game and own-extraversion has an insignificant negative impact on both (Table 1.A.9). However, given the scope for endogeneity bias, the IV approach is likely to be more appropriate.

[17]Soto and John, 2009 propose forming 10 facet scores, 2 for each of the Big Five traits, by dividing the 44 items in the BFI questionnaire. Assertiveness and activity facet scores are formed for each individual based on their responses to specific items in the BFI.

is responsible for the negative effect of the player's extraversion on beliefs about partner's contribution, as well as own contribution in the public goods game.

1.3.3 Result 3: Talkativeness and extraversion beliefs

In this section we will consider how a short period of small talk can work to foster belief formation with a direct examination of the text used by participants in the experiment. In particular, in line with *hypothesis 3* we will focus on the effect of the number of words spoken by the partner on beliefs to check whether more talkative players are indeed believed to be extraverted. The number of words is perhaps the easiest language characteristic to calculate. We also examine the scores for three affective or emotional components of the partner's language use, namely *valence*, *arousal* and *dominance*, using the score-ratings proposed by Warriner et al., 2013. The valence rating of a word refers to the pleasant emotion conveyed by a word, with the rating increasing as it moves from unhappy to happy. Arousal rating of a word increases with the degree of excitement emoted by it. Finally, the dominance rating of a word increases with the degree to which it conveys the emotion of *being in control*.[18]

Table 1.8 reports the results for the quantitative language characteristics that we consider. The dependant variables are beliefs about the partner's fundamental personality traits. Column 1 shows that beliefs about partner's extraversion increase with the number of words spoken by the partner (p-value < 0.01), consistent with *hypothesis 3* and the findings of Mehl et al., 2006 who also find talkative subjects are rated as more extraverted. The coefficient for number of words remains similar even after adding valence, arousal and dominance as explanatory variables in column 2. Column 3 shows that the result persists even after controlling for the player's IQ, eyes test score, age, gender, beliefs about partner's IQ, a dummy for non-native speaker (equals 1 if the player is a non-native English speaker and 0 otherwise) and a dummy for first speaker (equals 1 if the player started the conversation and 0 otherwise). Columns 4-6 show that

[18]We might also consider the choice of words used by participants. Figure 1.A.4 shows a word cloud of the words spoken by the subjects during the pre-game small talk communication which depicts the very general and trivial nature of small talk. Figure 1.A.5 attempts to distinguish between the most frequently used words by subjects believed to have different personalities. Through a simple examination of word usage, it's hard to distinguish between the nature of language used by subjects believed to have different personalities. Those who are believed to be highly extraverted (believed to have above median extraversion scores) have a similar set of most frequently used words when compared to those who are believed to be less extraverted (believed to have below median extraversion scores) which are likely to reflect the social norms of small talk (Figures 1.A.5 (a) and (b)). Figures 1.A.5 (c) and (d) show a similar story for neuroticism beliefs. This is not surprising given the unstructured nature of the small talk but we know from our results and experimental design that language is playing an important role, so we will focus on more quantitative measures in this section.

Table 1.8: Impact of number of words and emotional content of the text spoken by the partner on beliefs about partner's personality

	Extraversion Belief			Neuroticism Belief		
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Words	0.2744***	0.2604***	0.2355***	-0.0573	-0.0368	-0.0455
	(0.079)	(0.076)	(0.078)	(0.072)	(0.076)	(0.073)
Valence		-0.2723	-0.2545		-0.0578	0.0328
		(0.233)	(0.234)		(0.263)	(0.267)
Arousal		0.1718	0.2191		-0.0763	-0.1241
		(0.149)	(0.147)		(0.116)	(0.115)
Dominance		0.1918	0.1728		0.0361	-0.0407
		(0.256)	(0.253)		(0.256)	(0.255)
Own IQ			-0.1178			0.1278
			(0.086)			(0.079)
Eyes Test Score			0.0603			0.0276
			(0.058)			(0.097)
Age			0.0227			-0.0429**
			(0.022)			(0.020)
Female			-0.0659			-0.1659
			(0.157)			(0.157)
IQ Belief			0.1327			-0.0978
			(0.081)			(0.086)
Non-Native Speaker			0.3788**			-0.2491
			(0.151)			(0.158)
First Speaker			-0.0036			-0.3199**
			(0.140)			(0.152)
N	168	168	168	168	168	168

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The explanatory variables, namely number of words, and valence, arousal and dominance ratings are standardised for comparability across coefficients.

beliefs about partner's neuroticism decrease with the number of words spoken by the partner, although the impact is insignificant. Valence, arousal and dominance ratings did not have a significant impact on beliefs about either of the two fundamental personality traits. We also consider whether the beliefs formed by examining the number of words used in communication provide an *accurate* picture of someone's true personality type. What we see from the results in Table 1.A.12 is that extraverts genuinely do seem to use more words, a result which is significant (p-value < 0.05), with and without the addition of valence, arousal and dominance as explanatory variables and a list of sensible control variables.

1.4 Discussion

While the association between an individual's personality and strategic behaviour has been well established within Economics (Proto, Rustichini, and Sofianos, 2019; Proto and Rustichini, 2014; Rustichini et al., 2016; Johnson et al., 2009; Hirsh and Peterson, 2009), what remains unexplored is the impact of perceptions about another individual's personality and how these might influence subsequent strategic interactions. We examine this by providing subjects the opportunity to develop beliefs about the fundamental personality traits of their partner in a controlled laboratory setting, through a brief (4-minute) chat via their computer screens. We label this type of communication as *small talk* since there is no prior knowledge of the definitive rules of any future strategic interaction between the pair and also based on our observations of the nature of the communication. Following the short period of small talk and subsequent personality belief elicitation, the pair engage in two well-known one-shot strategic decision making tasks: the 11-20 money request game (Arad and Rubinstein, 2012), which examines level-k choices, and the public goods game, which is a game of cooperation. We examine the causal effect of beliefs about the partner's personality on decisions made in the two tasks through comparison with a control group, where subjects participated in an independent placebo task instead of engaging in small talk.

An examination of personality beliefs in section 1.3.1 revealed that beliefs about an individual's extraversion are not only determined by their true extraversion trait, but are also enhanced by the predictor's extraversion: extraverts tend to believe that their partners are also extraverted. This finding lends support to our first hypothesis which states that while formulating beliefs about someone's personality, individuals tend to project their own traits. This effect is significantly stronger in the treatment group than in the control group. This links closely with the psychological literature on extraversion: an extraverted person, who is subject to positive emotions, fosters a positive social environment around them and projects their extraversion or sociability onto others (Eaton and Funder, 2003; Thorne, 1987), making them prone to *complementary self projection bias*. No such projection was observed for neuroticism.[19] We also found that personality beliefs developed about a partner, after engaging in small talk with them, were only a reliable or accurate measure of the partner's extraversion, but not their neuroticism. This is consistent with the findings of Eaton and Funder,

[19]This is contradictory to the theory of *neurotic projection* which is a form of defence mechanism through which people tend to project negative feelings, motives or behaviour they might possess and are uncomfortable with, onto others. This can be attributed to the negative connotations of the trait neuroticism. Individuals are less keen to project trait neuroticism as it is likely to draw attention to their own neuroticism.

2003 who also observed accurate perceptions about a stranger's extraversion after a 5-minute in-person face-to-face conversation.

In section 1.3.2 *result 2a*, we observed that the perceived similarity or difference between the personalities of the players and their partners influenced decision making in level-k reasoning games, consistent with *hypothesis 2a*. Particularly, perceived differences in the pair's extraversion traits inversely affect the player's level-k choice, an effect which is significantly stronger in the treatment group, compared to the control. In level-k reasoning games a player's strategy reflects the player's beliefs about the opponent's type. The player best responds to these beliefs, attempting to out-reason or out-think their opponent. In accordance with simulation-based theories of mental modelling, perceptions anchor onto own-reasoning processes and likely choices and are then adjusted for any discrepancy between self and other, while inferring choices of similar others (Tamir and Mitchell, 2013). Thus, in level-k games, the perceived similarity or differences between the type of player and their partner, play a crucial role in predicting how the opponent might behave and in turn determine own strategy choice. The *perceived similarity hypothesis* (Thomas et al., 2014) states that when a player thinks they are faced by a similar opponent, they believe the opponent will reason and act in ways similar to themselves. Thus, when the player assumes the partner's type is similar to their own, it becomes harder for them to out-reason the partner in the level-k game. When faced by a similar other, player believes that the opponent, undergoing the same thinking process, will reason harder and pick a higher level which in turn should make the player choose a higher level as well. Consequently, when the player suspects their partner's type is similar to their own, the probability of them best responding to the distribution of level-k beliefs falls. This result holds only when the players engage in small talk as in the control condition the player has no reliable indicator of perceived similarity with the opponent.

In section 1.3.2 *result 2b*, we found that when a player thinks that their opponent is extraverted, they believe that their opponent will cooperate more, a result only observed in the small talk treatment. The result that extraverts are expected to cooperate more in social situations, is consistent with the finding in psychology that higher levels of the extraversion trait are associated with prosocial behaviour (Carlo et al., 2005; Burke and Hall, 1986). Thus, the player themselves cooperate, expecting cooperation from their opponent. In contrast, the literature is conflicted on the effect of a subject's own extraversion on cooperation. While Hirsh and Peterson, 2009 and Ross et al., 2003 find a positive effect of extraversion on cooperation, Koole et al., 2001 find the opposite. Hirsh and

Peterson, 2009 posit that individuals who score highly on the enthusiasm facet of extraversion, owing to their positive outlook, view cooperation as rewarding and expect cooperative behaviour from their partners as well. The opposing argument is that introverts, and not extraverts, are likely to cooperate more as they are more inclined to avoid conflicts (Koole et al., 2001). This paper supports the latter argument. We would also argue that some of the contradictions seen in the literature stem from missing the subtle interactions with beliefs that are highlighted in our results. Further, this negative effect of extraversion is driven by the assertive facet of an extravert's personality. Lastly, beliefs about opponent's extraversion have a relatively larger effect on decision-making in the public goods game than own-extraversion. Since these effects work in opposite directions they may partly explain the apparent contradictions seen in the general literature on extraversion and cooperation since they only become apparent when we disentangle the impact of beliefs and own-characteristics.

Consistent with *hypothesis 2*, we show that beliefs about a partner's personality - specifically beliefs about partner's extraversion - developed after engaging in small talk, significantly impact choices made in subsequent strategic interactions. The reason why extraversion plays a big role in our study is likely because, out of the two fundamental personality traits, subjects could only form reasonably accurate beliefs about the partner's extraversion. Extraverts, characterised by their sociability, enthusiasm and gregariousness, tend to stand out by nature, making extraversion the most detectable trait, especially after a brief chat. Extraversion, as one of the principal dimensions of personality, can explain a wide variety of outcomes, such as subjective well-being measures (Costa and McCrae, 1980), health outcomes (Lai and Qin, 2018), relationship satisfaction (Tov et al., 2016) and occupational choices (King et al., 2017). This may explain why beliefs about the extraversion of a partner is crucial for explaining strategic behaviour in our study.

The brief period of small talk, as the key experimental manipulation in our study, was the only opportunity for the players to interact and hence the primary basis for developing personality beliefs. Analysis of the text data from the chat in Section 1.3.3 revealed that partners who talk more during the chat are believed to be extraverted, as expressed in our final hypothesis. This result is in keeping with Mehl et al., 2006 who in a study observed that personality judges rated talkative participants as more extraverted, since extraversion is characterised by attributes such as sociability and talkativeness (Goldberg, 1990; Costa and McCrae, 1992).

Overall, we find evidence suggesting that beliefs about a partner's fundamental personality traits, particularly extraversion, are a significant determinant of decisions made in any subsequent strategic interaction with them. This impact of beliefs on choices can either be through the absolute value of beliefs about partner's extraversion (as in the public goods game) or the perceived differences in the pair's extraversion (as in the 11-20 money request game). We hope that our study might open avenues for future research exploring how beliefs about other's personality traits affect choices made in various strategic interactions with them. Being the first study of its kind, our work is limited in scope due to the controlled laboratory setting and limited communication time, making it impossible to examine the role of the remaining personality traits, agreeableness, conscientiousness and openness, and even intelligence. In order to give these traits a more reasonable chance of playing a role, a longer, more sustained series of small talk conversations, something more akin to what occurs in the real-world seems sensible, and this is only going to be feasible in the setting of a field experiment. Our hope is that our results will give impetus to new research that looks at repeated interactions in a more realistic setting.

1.5 Conclusion

The link between personality and strategic behaviour has garnered much attention in recent Economic literature. We expand on this by providing evidence of the impact of impressions about others' personalities on subsequent strategic interactions with them. In a laboratory setting we show that, when subjects engage in brief small talk interaction with strangers via an instant messaging software, they develop beliefs about the stranger's personality traits, particularly extraversion, which affect their ensuing strategic behaviour. Extraverts, who are characterised by sociability and gregariousness, tend to be distinctive by nature, making extraversion the most detectable trait in a short bout of communication. Perceptions of trait extraversion, thus, played a crucial role in two well-known strategic decision making tasks - the 11-20 money request game which examines level-k reasoning and the public goods game which is a game of cooperation. Analysis of the pre-game interaction revealed that subjects use the number of words spoken as a a mechanism for detecting extraverts, which does indeed provide a reasonably accurate forecast of type. However, perceptions about extraversion can be coloured by *complementary self projection bias* which makes extraverts prone to projecting their extraversion or positive affect onto those they interact with. Overall, we hope that this study paves the way for future research exploring the association between personality impressions and strategic

behaviour in a variety of tasks and real-world contexts.

References

Albright, L., Kenny, D., & Malloy, T. (1988). Consensus in personality judgments at zero acquaintance. *Journal of personality and social psychology*, *55*(3), 387–395. https://doi.org/10.1037//0022-3514.55.3.387

Arad, A., & Rubinstein, A. (2012). The 11-20 money request game: A level-k reasoning study. *American Economic Review*, *102*(7), 3561–73. https://doi.org/10.1257/aer.102.7.3561

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry*, *42*(2), 241–251. https://doi.org/10.1111/1469-7610.00715

Blais, A.-R., & Weber, E. U. (2006). A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision Making*, *1*(1), 33. https://doi.org/10.1037/t13084-000

Bochet, O., Page, T., & Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior and Organization*, *60*(1), 11–26. https://doi.org/10.1016/j.jebo.2003.06.006

Burke, D. M., & Hall, M. (1986). Personality characteristics of volunteers in a companion for children program. *Psychological Reports*, *59*(2), 819–825. https://doi.org/10.2466/pr0.1986.59.2.819

Bursztyn, L., Ferman, B., Fiorin, S., Kanz, M., & Rao, G. (2017). Status Goods: Experimental Evidence from Platinum Credit Cards. *The Quarterly Journal of Economics*, *133*(3), 1561–1595. https://doi.org/10.1093/qje/qjx048

Canli, T. (2004). Functional brain mapping of extraversion and neuroticism: learning from individual differences in emotion processing. *Journal of personality*, *72*(6), 1105–1132. https://doi.org/10.1111/j.1467-6494.2004.00292.x

Carlo, G., Okun, M. A., Knight, G. P., & de Guzman, M. R. T. (2005). The interplay of traits and motives on volunteering: Agreeableness, extraversion and prosocial value motivation. *Personality and Individual Differences*, *38*(6), 1293–1305. https://doi.org/10.1016/j.paid.2004.08.012

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. *Econometrica*, *74*(6), 1579–1601. https://doi.org/10.1111/j.1468-0262.2006.00719.x

Cooper, R., DeJong, D. V., Forsythe, R., & Ross, T. W. (1992). Communication in Coordination Games. *The Quarterly Journal of Economics*, *107*(2), 739–771. https://doi.org/10.2307/2118488

Coricelli, G., & Nagel, R. (2009). Neural correlates of depth of strategic reasoning in medial prefrontal cortex. *Proceedings of the National Academy of Sciences*, *106*(23), 9163–9168. https://doi.org/10.1073/pnas.0807721106

Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678. https://doi.org/10.1037//0022-3514.38.4.668

Costa, P. T., & McCrae, R. R. (1992). *Neo PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: an experimental study. *Econometrica*, *69*, 1193–1235. https://doi.org/10.1111/1468-0262.00239

Das, S. R., & Chen, M. Y. (2007). Yahoo! for amazon: Sentiment extraction from small talk on the web. *Management science*, *53*(9), 1375–1388. http://www.jstor.org/stable/20122297

Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, *35*, 1–11.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of personality and social psychology*, *93*(5), 880–896. https://doi.org/10.1037/0022-3514.93.5.880

Eaton, L. G., & Funder, D. C. (2003). The creation and consequences of the social world: An interactional analysis of extraversion. *European Journal of Personality*, *17*(5), 375–395. https://doi.org/10.1002/per.477

Eysenck, H. J. (1947). *Dimensions of personality*. Kegan Paul.

Fe, E., Gill, D., & Prowse, V. (2019). *Cognitive Skills and the Development of Strategic Sophistication* (CAGE Online Working Paper Series No. 448). Competitive Advantage in the Global Economy (CAGE). https://ideas.repec.org/p/iza/izadps/dp11326.html

Fehr, E., & Gaechter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. https://doi.org/10.1038/415137a

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171–178.

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory*, *159*, 369–400. https://doi.org/10.1016/j.jet.2015.07.012

Gill, D., & Prowse, V. (2016). Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy*, *124*(6), 1619–1676. https://doi.org/10.1086/688849

Goldberg, L. R. (1990). An alternative "description of personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*(6), 1216–1229. https://doi.org/10.1037/0022-3514.59.6.1216

Hall, J. A., Andrzejewski, S. A., Murphy, N. A., Mast, M. S., & Feinstein, B. A. (2008). Accuracy of judging others' traits and states: Comparing mean levels across tests. *Journal of Research in Personality*, *42*(6), 1476–1489. https://doi.org/10.1016/j.jrp.2008.06.013

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science*, *319*(5868), 1362–1367. https://doi.org/10.1126/science.1153808

Hirsh, J. B., & Peterson, J. B. (2009). Extraversion, neuroticism, and the prisoner's dilemma. *Personality and Individual Differences*, *46*(2), 254–256. https://doi.org/10.1016/j.paid.2008.10.006

Hoffman, C., Mischel, W., & Mazze, K. (1981). The role of purpose in the organization of information about behavior: Trait-based versus goal-based categories in person cognition. *Journal of Personality and Social Psychology*, *40*(2), 211–225. https://doi.org/10.1037/0022-3514.40.2.211

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, *2*(1999), 102–138.

Johnson, M. K., Rustichini, A., & MacDonald, A. W. (2009). Suspicious personality predicts behavior on a social decision-making task. *Personality and Individual Differences*, *47*(1), 30–35. https://doi.org/10.1016/j.paid.2009.01.050

King, D. D., Ott-Holland, C. J., Ryan, A. M., Huang, J. L., Wadlington, P. L., & Elizondo, F. (2017). Personality Homogeneity in Organizations and Occupations: Considering Similarity Sources. *Journal of Business and Psychology*, *32*(6), 641–653. https://doi.org/10.1007/s10869-016-9459-4

Koole, S. L., Jager, W., van den Berg, A. E., Vlek, C. A., & Hofstee, W. K. (2001). On the Social Nature of Personality: Effects of Extraversion, Agreeableness, and Feedback About Collective Resource Use on Cooperation in a Resource Dilemma. *Personality and Social Psychology Bulletin*, *27*(3), 289–301. https://doi.org/10.1177/0146167201273003

Lai, D. W. L., & Qin, N. (2018). Extraversion personality, perceived health and activity participation among community-dwelling aging adults in Hong

Kong. *PloS one*, *13*(12), 1–13. https://doi.org/10.1371/journal.pone.0209154

Little, A. C., & Perrett, D. I. (2007). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology*, *98*(1), 111–126. https://doi.org/10.1348/000712606X109648

McCrae, R. R., & Costa, P. T. (1999). A Five-Factor theory of personality. *Handbook of personality: Theory and research, 2nd ed.* (pp. 139–153). Guilford Press.

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, *90*(5), 862–877. https://doi.org/10.1037/0022-3514.90.5.862

Moskowitz, G. B., & Olcaysoy Okten, I. (2016). Spontaneous goal inference (SGI). *Social and Personality Psychology Compass*, *10*(1), 64–80. https://doi.org/10.1111/spc3.12232

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, *85*(5), 1313–1326. http://www.jstor.org/stable/2950991

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance [PMID: 19762717]. *Personality and Social Psychology Bulletin*, *35*(12), 1661–1671. https://doi.org/10.1177/0146167209346309

Nettle, D., & Liddle, B. (2008). Agreeableness is related to social-cognitive, but not social-perceptual, theory of mind. *European Journal of Personality*, *22*(4), 323–335. https://doi.org/10.1002/per.672

Proto, E., & Rustichini, A. (2014). Cooperation and Personality. *The Warwick Economics Research Paper Series (TWERPS)*, (1045). https://ideas.repec.org/p/wrk/warwec/1045.html

Proto, E., Rustichini, A., & Sofianos, A. (2019). Intelligence, Personality, and Gains from Cooperation in Repeated Interactions. *Journal of Political Economy*, *127*(3), 1351–1390. https://doi.org/10.1086/701355

Proto, E., Sgroi, D., & Nazneen, M. (2019). Happiness, cooperation and language. *Journal of Economic Behavior & Organization*, *168*, 209–228. https://doi.org/10.1016/j.jebo.2019.10.006

Pullin, P. (2010). Small talk, rapport, and international communicative competence: Lessons to learn from belf. *The Journal of Business Communication (1973)*, *47*(4), 455–476. https://doi.org/10.1177/0021943610377307

Ragan, S. (2014). Sociable talk in women's health care contexts: two forms of non-medical talk. In J. Coupland (Ed.), *Small talk* (pp. 289–307). Routledge. https://doi.org/10.4324/9781315838328-15

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*(1), 203–212. https://doi.org/10.1016/j.jrp.2006.02.001

Raven, J. (2003). Raven progressive matrices. In R. S. McCallum (Ed.), *Handbook of nonverbal assessment* (pp. 223–237). Springer US. https://doi.org/10.1007/978-1-4615-0153-4_11

Ross, S. R., Rausch, M. K., & Canada, K. E. (2003). Competition and cooperation in the five-factor model: Individual differences in achievement orientation. *The Journal of psychology*, *137*(4), 323–37. https://doi.org/10.1080/00223980309600617

Rustichini, A., DeYoung, C. G., Anderson, J. E., & Burks, S. V. (2016). Toward the integration of personality theory and decision theory in explaining economic behavior: An experimental investigation. *Journal of Behavioral and Experimental Economics*, *64*, 122–137. https://doi.org/10.1016/j.socec.2016.04.019

Sgroi, D. (2018). Theory of mind and strategic decision-making. *AEA RCT Registry*. http://www.socialscienceregistry.org/trials/2903/history/29367

Soto, C. J., & John, O. P. (2009). Ten facet scales for the Big Five Inventory: Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, *43*(1), 84–90. https://doi.org/10.1016/j.jrp.2008.10.002

Soto, C. J., & John, O. P. (2012). Development of big five domains and facets in adulthood: Mean-level age trends and broadly versus narrowly acting mechanisms. *Journal of Personality*, *80*(4), 881–914. https://doi.org/10.1111/j.1467-6494.2011.00752.x

Stahl, D., & Wilson, P. (1995). On Players' Models of Other Players: Theory and Experimental Evidence. *Games and Economic Behavior*, *10*(1), 218–254. https://EconPapers.repec.org/RePEc:eee:gamebe:v:10:y:1995:i:1:p:218-254

Stahl, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, *25*(3), 309–327. https://doi.org/10.1016/0167-2681(94)90103-1

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, *65*(3), 557–586. http://www.jstor.org/stable/2171753

Stiller, J., & Dunbar, R. I. (2007). Perspective-taking and memory capacity predict social network size. *Social Networks*, *29*(1), 93–104. https://doi.org/10.1016/j.socnet.2006.04.001

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General*, *142*(1), 151–162. https://doi.org/10.1037/a0028232

Thomas, K., DeScioli, P., Sultan Haque, O., & Pinker, S. (2014). The Psychology of Coordination and Common Knowledge. *Journal of Personality and Social Psychology*, *107*(4), 657–676. https://doi.org/10.1037/a0037037

Thorne, A. (1987). The press of personality: A study of conversations between introverts and extraverts. *Journal of Personality and Social Psychology*, *53*(4), 718–726. https://doi.org/10.1037/0022-3514.53.4.718

Tov, W., Nai, Z. L., & Lee, H. W. (2016). Extraversion and Agreeableness: Divergent Routes to Daily Satisfaction With Social Relationships. *Journal of Personality*, *84*(1), 121–134. https://doi.org/10.1111/jopy.12146

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45 VN - r*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Watson, D., & Clark, L. A. (1992). On traits and temperament: General and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality*, *60*(2), 441–476. https://doi.org/10.1111/j.1467-6494.1992.tb00980.x

Watson, D., Wiese, D., Vaidya, J., & Tellegen, A. (1999). The two general activation systems of affect: Structural findings, evolutionary considerations, and psychobiological evidence. *Journal of Personality and Social Psychology*, *76*(5), 820–838. https://doi.org/10.1037/0022-3514.76.5.820

Appendix

1.A Additional Tables and Figures

Table 1.A.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Own Extraversion	3.372	0.814	1.25	5	338
Own Neuroticism	2.935	0.811	1	5	338
Extraversion Belief	3.499	0.827	1	5	338
Neuroticism Belief	2.818	0.865	1	5	338
Perceived diff Extraversion	0.882	0.689	0	3.25	338
Perceived diff Neuroticism	0.899	0.714	0	3.125	338
Level Chosen in 11-20 game	2.891	2.522	0	9	338
Level Belief in 11-20 game	2.787	2.566	0	9	338
Own Contribution in PGG	11.163	7.363	0	20	338
Contribution Belief in PGG	11.633	6.956	0	20	338
Own IQ	18.604	4.464	4	28	338
IQ Belief	18.213	4.825	1	30	338
Eyes Test Score	27.817	3.759	11	35	338
Age	21.154	3.622	17	42	338
Risk Aversion	4.317	0.767	1.533	6	338
Female	0.615	0.487	0	1	338
Non-native English speaker	0.349	0.477	0	1	338

Table 1.A.2: Balance Test for Treatment and Control groups

Variable	(1) Control Mean/SE	(2) Treatment Mean/SE	T-test P-value
Own Extraversion	3.421	3.322	0.263
	(0.064)	(0.061)	
Own Neuroticism	2.953	2.917	0.688
	(0.061)	(0.064)	
Eyes Test Score	27.471	28.167	0.089*
	(0.252)	(0.322)	
Age	20.935	21.375	0.265
	(0.276)	(0.280)	
Female	0.653	0.577	0.154
	(0.037)	(0.038)	
Own IQ	18.206	19.006	0.100*
	(0.343)	(0.342)	
Risk Aversion	4.286	4.347	0.465
	(0.060)	(0.058)	
Non-native English speaker	0.359	0.339	0.707
	(0.037)	(0.037)	
N	170	168	
F-test of joint significance (p-value)			0.236
Number of observations			338

Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.A.3: Overestimation and inaccuracy of personality beliefs

	Overestimation of Extraversion Belief		Inaccuracy of Extraversion Belief	Overestimation of Neuroticism Belief		Inaccuracy of Neuroticism Belief
	(1)	(2)	(3)	(4)	(5)	(6)
OwnExtraversion × Treatment	0.1601*	0.2170**	-0.0132	-0.0760	-0.0954	0.0489
	(0.086)	(0.092)	(0.112)	(0.092)	(0.100)	(0.118)
OwnNeuroticism × Treatment	0.1040	0.1121	0.1135	-0.0404	-0.0321	-0.0213
	(0.093)	(0.096)	(0.119)	(0.085)	(0.083)	(0.117)
PartnerExtraversion × Treatment	0.3031***	0.3075***	-0.3722***			
	(0.079)	(0.081)	(0.124)			
PartnerNeuroticism × Treatment				0.0169	-0.0004	0.4717***
				(0.078)	(0.078)	(0.131)
Eyes Test Score × Treatment	0.0663	0.0773	-0.1817*	0.1146	0.1503	-0.0833
	(0.072)	(0.072)	(0.105)	(0.099)	(0.102)	(0.125)
Own Extraversion	0.0101	0.0181	0.1391*	-0.0740	-0.0552	-0.0261
	(0.052)	(0.059)	(0.082)	(0.058)	(0.058)	(0.073)
Own Neuroticism	-0.0061	0.0006	0.0144	0.0343	0.0461	-0.0194
	(0.062)	(0.064)	(0.094)	(0.062)	(0.061)	(0.091)
Partner's Extraversion	-0.8160***	-0.8189***	-0.0259			
	(0.052)	(0.055)	(0.091)			
Partner's Neuroticism				-0.6530***	-0.6395***	-0.1162
				(0.053)	(0.054)	(0.098)
Eyes Test Score	-0.0459	-0.0368	0.1731**	-0.0930	-0.1352*	-0.0816
	(0.054)	(0.057)	(0.080)	(0.070)	(0.074)	(0.090)
Treatment	0.2609***	-0.2290	1.0108	-0.3866***	-0.1525	-0.4028
	(0.071)	(0.463)	(0.613)	(0.080)	(0.423)	(0.576)
Controls	No	Yes	Yes	No	Yes	Yes
N	338	338	338	338	338	338

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependant variable, Overestimation of personality beliefs, is computed by taking the difference between the player's beliefs about their partner's personality and the partner's true personality scores. This difference is then standardised. The dependent variable is thus a measure of exaggeration or overestimation of the partner's personality by the player. The dependent variable, Inaccuracy of personality beliefs, is computed by taking the absolute difference between the player's beliefs about their partner's personality and the partner's true personality scores. This difference is then standardised. This dependent variable is thus a measure of the error or inaccuracy in the player's beliefs about their partner's personality. The independent variables are the player's own personality traits, the true personality trait score of the partner, the player's eyes test score and these variables interacted with the treatment dummy. The control variables are the player's IQ, gender, age and risk aversion and these variables interacted with the treatment dummy. Columns 1 and 2 show that overestimation of partner's extraversion increases with the player's own extraversion, an effect which is significantly stronger in the treatment group compared to the control group. In column 3, the negative significant (p-value < 0.10) interaction term between the player's eyes test score and the treatment dummy shows that with increasing eyes test score, the inaccuracy in the player's beliefs about partner's extraversion is significantly lower in the treatment group compared to the control. Columns 4 and 5 show no significant effect of own extraversion or neuroticism on overestimation of the partner's neuroticism in either of the two groups. Column 6 shows that the player's performance in the eyes test has no significant impact on the inaccuracy of their beliefs about partner's neuroticism.

Table 1.A.4: Impact of beliefs about own cognitive ability on beliefs about partner's cognitive ability

	IQ Belief		Overestimation of IQ Belief		Inaccuracy of IQ Belief	
	(1)	(2)	(3)	(4)	(5)	(6)
Own IQ Belief × Treatment	-0.0588	-0.0626	-0.0445	-0.0474	-0.1807	-0.3183**
	(0.086)	(0.116)	(0.065)	(0.088)	(0.112)	(0.143)
Partner's IQ × Treatment	-0.0345	-0.0186	-0.0261	-0.0141	0.0912	0.0881
	(0.081)	(0.082)	(0.061)	(0.062)	(0.148)	(0.145)
Own IQ belief	0.6706***	0.7319***	0.5077***	0.5541***	-0.1120	0.0198
	(0.060)	(0.078)	(0.045)	(0.059)	(0.079)	(0.105)
Partner's IQ	0.0937*	0.0894*	-0.6296***	-0.6328***	-0.1668**	-0.1588**
	(0.050)	(0.050)	(0.038)	(0.038)	(0.082)	(0.077)
Treatment	-0.0833	0.4362	-0.0631	0.3303	0.0693	0.3750
	(0.082)	(0.506)	(0.062)	(0.383)	(0.108)	(0.625)
Own IQ × Treatment		-0.0172		-0.0130		0.1404
		(0.110)		(0.083)		(0.122)
Eyes Test Score × Treatment		0.0276		0.0209		0.1534
		(0.099)		(0.075)		(0.122)
Own IQ		-0.0714		-0.0541		-0.1146
		(0.069)		(0.053)		(0.087)
Eyes Test Score		0.0194		0.0147		-0.1784**
		(0.077)		(0.058)		(0.082)
Controls	No	Yes	No	Yes	No	Yes
N	338	338	338	338	338	338

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Column 1 examines the impact of the player's beliefs about own IQ, partner's true IQ and their interaction with the treatment dummy, on beliefs about the partner's IQ. While own IQ belief interacted with treatment dummy has no significant effect, own IQ belief positively impacts beliefs about partner's IQ. Column 2 includes the player's (i.e. the predictor's) true IQ as measured by the Raven's test, the player's eyes test score, along with their interactions with the treatment dummy. Columns 2 also includes the control variables - player's age, gender and risk aversion - and the 3 control variables interacted with the treatment dummy. For columns 3 and 4 the dependant variable is the standardised difference between the beliefs about partner's IQ and the partner's true IQ (as measured by the partner's performance in the Raven's test). Hence, for columns 3 and 4 the dependant variable is a measure of the degree by which the player overestimates their partner's IQ. Columns 3 and 4 indicate that an increase in player's own IQ belief leads to overestimation of the partner's IQ, irrespective of being in the treatment or control group i.e. players project beliefs about their own IQ onto their partner. For columns 5 and 6 the dependant variable is the standardised absolute difference between the beliefs about partner's IQ and the partner's true IQ. Hence, for columns 5 and 6 the dependant variable is a measure of the inaccuracy in the player's beliefs about their partner's IQ. In column 6, the significant (p-value < 0.05) negative interaction between own IQ belief and the treatment dummy, implies that as own IQ belief increases, the inaccuracy in beliefs about partner's IQ is significantly lower in the treatment group compared to the control.

Table 1.A.5: Impact of (absolute) difference between own personality and beliefs about partner's personality on level-k strategy chosen

	Level Belief			Level Chosen		
	(1)	(2)	(3)	(4)	(5)	(6)
DiffExtraversion × Treatment	-0.5302*	-0.5600*	-0.5396*	-0.6597***	-0.7395***	-0.6505**
	(0.269)	(0.290)	(0.299)	(0.237)	(0.242)	(0.254)
DiffNeuroticism × Treatment	0.1879	0.2106	0.3353	-0.0415	-0.0060	0.1645
	(0.248)	(0.263)	(0.292)	(0.248)	(0.244)	(0.262)
DiffExtraversion	0.1470	0.1806	0.1448	0.2046	0.2089	0.1663
	(0.198)	(0.201)	(0.204)	(0.177)	(0.180)	(0.183)
DiffNeuroticism	-0.1579	-0.1491	-0.2589	-0.1604	-0.1499	-0.2929
	(0.183)	(0.190)	(0.211)	(0.174)	(0.182)	(0.189)
Treatment	0.1668	0.0079	-2.9732	0.0677	-0.0799	-2.3533
	(0.267)	(0.285)	(2.048)	(0.279)	(0.286)	(1.854)
Own Extraversion × Treatment		-0.0578	0.0509		-0.1528	0.0320
		(0.306)	(0.364)		(0.289)	(0.317)
Own Neuroticism × Treatment		-0.1846	-0.1402		-0.4226	-0.4142
		(0.277)	(0.303)		(0.280)	(0.280)
Own Extraversion		-0.0612	-0.1633		-0.1773	-0.2771
		(0.196)	(0.203)		(0.214)	(0.216)
Own Neuroticism		0.0227	-0.0960		0.2069	0.0490
		(0.197)	(0.215)		(0.201)	(0.200)
Extraversion Belief × Treatment		-0.2863	-0.2942		-0.2422	-0.2628
		(0.272)	(0.284)		(0.255)	(0.263)
Neuroticism Belief × Treatment		-0.2533	-0.1483		-0.2287	-0.1007
		(0.292)	(0.303)		(0.282)	(0.297)
Extraversion Belief		0.1724	0.1403		0.1498	0.1261
		(0.194)	(0.197)		(0.183)	(0.179)
Neuroticism Belief		-0.1412	-0.1978		-0.0924	-0.1278
		(0.195)	(0.196)		(0.203)	(0.209)
Eyes Test Score × Treatment			0.5905*			0.6297**
			(0.301)			(0.311)
Own IQ × Treatment			-0.2809			-0.3072
			(0.296)			(0.307)
IQ Belief × Treatment			0.3462			0.2073
			(0.324)			(0.270)
Female × Treatment			-0.8057			-0.8928
			(0.612)			(0.564)
Order × Treatment			1.1902**			1.1083*
			(0.594)			(0.605)
Eyes Test Score			-0.4462*			-0.4522*
			(0.249)			(0.254)
Own IQ			0.2175			0.2606
			(0.203)			(0.220)
IQ Belief			-0.3538*			-0.3321*
			(0.210)			(0.200)
Female			1.1196**			1.4361***
			(0.435)			(0.388)
Order			-0.7462*			-0.9827**
			(0.401)			(0.420)
Controls	No	No	Yes	No	No	Yes
N	338	338	338	338	338	338

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Here, 'Controls' imply the player's risk aversion, age and their interactions with the treatment dummy.

Table 1.A.6: Impact of (absolute) difference between own personality and beliefs about partner's personality on the probability of choosing the best response - Probit Model

	Control		Treatment	
	(1) Pr(Level=2)	(2) Pr(Level=2)	(3) Pr(Level=2)	(4) Pr(Level=2)
DiffExtraversion	-0.0453	-0.0550	0.0846***	0.0992***
	(0.038)	(0.035)	(0.030)	(0.032)
DiffNeuroticism	-0.0008	-0.0077	-0.0459	-0.0358
	(0.031)	(0.031)	(0.032)	(0.033)
Own Extraversion		0.0123		0.0168
		(0.030)		(0.046)
Own Neuroticism		0.0543*		-0.0438
		(0.032)		(0.036)
Extraversion Belief		-0.0165		-0.0109
		(0.035)		(0.031)
Neuroticism Belief		0.0296		0.0656*
		(0.033)		(0.035)
Own IQ		0.0587*		0.0558
		(0.036)		(0.038)
IQ Belief		-0.0441		-0.0028
		(0.029)		(0.036)
Eyes Test Score		0.0549		0.0497
		(0.037)		(0.031)
Controls	No	Yes	No	Yes
N	170	170	168	168

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the average marginal effects from Probit regressions. 'Controls' imply the player's age, gender, risk aversion and the order of play of the two games.

Table 1.A.7: Impact of (absolute) difference between own personality and beliefs about partner's personality on the probability of choosing the best response - Logit Model

	Control		Treatment	
	(1) Pr(Level=2)	(2) Pr(Level=2)	(3) Pr(Level=2)	(4) Pr(Level=2)
DiffExtraversion	-0.0486	-0.0547	0.0843***	0.1016***
	(0.041)	(0.040)	(0.029)	(0.030)
DiffNeuroticism	-0.0019	-0.0074	-0.0459	-0.0370
	(0.030)	(0.031)	(0.032)	(0.032)
Own Extraversion		0.0093		0.0185
		(0.030)		(0.047)
Own Neuroticism		0.0548		-0.0422
		(0.034)		(0.037)
Extraversion Belief		-0.0154		-0.0102
		(0.039)		(0.032)
Neuroticism Belief		0.0321		0.0665*
		(0.035)		(0.035)
Own IQ		0.0618		0.0583
		(0.038)		(0.040)
IQ Belief		-0.0428		-0.0042
		(0.027)		(0.037)
Eyes Test Score		0.0531		0.0454
		(0.038)		(0.033)
Controls	No	Yes	No	Yes
N	170	170	168	168

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The table reports the average marginal effects from Logit regressions. 'Controls' imply the player's age, gender, risk aversion and the order of play of the two games.

Table 1.A.8: Impact of beliefs about partner's personality on beliefs about partner's contribution and own contribution in the public goods game

	Control Order 1				Treatment Order 1			
	(1) Contribution Belief	(2) Contribution Belief	(3) Own Contribution	(4) Own Contribution	(5) Contribution Belief	(6) Contribution Belief	(7) Own Contribution	(8) Own Contribution
Extraversion Belief	0.0430	0.0575	0.0951	0.1042	0.1964*	0.1879*	0.1882**	0.1667*
	(0.083)	(0.082)	(0.087)	(0.101)	(0.101)	(0.100)	(0.087)	(0.083)
Neurotcisim Belief	0.0440	0.0456	-0.0207	-0.0275	0.1771	0.1627	0.1591	0.1697
	(0.090)	(0.109)	(0.087)	(0.101)	(0.111)	(0.109)	(0.117)	(0.112)
Own IQ		-0.0664		-0.0114		0.1265		0.1782*
		(0.106)		(0.087)		(0.088)		(0.101)
IQ Belief		0.1329		0.1016		0.0964		0.2512**
		(0.097)		(0.107)		(0.096)		(0.097)
Eyes Test Score		-0.0256		0.0221		0.1197		0.1694
		(0.096)		(0.130)		(0.090)		(0.117)
Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	110	110	110	110	106	106	106	106

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table shows that of the two fundamental personality traits - extraversion and neuroticism - only beliefs about partner's extraversion affect decision making in the public goods game, for treatment group subjects. 'Controls' refers to the player's age, gender and risk aversion.

Table 1.A.9: Impact of beliefs about partner's personality and own personality on beliefs about partner's contribution and own contribution in Public Goods Game - OLS approach

	Control OLS		Treatment OLS	
	(1) Contribution Belief	(2) Own Contribution	(3) Contribution Belief	(4) Own Contribution
ExtraversionBelief	0.0601	0.1110	0.2036**	0.1599*
	(0.082)	(0.092)	(0.099)	(0.085)
OwnExtraversion	-0.0733	-0.2041**	-0.1831	-0.0919
	(0.095)	(0.088)	(0.118)	(0.117)
Own IQ	-0.0583	-0.0417	0.0783	0.1484
	(0.096)	(0.084)	(0.086)	(0.099)
IQ Belief	0.1250	0.1140	0.0953	0.2474**
	(0.091)	(0.100)	(0.096)	(0.099)
Eyes Test Score	-0.0431	-0.0015	0.1328	0.1754
	(0.096)	(0.118)	(0.099)	(0.127)
Controls	Yes	Yes	Yes	Yes
N	110	110	106	106

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Controls' refers to the player's age, gender and risk aversion.

Table 1.A.10: Impact of beliefs about partner's personality and own personality facets on beliefs about partner's contribution and own contribution in Public Goods Game

	Control OLS		Treatment IV	
	(1) Contribution Belief	(2) Own Contribution	(3) Contribution Belief	(4) Own Contribution
ExtraversionBelief	0.0542	0.1036	0.6169**	0.5262**
	(0.084)	(0.093)	(0.265)	(0.251)
OwnAssertiveness	-0.1258	-0.2271*	-0.3287**	-0.3095**
	(0.113)	(0.114)	(0.128)	(0.124)
OwnActivity	0.0593	0.0333	0.0255	0.1562
	(0.122)	(0.123)	(0.125)	(0.106)
Own IQ	-0.0497	-0.0323	0.0781	0.1396
	(0.099)	(0.088)	(0.098)	(0.105)
IQ Belief	0.1391	0.1301	0.1041	0.2708***
	(0.089)	(0.102)	(0.091)	(0.092)
Eyes Test Score	-0.0342	0.0114	0.1193	0.1751
	(0.102)	(0.122)	(0.118)	(0.139)
Controls	Yes	Yes	Yes	Yes
N	110	110	106	106

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

'Controls' refers to the player's age, gender and risk aversion. Columns 1 and 2 report the OLS regression results for the control group. Column 2 shows that the players own assertiveness has a negative significant effect (p-value < 0.05) on contribution levels whereas facet activity has an insignificant positive effect. None of the facets significantly impact beliefs about partner's contribution. Columns 3 and 4 present the results from 2SLS IV regression for the treatment group. For the treated subjects, beliefs about partner's extraversion positively and significantly (p-value < 0.05) affects beliefs about partner's contribution as well as own-contribution. With regards to the player's own personality, facet assertiveness has a significant negative effect (p-value < 0.05) on both contribution belief and own-contribution, whereas facet activity has an insignificant positive effect.

Table 1.A.11: Impact of beliefs about partner's personality and own personality on beliefs about partner's contribution and own contribution in Public Goods Game - Order 2

	Control OLS		Treatment IV	
	(1) Contribution Belief	(2) Own Contribution	(3) Contribution Belief	(4) Own Contribution
Extraversion Belief	-0.0357	-0.2345*	0.1273	1.2682
	(0.147)	(0.121)	(1.065)	(1.986)
Own Extraversion	0.1603	0.0317	0.1219	-0.1167
	(0.158)	(0.158)	(0.189)	(0.321)
Own IQ	0.1372	0.0435	-0.0345	-0.0495
	(0.203)	(0.162)	(0.120)	(0.223)
IQ Belief	0.1792	0.0170	-0.0657	-0.1679
	(0.159)	(0.133)	(0.142)	(0.209)
Eyes Test Score	-0.2673	0.2327	0.2574	0.0801
	(0.174)	(0.164)	(0.157)	(0.330)
Controls	Yes	Yes	Yes	Yes
N	60	60	62	62

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

This table replicates the IV regression results from the main paper but only for those subjects which played the 11-20 game first. 'Controls' refers to the player's age, gender and risk aversion.

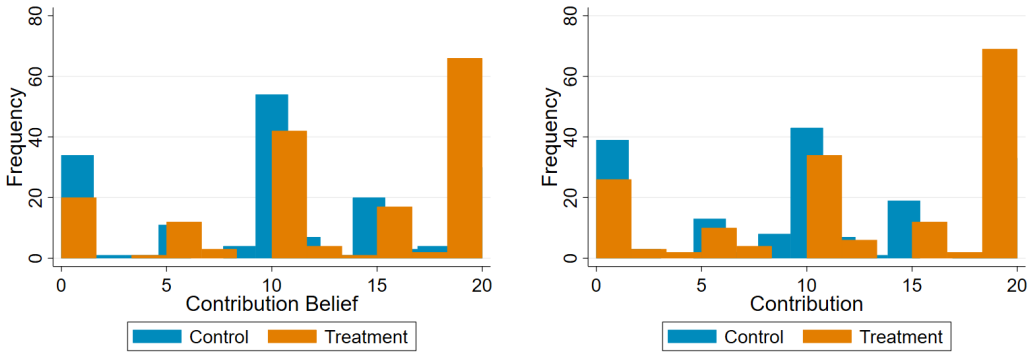
Table 1.A.12: Relationship between number of words and emotional content of text spoken by the subject and the subject's own personality

	Own Extraversion			Own Neuroticism		
	(1)	(2)	(3)	(4)	(5)	(6)
Own Number of Words	0.1439**	0.1733**	0.1781**	0.1289*	0.1282*	0.0814
	(0.071)	(0.073)	(0.075)	(0.071)	(0.071)	(0.073)
Own Valence		-0.5325*	-0.5533**		0.2299	0.3302
		(0.284)	(0.276)		(0.269)	(0.255)
Own Arousal		0.0296	0.0358		-0.2117	-0.2409*
		(0.150)	(0.152)		(0.135)	(0.132)
Own Dominance		0.4011	0.4033		-0.0124	-0.0905
		(0.279)	(0.272)		(0.258)	(0.252)
Own IQ			-0.1804**			-0.0031
			(0.085)			(0.076)
Eyes Test Score			-0.0069			0.1478*
			(0.090)			(0.084)
Age			0.0210			0.0021
			(0.027)			(0.023)
Female			0.0201			0.3858**
			(0.169)			(0.162)
Non-Native Speaker			0.0798			-0.1319
			(0.156)			(0.174)
First Speaker			0.1562			0.1071
			(0.156)			(0.162)
N	168	168	168	168	168	168

Standard errors in parentheses

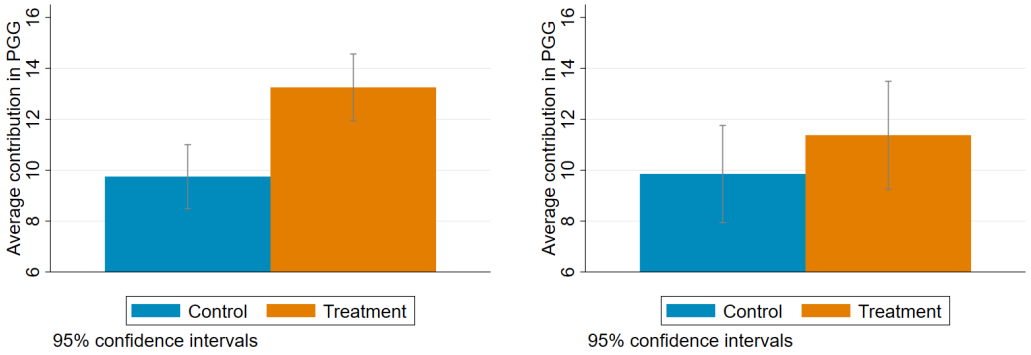
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The explanatory variables, namely the number of words spoken by the subject, and valence, arousal and dominance ratings of language used, are standardised for comparability across coefficients. Number of words used by the subject is a positive indicator of the subject's extraversion, even after adding valence, arousal and dominance ratings of the text used as explanatory variables. The result persists after controlling for the subject's IQ, eyes test score, age, gender, a dummy for non-native speaker (equals 1 if the subject is a non-native English speaker and 0 otherwise) and a dummy for first speaker (equals 1 if the subject started the conversation and 0 otherwise). Trait neuroticism also appears to be positively associated with number of words used, although the coefficient becomes insignificant after adding sensible control variables.



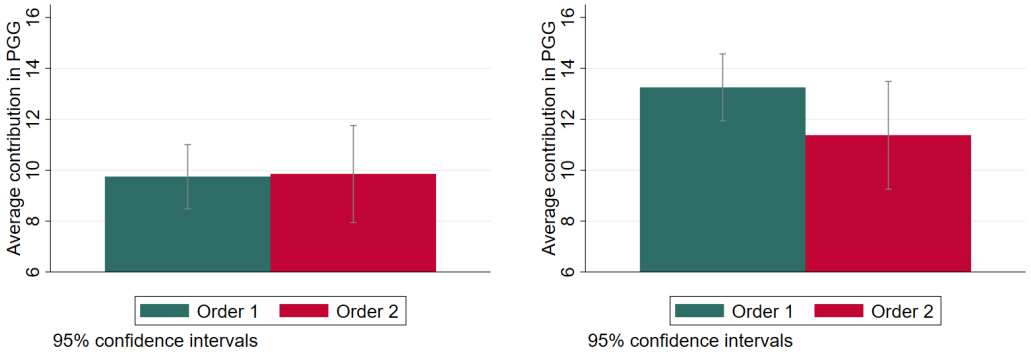
(a) Contribution Belief (b) Own Contribution

Figure 1.A.1: Distribution of (a) Beliefs about Partner's Contribution and (b) Own Contribution in the Public Goods Game



(a) Order 1 (b) Order 2

Figure 1.A.2: Average contribution in Public Goods Game (PGG) (a) when PGG is played first (order 1) and (b) when the 11-20 game is played first (order 2). Treated subjects contribute more than control group subjects in order 1. The average contribution of treated subjects is 13.2 EP where as that of control group subjects is 9.7 EP in order 1. The difference is statistically significant with t-statistic of -3.8060 and p-value < 0.01. There is no significant difference in contribution levels between the treatment and control groups in order 2.



(a) Control (b) Treatment

Figure 1.A.3: Average contribution in Public Goods Game (PGG) for different orders of play of the two games for (a) Control and (b) Treatment groups. Order 1 is when PGG is played first and order 2 is when the 11-20 game is played first. On average players contribute more in the treatment group (figure (b)) when PGG is played first (order 1) compared to when 11-20 is played first (order 2). In a one-tailed t-test, we reject the null of no significant difference in contribution between treated players in order 1 and treated players in order 2 in favour of the alternative that treated players in order 1 contribute more at the 10% significance level (t-statistic = 1.5752, p-value =0.0586). There is no significant difference for control group subjects (figure (a)).

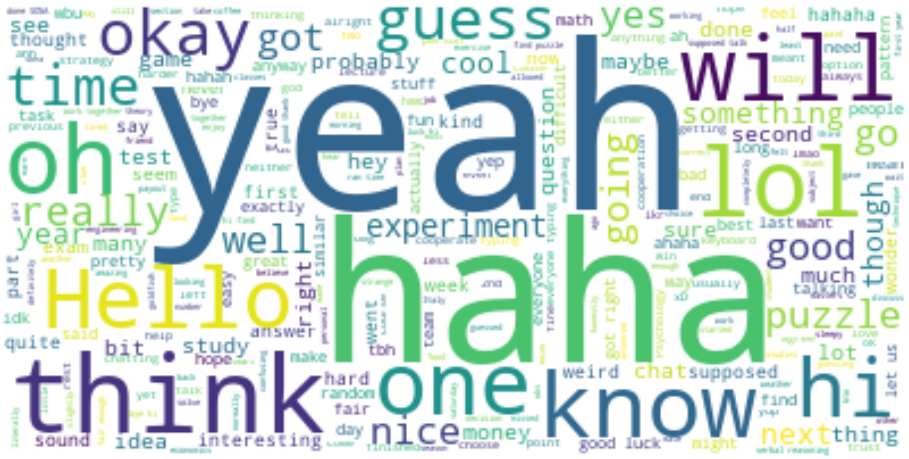


Figure 1.A.4: Most frequently used words by subjects during small talk communication

(a) Highly Extraverted

(b) Less Extraverted

(c) Highly Neurotic

(d) Less Neurotic

Figure 1.A.5: Most frequently used words during small talk communication by subjects who are believed to be (a) highly extraverted (b) less extraverted (c) highly neurotic and (d) less neurotic.

1.B Examples of Small Talk Communication

Example 1

Player 1: *hey*
Player 2: *Hey how are you doing :)*
Player 1: *lol alright*
Player 1: *you*
Player 2: *yeah fine haha*
Player 1: *tbh this is strange*
Player 2: *this is strange*
Player 2: *exactly haha*
Player 1: *omg*
Player 1: *so...*
Player 1: *do you have any pets?*
Player 2: *probably they want to see if we will cooperate depending on our chat or something haha*
Player 2: *nope and you?*
Player 1: *trying to make conversation :D*
Player 1: *yep, two cats*
Player 2: *I had fish when I was little haha*
Player 2: *What are their names?*
Player 1: *aww like goldfish?*
Player 1: *Cosmos and Titan*
Player 2: *Yes a goldfish and one more but I forgot the type lol*
Player 2: *That is great!*
Player 1: *i used to have goldfish*
Player 1: *but we could not keep them cause of the cats*
Player 2: *Goldfish live a long I think generally haha*
Player 2: *Oh no!*
Player 1: *we had 4 goldfish*
Player 2: *Cats is more interesting haha*
Player 2: *are**
Player 1: *yeah i know*
Player 1: *only problem is they scratch you*
Player 1: *a lot*
Player 2: *Ahaha yes*
Player 2: *scars all the time*
Player 1: *so now i have lots of marks on me*
Player 2: *This keyboard is so bad*

Player 2: *Oh no*
Player 2: *The pain of being a cat owner haha*
Player 1: *the keyboard never crossed my mind lol*
Player 2: *I barely can type on it haha*
Player 2: *It was nice chatting to you haha*
Player 1: *aww goodbye*

Example 2

Player 1: *hi*
Player 2: *hey*
Player 1: *what is up?*
Player 2: *not much, you?*
Player 1: *same, just waiting haha*
Player 2: *same, it is a bit dead is not it*
Player 1: *it really is...*
Player 2: *think I mucked up most of those puzzles tbh*
Player 1: *although everyone is now typing fervently*
Player 1: *you think you did that bad?*
Player 2: *not that bad, but some of them I just did not get*
Player 2: *or I almost got them and then the time ran out*
Player 1: *there were some really weird ones though*
Player 2: *yeah igy*
Player 1: *yeah same, 30 seconds is a bit too quick for some of those*
Player 2: *some just made no sense to me*
Player 1: *true that*
Player 1: *but they take 2/30 anyway,*
Player 2: *seems like a bit of a waste of time*
Player 2: *to do 30 and then only 2 count*
Player 1: *and for some reason \ q random \ q selection always ends up in me being paid nothing xD*
Player 2: *same haha*
Player 1: *Ikr*
Player 2: *or i am in a team and the team does really badly and i get almost no money*
Player 1: *but yeah, pretty much a waste*
Player 2: *really**
Player 1: *omg yes....*
Player 2: *its a bit annoying*
Player 1: *These dictator games where in the end one person decides whether I*

can keep my money or get nothing
Player 2: *yes! so irritating*
Player 1: *Being paid £3 after 1,5 hours....*
Player 2: *what a drag*

Example 3

Player 1: *Hi*
Player 2: *Hello*
Player 1: *how are you?*
Player 2: *How are you?*
Player 2: *haha*
Player 1: *haha i'm good you?*
Player 2: *great*
Player 2: *How are exams going?*
Player 1: *yeah not too bad, some have gone worse that i had wanted, you?*
Player 2: *Most of them were alright, three more to go*
Player 2: *How about you?*
Player 2: *Any more left?*
Player 1: *i've got 1 more to go, thank god, i have 7 overall*
Player 1: *how many do you have overall?*
Player 2: *That's a lot. When is your last one?*
Player 2: *I have 6 in total*
Player 1: *next wednesday*
Player 1: *so i can go to circle and pop and celebrate by getting black out drunk haha*
Player 2: *Still some time to prepare. I have one this Saturday*
Player 2: *Yeah, pop is back on again next week*
Player 1: *that's grim, my boyfriend does to, i don't get why exams on saturday is a thing*
Player 1: **too*
Player 2: *None of your 7 exams were on Saturday?*
Player 1: *nope, i had 1 in week 3, 1 week 4, 3 last week, 1 this week and one next week*
Player 2: T*ime is running out heh*

1.C Experiment Script

This following part is read out by the experimenter (Note that the script presented includes certain comments in italics which were not shown to subjects during the

experiment).

Thank you everyone for coming to our experiment today. Before we begin, please check that the number on the card handed to you matches with the number on the cubicle that you are seated in.

During the whole experiment, please do not speak with each other. If you do not understand something, please ask the experimenter by raising your hand. We will come to you and answer your question individually. Please also refrain from using your mobile phones during the experiment.

Also bear in mind that you may have to wait a few moments during the experiment, as we want everyone to finish at the same time. You will see the message 'Please wait until the experiment continues' on your screen when this is applicable.

Before we begin, I would just like to say, that your participation is very crucial for our research and we truly appreciate all of you being here. Thank you.

We will now begin the experiment.

General Instructions

In the laboratory experiment you are taking part in, you can - depending on your decisions and the decisions of your fellow players - earn money in addition to the show-up fee of £4. It is, therefore, of importance that you read these instructions carefully.

Today's experiment consists of the following: In the first section, you will be asked to answer a few questions and solve some puzzles. In the second section, you will be asked to make decisions in a few tasks. Lastly, there will be some questions for you to answer.

Please note that the experiment will not involve any deception and your answers today will remain strictly anonymous. The generated anonymous data will only be used for the purpose of our study. Therefore, we request you to answer to the best of your ability as it is integral to our research.

The outcomes from each task will be disclosed at the end of the experiment.

Detailed instructions for each part will follow.

We will now begin the experiment.

(a) Questionnaire: Personality (44 questions)

You will be asked to answer some questions about yourself. Your payment will not be affected by this. Just to remind you, your answers will remain anonymous so please answer as truthfully as possible as this is critically important for our research.

You will see a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please pick an option next to each statement to indicate the extent to which you agree or disagree with that statement.

I see myself as someone who...

START BFI QUESTIONNAIRE

(b) PUZZLES: Raven's Test (30 items)

You will be asked to solve some puzzles, a pattern game.

On the screen, you will see a set of abstract pictures with one of the pictures missing. You need to choose a picture from the choices given below to complete the pattern. You will have 30 seconds to complete each set of pictures. The first picture you will see will be an example, no input is required. You will then be asked to solve a total of 30 such puzzles. 2 of these 30 puzzles will randomly be selected. For each correct answer, from the random 2, you will receive £1. Please make sure to click 'submit answer', as otherwise your answer will not be recorded, and you might lose money.

START RAVEN TEST

Out of the 30 puzzles you just saw, how many puzzles do you think you correctly solved?

If your answer to this question is correct, then you will win an additional £1.

Now subjects will be allocated to one of 2 groups - control or treatment.

Control Group

Placebo Task[20]

Can you please indicate the title and summarize the story of the last movie you have seen? Please be as specific as possible and include as many details as possible. Please use a minimum of 250 characters. You will have 4 minutes to write the summary.

[20]This task has been adapted from the Placebo Task used in Bursztyn et al., 2017.

APPENDIX

Please write the summary in the box provided on the next screen.

(next screen) Please make sure to click 'Submit' after you are done, as otherwise your answer will not be recorded.

Beliefs

You have been randomly and anonymously matched with another person in this room who is participating in the experiment. Please answer a few questions about the other player to the best of your ability, before you proceed with the tasks.

1. You will see a number of characteristics that may or may not apply to the other player. For example, do you agree that the other player is someone who likes to spend time with others? Please pick an option next to each statement to indicate the extent to which you agree or disagree with the statement regarding the other player.

   You will see 11 statements about the other player.

   1 out of these 11 statements will be randomly chosen and if your answer matches that of the other player, then you will win an additional £1.

   START PERSONALITY PREDICTION QUESTIONNAIRE

2. Recall the visual puzzle task from earlier in the experiment. On the screen, you saw a set of abstract pictures with one of the pictures missing. You had to choose a picture from the choices given below to complete the pattern. You had 30 seconds to complete each set of pictures. You were asked to solve a total of 30 such puzzles. How many puzzles do you think the other player, with whom you have been matched, correctly solved? Please indicate a (whole) number between 0 and 30.

   If your answer to this question is correct, then you will win an additional of £1.

   Tasks (note that the order of the two tasks below were randomised)

   You will now take part in a few decision-making tasks with the player with whom you have already been matched. Note that you will be participating in all tasks with the same player. Your payoff from these tasks will be calculated in Experimental Pounds (EP). The exchange rate between £ and EP is 1:5, i.e. 5 EP = £1.

   The outcomes from each task will be disclosed at the end of the experiment. You will receive payment based on your results from one of the tasks randomly

selected from the tasks in this part of the experiment. Please note that each task is equally likely to be chosen for payment.

Task 1: PGG

You will now participate in a task with the player with whom you have been matched. You have 20 EP and the other player has 20 EP as well. Your task in the game, and also the other player's task, is to decide how much to contribute to a joint project. You can choose to contribute any amount between 0 and 20 EP (only integer numbers). Your earnings from the project is the total contribution to the project, made by you and the other player, multiplied by a factor of 3/4. Your payoff from this task will be your earnings from the project, plus the amount you did not contribute. Thus, your final payoffs (in EP) will be given by:

Your payoff = (20 – your contribution)+3/4(your contribution + the other player's contribution)

Other player's payoff = (20 – the other player's contribution)+3/4(your contribution + the other player's contribution)

If for example, you contribute 20 EP to the project and the other player contributes 20 EP then,

Your payoff will be: $20–20 + 3/4(20 + 20) = 30$

The other player's payoff will be: $20–20 + 3/4(20 + 20) = 30$

If for example, you contribute 0 EP to the project and the other player contributes 20 EP then,

Your payoff will be: $20–0 + 3/4(0 + 20) = 35$

The other player's payoff will be: $20–20 + 3/4(0 + 20) = 15$

If you have a question, please raise your hand.

If you have read the instructions and do not have any questions, please click 'OK' to proceed to a practice quiz. The quiz is to make sure that you understand the task and your answers will not affect your payoffs from the experiment.

Suppose you choose to contribute 20 EP and the other player chooses to contribute 0 EP.

Your payoff will be:

The other player's payoff will be:

Suppose you choose to contribute 10 EP and the other player chooses to contribute 14 EP.

Your payoff will be:

The other player's payoff will be:

You have correctly answered the practice quiz. Click 'Continue' to proceed with the task.

How much money do you think the other player will contribute? Please indicate a number (an integer) between 0 and 20.

If your answer to this question matches that of the other player, then you will win an additional £1.

How much would you like to contribute? Please choose a number (an integer) between 0 and 20.

Task 2: 11-20 money request game

You will now participate in a different task with the same player.

You and the other player are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 Experimental Pounds. Each player will receive the amount he or she requests. A player will receive an additional amount of 20 Experimental Pounds if he or she asks for exactly one Experimental Pound less than the other player.

If for example, you request 19 EP and the other player requests 20 EP then,

Your payoff will be: $19 + 20 = 39$

The other player's payoff will be: 20

If for example, you request 17 EP and the other player requests 16 EP then,

Your payoff will be: 17

The other player's payoff will be: $16 + 20 = 36$

If you have a question, please raise your hand.

If you have read the instructions and do not have any questions, please click 'OK' to proceed to a practice quiz. The quiz is to make sure that you understand the task and your answers will not affect your payoffs from the experiment.

Suppose you choose to request 13 EP and the other player chooses to request 14 EP.

Your payoff will be:

The other player's payoff will be:

Suppose you choose to request 15 EP and the other player chooses to request 18 EP.

Your payoff will be:

The other player's payoff will be:

You have correctly answered the practice quiz. Click 'Continue' to proceed with the task.

How much money do you think the other player will request? Please indicate a number (an integer) between 11 and 20.

If your answer to this question matches that of the other player, then you will win an additional £1.

What amount of money would you request? Please choose a number (an integer) between 11 and 20.

Treatment Group

Chat Instructions

You have been randomly and anonymously matched with another person in this room who is participating in the experiment.

Before you proceed with the tasks, you are allowed to chat with the other player for 4 minutes. You can type in the box provided at the bottom of the screen and press Enter on your keyboard to send your messages.

Your message should not contain any personal information such as your name or your computer ID. The purpose is to preserve anonymity throughout the experiment. You are allowed to chat freely in English and in a non-abusive manner.

Beliefs

Now that you have chatted with the other player please answer a few questions about the other player, before you proceed with the tasks.

APPENDIX

1. You will see a number of characteristics that may or may not apply to the other player. For example, do you agree that the other player is someone who likes to spend time with others? Please pick an option next to each statement to indicate the extent to which you agree or disagree with the statement regarding the other player.

   You will see 11 statements about the other player.

   1 out of these 11 statements will be randomly chosen and if your answer matches that of the other player, then you will win an additional £1.

   START PERSONALITY PREDICTION QUESTIONNAIRE

2. Recall the visual puzzle task from earlier in the experiment. On the screen, you saw a set of abstract pictures with one of the pictures missing. You had to choose a picture from the choices given below to complete the pattern. You had 30 seconds to complete each set of pictures. You were asked to solve a total of 30 such puzzles. How many puzzles do you think the other player, with whom you chatted, correctly solved? Please indicate a (whole) number between 0 and 30.

   If your answer to this question is correct, then you will win an additional £1.

   Tasks (note that the order of the two tasks below were randomised)

   You will now take part in a few decision-making tasks with the player you chatted with. Note that you will be participating in all tasks with the same player. Your payoff from these tasks will be calculated in Experimental Pounds (EP). The exchange rate between £ and EP is 1:5, i.e. 5 EP = £1.

   The outcomes from each task will be disclosed at the end of the experiment. You will receive payment based on your results from one of the tasks randomly selected from the tasks in this part of the experiment. Please note that each task is equally likely to be chosen for payment.

   Task 1: PGG

   You will now participate in a task with the player you chatted with. You have 20 EP and the other player has 20 EP as well. Your task in the game, and also the other player's task, is to decide how much to contribute to a joint project. You can choose to contribute any amount between 0 and 20 EP (only integer numbers). Your earnings from the project is the total contribution to the project, made by you and the other player, multiplied by a factor of ¾. Your payoff from this task will be your earnings from the project, plus the amount you did not contribute. Thus, your final payoffs (in EP) will be given by:

Your payoff = (20– your contribution)+$^3/_4$(your contribution + the other player's contribution)

Other player's payoff = (20– the other player's contribution)+$^3/_4$(your contribution + the other player's contribution)

Examples and quiz related to the game, then partner's strategy belief and own choice

Task 2: 11-20 money request game

You will now participate in a different task with the same player.

You and the other player are playing a game in which each player requests an amount of money. The amount must be (an integer) between 11 and 20 Experimental Pounds. Each player will receive the amount he or she requests. A player will receive an additional amount of 20 Experimental Pounds if he or she asks for exactly one Experimental Pound less than the other player.

Examples and quiz related to the game, then partner's strategy belief and own choice

FOR BOTH CONTROL AND TREATMENT:

Eyes Test (36 questions)

In this section, you will be asked to look at 36 pictures of different pairs of eyes.

For each set of eyes, choose the word which best describes what the person in the picture is thinking or feeling. You may feel that more than one word is applicable but please choose just one word, the word which you consider to be most suitable. Before making your choice, make sure that you have read all 4 words. You should try to do the task as quickly as possible, but you will not be timed. If you do not know what a word means you can read the meaning of the word provided at the bottom of the screen.

2 of these 36 questions you answer will randomly be selected. For each correct answer, from the random 2, you will receive £1.

You will first see a practice question with four options. The correct option will be highlighted. After that you may proceed to the questions.

Which word best describes what the person in the picture is thinking or feeling?

START EYES TEST

APPENDIX

Questionnaire

Thank you. Now, in the final section, you will be asked to answer some questions about yourself.

(a) Risk

Please indicate the likelihood that you would engage in the described activity or behaviour if you were to find yourself in that situation.

START DOSPERT

(b) Personal information

1. How old are you? (in years)
2. What is your year of study? (1, 2, 3, Post-graduate Other)
3. What is your gender? (M, F, Other, Prefer not to say)
4. What is your nationality?
5. Is English your Native language? (Yes, No)
6. What is your current degree course?
7. Would you consider your degree course mostly: (quantitative, qualitative)
8. Have you ever taken any game theory modules/courses? (Yes, No)
9. How dissatisfied or satisfied are you with your life in general? (1-7 scale from completely dissatisfied to completely satisfied)

Profit display screen

1. Number of correct answers from the visual puzzles task (out of 30):
2. Your payoff (in EP) from the first decision-making task:
3. Your payoff (in EP) from the second decision-making task:
4. Number of correct answers from the eyes task (out of 36):
5. Additional amount earned (in £):
6. Total earnings (in £):

Thank you for completing the experiment successfully. Please queue at the marked line once you are done, show the number card and collect your payment in cash.

1.D Personality Beliefs Questionnaire

The personality beliefs questionnaire used in the study was adapted from Rammstedt and John, 2007 and is presented below:
Please pick an option next to each statement to indicate the extent to which you agree or disagree with the statement **regarding the other player**.
For each of the below statements the subject could pick any one of five options - Disagree strongly, Disagree a little, Neither agree nor disagree, Agree a little and Agree strongly

1. The other player is reserved.
2. The other player is generally trusting.
3. The other player tends to be lazy.
4. The other player is relaxed, handles stress well.
5. The other player has few artistic interests.
6. The other player is outgoing, sociable.
7. The other player tends to find fault with others.
8. The other player does a thorough job.
9. The other player gets nervous easily.
10. The other player has an active imagination.
11. The other player is considerate and kind to almost everyone.

2 Climate Change and Diet[1]

with Thomas Hills and Daniel Sgroi

Despite extensive scientific evidence, the public remains largely uninformed about the detrimental environmental effects of meat consumption. We conducted a pre-registered online randomised control trial, involving 1220 subjects, of which only 9% reported diet as impacting climate change. Evaluating 6 information interventions to reduce planned meat consumption, the most effective interventions were based on scientific knowledge, and efficacy salience i.e. providing concrete information about the consequences of one's actions. A targeted messaging approach that highlights the health impact of a plant-based diet for individuals with health concerns was also effective. The greatest resistance to the interventions was associated with motivated reasoning around meat consumption: frequent meat eaters reported lower prior knowledge, lower responses to the evidence, and more moral offence at being informed. Further insight into the information interventions is provided by evaluating donations to a climate change charity and the recall value of information provided.

2.1 Introduction

Plant-based diets are of utmost importance in the path to mitigating climate change (Schiermeier, 2019). A global shift towards a diet with lower meat content (based on dietary recommendations by the Harvard Medical School for Public Health) is estimated to reduce the costs of climate change mitigation by up to 50% in 2050, compared to a reference case where there are no changes to diet or climate policy (Stehfest et al., 2009). Despite scientific consensus regarding the benefits of a plant-based diet, meat intake related climate change mitigation actions rank low among the public in perceived effectiveness compared to actions such as energy saving, recycling and reduced levels of driving (Vanhonacker et al., 2013; de Boer et al., 2016; Truelove and Parks, 2012).

Interventions aiming to enhance public awareness about the ecological impact of a

[1]Ethical approval for the study was granted by the University of Warwick (HSSREC 03/19-20). The experiment is registered at the AEA RCT registry https://doi.org/10.1257/rct.5069-1.0 (RCT ID AEARCTR-0005069) and https://doi.org/10.1257/rct.5584-1.0 (RCT ID AEARCTR-0005584). This work was supported by the Economic and Social Research Council (ESRC) CAGE Centre under grant RES-626-28-0001. The authors are grateful to seminar participants at the Warwick Business School and Department of Economics, Warwick University, for their helpful comments. The authors declare no conflicts of interest.

dietary shift and increase intention to reduce meat content in individual diets have been scattered and inconclusive. Harguess et al., 2020 offer a systematic review of experimental studies aiming to reduce meat consumption. These studies largely focus on providing information about the health effects of meat consumption (Bertolotti et al., 2016; Fehrenbach, 2015) or the environmental impact of meat production (Graham and Abrahamse, 2017) or both (Vainio et al., 2018; Verain et al., 2017). A branch of literature has also used animal welfare arguments, in order to invoke empathy, to persuade individuals to reduce their meat intake (Kunst and Palacios Haugestad, 2018; Tian et al., 2016). However, what remains absent in the literature is a comprehensive evaluation of different information interventions aiming to reduce meat consumption.

In this paper we evaluate on a level playing-field a range of information interventions in relation to individual differences among those receiving the interventions. We outline these briefly below as well as discussing policy ramifications.

One simple barrier to the adoption of a plant-based diet is low public awareness about the impact of dietary choices on climate change (Sanchez-Sabate and Sabaté, 2019; de Boer et al., 2016; Macdiarmid et al., 2016; Vanhonacker et al., 2013). The *deficit model of science communication* (Bauer et al., 2007), a widely used framework, promotes the enhancement of public knowledge about complex issues like climate change through scientific facts. With regards to meat consumption, there is a dearth of knowledge regarding not only the environmental impacts of meat but also the general health ramifications of a meat-rich diet (Lacroix and Gifford, 2019). Emphasizing the personal health impact of eating meat has an added benefit of making communication interventions more personally relevant to those who may feel climate change is an otherwise distant threat (Monroe et al., 2019; Ockwell et al., 2009). However, the benefits of the deficit model can be moderated by barriers such as scepticism, perceived inefficacy and social norms. Furthermore, acceptance of scientific facts about climate change is dependent on perceived scientific (Lewandowsky, Gignac, et al., 2013) and social consensus (Lewandowsky, Cook, et al., 2019). Additionally, it has been observed that individuals have a tendency to reject scientific information about climate change when it contradicts pre-existing beliefs (Druckman and McGrath, 2019; Hart and Nisbet, 2012). Such *directional motivated reasoning* (Kunda, 1990) or *strategic ignorance* (Carrillo and Mariotti, 2000) could further blunt the effectiveness of the deficit model.

The extended Parallel Process Model (EPPM) (Witte, 1992) posits that persuasive messages highlighting potential threats, such as from Climate Change, should

also include efficacy information to promote attitude and behaviour change. Efficacy describes an individual's perception that an issue is addressable and that they are capable of taking necessary action to tackle the issue (Balch, 1974). Efficacy includes both *efficacy expectation* and *outcome expectancy*. While efficacy expectation (or self efficacy) is the belief that one is capable of performing a certain behaviour, outcome expectancy (or response efficacy) is an individual's estimate that the behaviour will lead to specific outcomes (Bandura, 1977). Prior literature has found that messages with high self efficacy (Xue et al., 2016) and response efficacy (Hart and Feldman, 2016b) content can have a positive impact on encouraging action to combat climate change. In this study, we focus on response efficacy information. We combine response efficacy with (1) salience or vividness (Chang and Lee, 2010; Gonzales et al., 1988) (what we call *efficacy salience*) and (2) capacity for an individual to affect others (Wolske et al., 2020; Weeks et al., 2015) (what we call *social efficacy*).

Many meat-eaters enjoy eating meat but dislike the notion that their diet involves killing or harming of animals. This state of inconsistent beliefs is termed the "meat paradox" (Loughnan et al., 2010). Meat-eaters can resolve this form of *cognitive dissonance*, where beliefs and behaviour are inconsistent (Festinger, 1957), by not associating meat-eating with living and sentient animals or by withdrawing moral status from animals (Loughnan et al., 2010). Inducing higher cognitive dissonance through statements and images which connect meat with living animals has been shown to have a positive impact in reducing willingness to eat meat (Tian et al., 2016; Kunst and Hohle, 2016) and provides another possible intervention. However, this may be reduced by *hedonic motivations* like pleasure derived from meat consumption (Feinberg et al., 2019).

Doherty and Webler, 2016 posit that descriptive social norms (information about others' behaviour) are an important predictor of public engagement in climate actions. In particular, eating meat is a well-reinforced norm as meat is often served in social gatherings and restaurants (Sparkman and Walton, 2017). Information about social norms have shown a positive effect on energy conservation (Allcott, 2011; Hafner, Elmes, Read, and White, 2019), littering (Cialdini et al., 1990), recycling (Schultz, 1999) and towel reuse in hotels (Goldstein et al., 2008). Sparkman and Walton, 2017 found that dynamic descriptive social norms (which involve considering how the behaviour of others is changing over time) are more effective than static social norms (which considers the current behaviour of other people) in encouraging lower levels of meat consumption.

The information interventions used in this paper address all these major barriers

to reducing meat intake: the scientific knowledge gap, perceived inefficacy of action, personal relevance, cognitive dissonance and social norms. A prominent drawback of prior information interventions promoting reduced meat intake is the frequent reliance on a 'no information' control group for estimating effect sizes of the interventions (for instance, Graham and Abrahamse, 2017; Verain et al., 2017), potentially leading to overestimation or bias since it is hard to know participants' prior knowledge. This study avoids this issue by using a control group which was offered a baseline amount of scientific information about the greenhouse gas emissions caused by farm animals. This control group is compared with 6 treatment conditions each of which is provided an additional statement pertaining to the barriers discussed, along with the control group information. A list of the treatment conditions used in this study, along with their supporting theories, is presented in table 2.1. Following our pre-registration, our principal objective is not to compare treatments, but to evaluate among the possible set of presently popular alternatives which might be likely to have at least a short term effect and why. This falls down to the principle that if any of these approaches are effective, they should have at least a short term impact, and therefore provide a basis for more substantial change.

Table 2.1: Information interventions to encourage reduced meat consumption to combat climate change

Intervention	**Supporting Theory**	**Additional Information**	**Relevant Papers**
More scientific information	Deficit model of science communication	Effect of carbon dioxide emissions	Graham and Abrahamse, 2017
Efficacy salience	Response efficacy, Salience	CO_2 equivalent of eating less meat expressed in terms of miles driven in a car	Salomon et al., 2017; Hart and Feldman, 2016b; Chang and Lee, 2010
Health information	Personal relevance	Health impact of meat	Bertolotti et al., 2016; Vainio et al., 2018
Animal welfare	Cognitive Dissonance	Potential animal lives saved through reduced meat consumption	Tian et al., 2016; Kunst and Hohle, 2016; Feinberg et al., 2019
Social norms	Social norms approach	Proportion of the public that is actively eating more plant-based foods	Sparkman and Walton, 2017
Social Efficacy	Response efficacy, Social influence	Social effect of less meat intake	Salomon et al., 2017; Hart and Feldman, 2016b; Wolske et al., 2020; Weeks et al., 2015

The study was run as an online survey experiment. Subjects were initially asked about actions they could take to mitigate climate change and then the number of days they consumed meat in a week. Subjects then received the intervention followed by a set of questions similar to those seen at the beginning. The paper examines the influence of the interventions on the difference between post and prior beliefs as well as the intended reduction in weekly meat consumption (referred to as 'Δ Days'). Together, these provide insight into the instantaneous impact of each intervention.

Prior research (Lorenzoni et al., 2007; Bostrom et al., 2013) has highlighted the need for targeting and tailoring information to specific subject groups to encourage public engagement in climate action. Accordingly, the paper collects information regarding baseline beliefs and practices to help formulate strategies for targeted messaging. As moral convictions related to meat intake have been observed to impact behaviour (Feinberg et al., 2019), the study also measures moral offence taken by subjects on being told the consequences of meat consumption. More moral offense among those who eat more meat would indicate that meat eating may be perceived as a sacred value (Hanselmann and Tanner, 2008), and therefore reduce the impact of information interventions.

Towards the end of the study, subjects are asked to report the evidence offered to them during the survey as a measure of the recall value of the different treatments. Finally, the paper adds a directly-incentivised element by allowing participants to donate funds given to them in the experiment to a climate change charity, and tests the impact of each treatment on decision to donate.

The results of the study show that compared to a control group, which was provided baseline scientific information about the environmental impact of meat, an additional statement providing scientific information or a statement designed to make individual efficacy salient led to the strongest impacts on planned meat intake. For example, these led individuals to intend to reduce their dietary carbon dioxide (CO_2) emissions by more than 30 kg/year. If we extrapolate the results to a population the size of the U.S., the interventions would reduce the production of CO_2-equivalents by approximately 10 million tonnes per year. Note that these estimates should be interpreted cautiously as they rely on large-scale and persistent impact of the information interventions. Nonetheless they can be useful for comparison with future studies. We also find results supporting the use of targeted information, based on prior beliefs and individual characteristics. Providing information on the health benefits of a plant-based diet was effective in encouraging subjects with prior health concerns to opt for reduced meat con-

sumption. Our results also indicate that more scientific information and messages about social norms can encourage donation to a climate change charity among those who report low baseline ecological concerns. Further, we measure the recall value of the information interventions using text-based cosine similarity between information provided during the study and information recalled by subjects at the end. We found that although efficacy salience had the highest recall value, the difference compared to the control group was insignificant.

In addition, the more meat a person consumed, the less they acknowledged the relationship between climate and diet, the more resistant they were to new information, the lower their efficacy beliefs, and the more likely they were to take moral offence at being told the consequences of their behaviour. This constellation of results strongly indicates directional motivated reasoning (Kunda, 1990) or strategic ignorance (Carrillo and Mariotti, 2000), impairing the ability of meat eaters to objectively consider scientific information about meat consumption.

This study contributes to ongoing research exploring effectiveness of interventions in encouraging a reduced-meat diet (Bertolotti et al., 2016; Fehrenbach, 2015; Graham and Abrahamse, 2017; Vainio et al., 2018; Verain et al., 2017), by conducting a comparative evaluation of 6 interventions against a control. The study also adds to research finding evidence of *motivated reasoning* among individuals exposed to information related to climate change (Druckman and McGrath, 2019; Hart and Nisbet, 2012). More broadly, the paper contributes to studies evaluating interventions to 'nudge' sustainable behaviours (Allcott and Rogers, 2014; Allcott, 2011; Cialdini et al., 1990; Hafner, Elmes, and Read, 2019), specifically studies using information interventions (Goldstein et al., 2008; Hafner, Elmes, Read, and White, 2019). Lastly, the paper complements research exploring narrative approaches to presenting information (Shiller, 2017; Slater et al., 2003; McQuiggan et al., 2008).

Overall, this study makes a novel attempt to evaluate the effectiveness of six information interventions - framed using six different theoretical viewpoints - against a control group. Our results offer guidance to future researchers and policy-makers about information strategies for large-scale communication campaigns that are likely to be most effective in reducing meat consumption and producing corresponding reductions in emissions.

The rest of the paper is structured as follows: Section 2.2 describes the experiment design used for the study. Next, Section 2.3 presents the results of the study and Section 2.4 discusses the key findings of the paper. Lastly, Section 2.5 concludes.

2.2 Methodology

2.2.1 Experiment Design

The study was conducted as an online survey-based experiment. The experiment began with a series of questions designed to assess baseline ecological concerns. For this the 7 item questionnaire proposed by Stedman, 2004 was used. This questionnaire is based on the "New Environmental Paradigm" (NEP), developed by Dunlap and Van Liere, 1978. Each item on the questionnaire could be answered on a scale of 1 to 5, with ecological concern rising from 1 to 5. The average of the answer to these 7 questions was calculated as a measure of prior ecological concern for each subject. Related to this, subjects were also asked how concerned were they about climate change and how concerned their friends thought they were about climate change on a 5-point scale where 1 was not at all concerned and 5 extremely concerned.[2] This was followed by asking subjects what they believe are effective ways of combating climate change, for which they were allowed to write any answers that came to mind. This provided us with a knowledge base from which to compare the effectiveness of our various interventions.

Next, subjects were asked their baseline efficacy beliefs. Efficacy beliefs were calculated by taking the average value of the answers to two questions. They were asked the degree to which they agreed with the following statements on a scale of 1 to 5 where 1 is Strongly Disagree and 5 is Strongly Agree: (1) Individuals can influence climate change and (2) Collectively humans have little influence on climate change (*reverse coded*). Subjects were then asked about the number of days they consume meat in a week to measure baseline or pre-intervention frequency of meat consumption. This was followed by asking participants the degree to which they agree with the statement that "there is a relationship between climate change and people's food choices" which provides a baseline level of issue importance and follows Hart and Feldman, 2016a.

Participants were then randomly allocated into either the control group or one of 6 treatment conditions. Depending on the condition they were assigned to, they faced a different information intervention related to meat consumption and climate change. The exact information provided to the subjects in the different conditions is presented in Appendix 2.B. Following the intervention, subjects

[2]These two measures were elicited simply to confirm positive correlations with the baseline or prior ecological concern measure. As expected, prior ecological concern was significantly positively correlated with answers to both questions, i.e., subject's concern about climate change (with Pearson correlation coefficient (r) = 0.6633 and $p < 0.01$) and how concerned their friends think they are about climate change (r = 0.4670 and $p < 0.01$).

were asked the degree to which they now agree with the statement that there is a relationship between climate change and people's food choices. When compared to baseline issue importance, this provided a basic measure of the effectiveness of each intervention in shifting perceptions about the importance of diet on climate change.

Post intervention subjects were asked the same efficacy belief questions as they were asked pre-intervention. We then asked participants the number of days in a week they planned to consume meat. This was subtracted from the baseline answer from before the intervention to produce a measure in days of the effect of each intervention (referred to as "Δ Days" in the paper) and provides us with the key dependent variable for the study. Participants were also again asked what likely actions they could undertake to reduce climate change together with the information that they could recall from the experiment, providing us with measures of salience and also a set of text data for each participant. To measure the degree of moral offence taken by subjects at being informed about the consequences of meat intake, they were asked if it was morally wrong to show people the consequences of their own behaviour. The subjects were further asked certain socio-demographic questions like age, gender, educational qualification, political beliefs and perceived social network (the number of people who might notice if they changed their diet). Subjects were also asked if any close family member had ever suffered from any heart disease, stroke, cancer, diabetes, high blood pressure or high cholesterol. This question was asked to ascertain if the subject had any genetic history of diseases.

The experiment concluded by asking participants if they would like to keep or donate to a climate change related charity a bonus amount offered to them. This provided us with an incentivised measure of the external validity of the various information interventions. A detailed experimental script for the entire experiment is provided in Appendix 2.C.

2.2.2 Logistics

The experiment was designed using oTree (Chen et al., 2016) and it was conducted online on Amazon Mechanical Turk (MTurk). The subject pool were registered MTurk workers in the U.S. The study was preregistered with the AEA RCT registry (Bose et al., 2019; Bose et al., 2020). Ethical approval for the study was granted by University of Warwick Humanities & Social Sciences Research Ethics Committee (HSSREC). The experiment took approximately 10 minutes. There were 1458 observations collected in total, after removing those who had technical issues owing to browser/device compatibility. In order to remove noise,

originating from insincere responses, bots and autofill, the final dataset only included those subjects that spent more than 10 seconds on the intervention page and wrote more than 3 words when asked to recall the evidence presented during the study.

The final dataset comprised of 1220 subjects: 177 in the Control condition, 175 in Treatment 1 (More scientific information), 179 in Treatment 2 (Efficacy salience), 174 in Treatment 3 (Health information), 177 in Treatment 4 (Animal welfare), 164 in Treatment 5 (Social norms) and 174 in Treatment 6 (Social efficacy). Data for the control and treatment groups 1, 3 and 6 were collected in December 2019 whereas data for treatment groups 2, 4 and 5 were collected in March 2020. This raises a potential issue: what if the salience of COVID19 or any other event which occurred between the two data-gathering exercises rendered our two sets of data incomparable? In order to exclude this possibility we collected an additional 88 control condition observations in March 2020 (and we used this data only as part of the following comparability exercise since it was not part of our pre-registered plan). We then compared the two sets of control group observations from December and March, failing to reject the null of equal means (using t-tests) and equal distributions (using Kolmogorov-Smirnov tests) for Δ Days, donation to charity, prior ecological concerns, efficacy beliefs (baseline and post intervention), and baseline and post-intervention meat intake of the two sets at the 5 and 10% significance levels. This suggests that there was no change in behaviour *related to our experiment* between December 2019 and March 2020.

All subjects received a participation fee of \$1.50 and a bonus of \$0.50 (which they could either keep for themselves or donate to a climate change charity in the final part of the experiment). The average earnings from the experiment was \$1.85, with 372 out of the 1220 subjects choosing to donate their bonus earnings.

2.3 Results

The results sections is divided into 3 main parts: the impact on beliefs, the impact on reduction in weekly meat consumption (Δ Days) and the impact on donations to charity. In the final part of the results section we also provide a brief discussion of the text used by participants when they are asked to recall the information they were shown, and a discussion of moral convictions. The summary statistics of the variables used in the paper are presented in the table 2.A.1 and the balance tests for the intervention groups are provided in table 2.A.2.

2.3.1 Baseline beliefs and practices

Only 111 (around 9%) of 1220 participants mentioned a dietary change from meat-based to plant-based foods when asked about personal actions they could take to mitigate climate change before the interventions (figure 2.A.1). The number of subjects mentioning eating less meat as an effective personal action went up to 633 (around 52%) post intervention. The interventions significantly increased the proportion of people that mentioned diet-related solutions as actions effective against climate change for the control and all 6 treatment groups. Paired t-tests rejected the null of no difference in baseline and post-intervention proportions for the control and all 6 treatment conditions ($p < 0.01$).

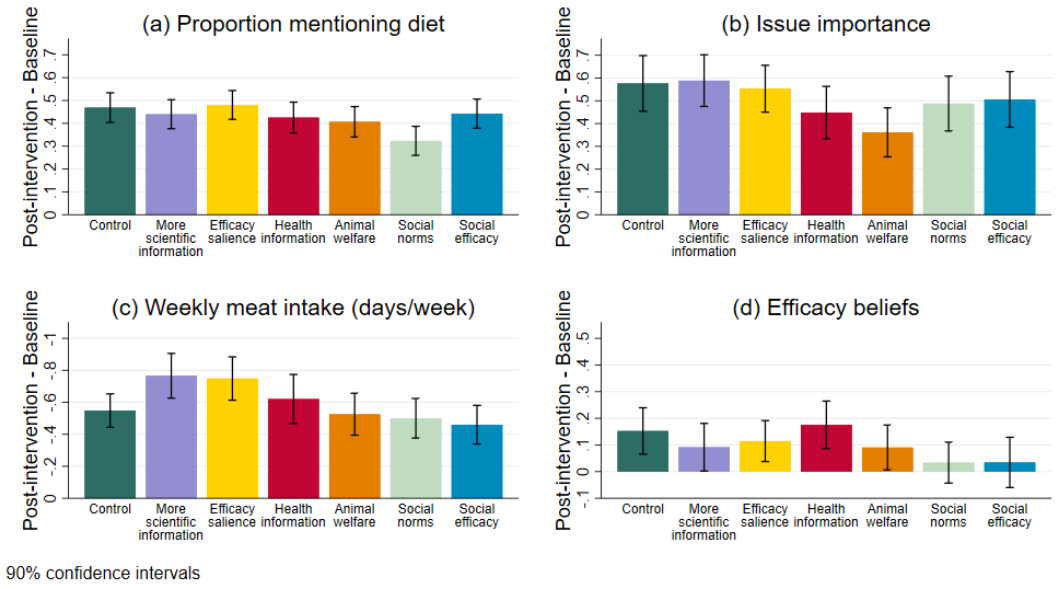
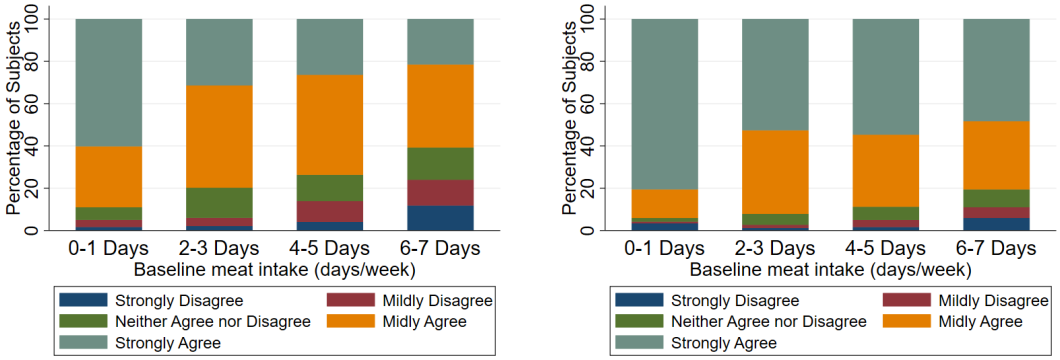


Figure 2.1: Difference between post-intervention and baseline values across the different interventions. Figure (a) plots the average increase in the proportion of people who mentioned dietary shift-related solutions as an action they can personally take to combat climate change post-intervention, compared to baseline. Figure (b) plots the average increase in perceived issue importance post-intervention compared to baseline perceived importance. Figure (c) plots the average difference between post-intervention and baseline weekly meat intake. Figure (d) plots the average increase in efficacy beliefs post-intervention compared to the baseline.

The difference between post-intervention and baseline beliefs and practices are presented in figure 2.1. The interventions significantly enhanced beliefs about issue importance (i.e. the degree to which subjects agree with the statement "There is a relationship between climate change and people's food choices") for the control and all 6 treatment conditions (results of a paired t-test for each condition significant with $p < 0.01$). It should be noted that the increase in

issue importance for the animal welfare group is significantly lower compared to the control group (t-statistic 2.1730 and $p < 0.05$) as shown in Figure 2.1 (b), which may be a result of individuals believing animal welfare is an unrelated moral argument (Feinberg et al., 2019). Furthermore, all 6 treatment conditions and the control led to reduced intended meat consumption post-intervention ($p < 0.01$). Efficacy beliefs were also significantly higher for all treatments except social norms and social efficacy (results of paired t-tests were significant for control, health information and efficacy salience groups with $p < 0.05$ and for more scientific information and animal welfare groups with $p < 0.10$).



(a) Baseline

(b) Post intervention

Figure 2.2: Relationship between baseline number of days of meat consumption per week and the (a) pre-intervention and (b) post-intervention degree to which subjects agree that there is a relationship between climate change and people's food choices.

We discuss meat consumption and beliefs in more detail below, but the baseline frequency of meat consumption was negatively correlated with perceived issue importance prior to the intervention (Pearson correlation coefficient (r) = -0.2848 with $p < 0.01$) as shown in Figure 2.2. The relation between baseline weekly meat consumption and post-intervention beliefs in the importance of diet to climate change was also negative (r = -0.1846 with $p < 0.01$). In other words, both before and after the intervention, more frequent meat-eaters found the issue less important. Nonetheless, because frequent meat-eaters have more potential range to change their views, frequency of meat consumption and the change in issue importance between post and pre-intervention is positively correlated (r = 0.1544 with $p < 0.01$). We also observed a significant negative correlation of baseline frequency of meat consumption with both baseline efficacy beliefs (r = -0.1107 with $p < 0.01$) and post-intervention efficacy beliefs (r =-0.1140 with $p < 0.01$). Finally, there was also a significant negative correlation between baseline frequency of meat intake and prior ecological concern (r = -0.2034 with $p < 0.01$).

Thus, frequent meat-eaters reported being less concerned about the climate and felt they had limited abilities to ameliorate climate outcomes.

2.3.2 Effect of the information interventions on Δ Days

Table 2.2 reports the results from an OLS regression with the dependent variable being the intended change in number of days of weekly meat consumption (Δ Days). All regressions in the paper were run with standardised explanatory variables and the tables report robust standard errors. Column 1 examines the impact of being a member of one of the different treatment (intervention) groups on Δ Days, relative to the control group. Columns 2 and 3 show that the results of column 1 are robust to the inclusion of control variables such as prior ecological concern, pre-intervention weekly meat consumption, baseline efficacy beliefs and the socio-demographic variables, age, female, democrat and education (a dummy which equals 1 if the subject reported that the highest degree obtained by them is greater than a school degree and 0 otherwise). Column 3 shows that the *more scientific information* and the *efficacy salience* interventions significantly reduce planned meat consumption per week by 0.22 days ($p < 0.05$) and 0.21 days ($p < 0.05$) more, respectively, compared to the control group.

To get a better grasp of the effect size of the treatments we convert the estimates to CO_2 emissions. In terms of CO_2 emissions, by shifting away from meat 1 day per week, an individual can reduce their CO_2 emissions by around 161.11 kg CO_2 per year (calculation based on estimates provided by Weber and Matthews, 2008; see Appendix 2.D). Thus, exposure to more scientific information and efficacy salience conditions can reduce CO_2 emissions by 35.4 kg/year and 33.8 kg/year more compared to the control group, respectively. The effect sizes of the treatments are presented in figure 2.3 (a).

The higher the prior ecological concerns and baseline efficacy beliefs of the participant, the greater the intended change in diet. A 1 standard deviation increase in prior ecological concern and baseline efficacy beliefs reduces planned weekly meat intake by 0.10 and 0.15 days (both with $p < 0.01$), respectively. Also, there is a significant ($p < 0.01$) positive impact of pre-intervention diet on the dependant variable. Among the demographic variables, being female has a positive significant ($p < 0.05$) effect on intention to reduce meat consumption. Indeed, females in our study consumed less meat than males prior to the study (4.04 days/week versus 4.36 days/week, difference significant with $p < 0.01$ and t-statistic 2.7207). The change in meat consumption is consistent with two separate empirical observations in the literature. Firstly, females have been observed to display greater concern and knowledge about climate change than males (McCright, 2010). Sec-

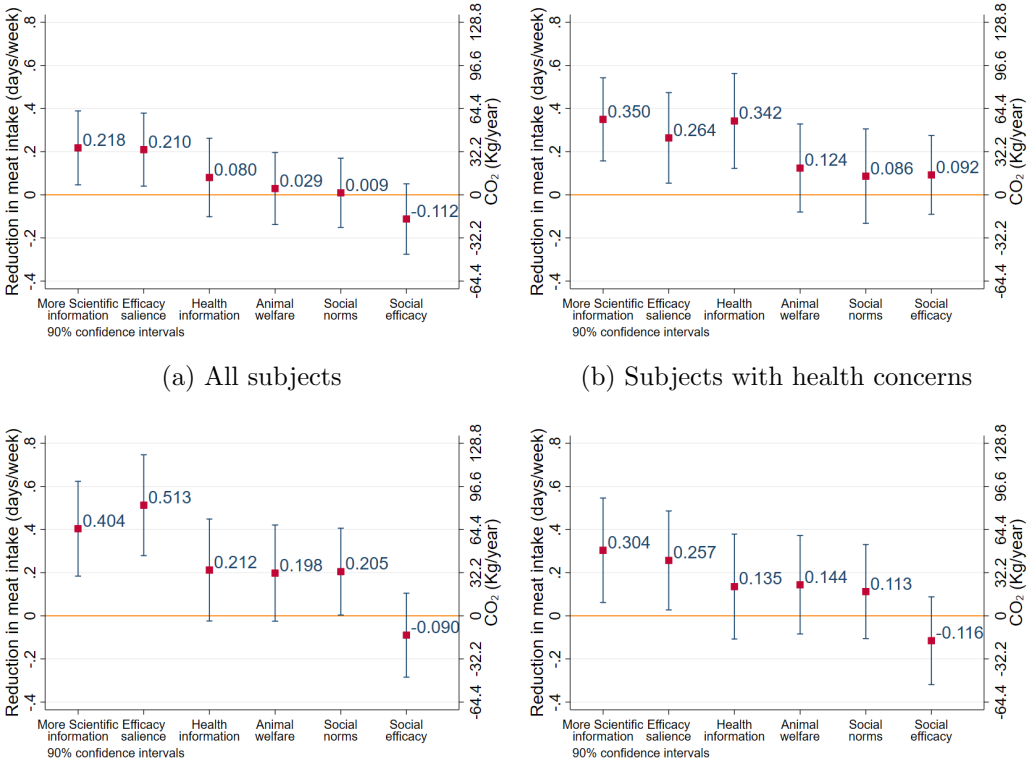
Table 2.2: Impact of information interventions on planned reduction in number of days of meat consumption per week

	Dependent Variable: Δ Days		
	(1)	(2)	(3)
More scientific information	0.2177**	0.2339**	0.2176**
	(0.1060)	(0.1039)	(0.1041)
Efficacy salience	0.2006*	0.2118**	0.2095**
	(0.1040)	(0.1027)	(0.1028)
Health information	0.0727	0.0828	0.0803
	(0.1127)	(0.1113)	(0.1106)
Animal welfare	-0.0226	0.0295	0.0293
	(0.1020)	(0.1013)	(0.1015)
Social norms	-0.0480	0.0130	0.0093
	(0.0984)	(0.0979)	(0.0975)
Social efficacy	-0.0883	-0.0963	-0.1124
	(0.0973)	(0.0983)	(0.0992)
Prior Ecological Concern		0.1007***	0.0975***
		(0.0298)	(0.0304)
Pre-intervention Diet		0.2115***	0.2193***
		(0.0286)	(0.0287)
Baseline Efficacy Beliefs		0.1559***	0.1541***
		(0.0320)	(0.0327)
Age			0.0008
			(0.0272)
Female			0.1344**
			(0.0636)
Democrat			0.0227
			(0.0626)
Education			0.0788
			(0.0663)
N	1220	1220	1220

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

ondly, males tend to more strongly associate meat with healthiness than females (Love and Sulikowski, 2018) and hence might be less inclined to reduce their weekly meat intake. Table 2.2 is replicated with the dependent variable as proportional planned reduction in weekly meat intake i.e $\frac{\Delta Days}{Pre-intervention\ Diet}$ and the results are presented in table 2.A.3. Exposure to the efficacy salience treatment has the biggest effect in terms of proportionate reduction in weekly meat consumption, compared to the control group, and the effect remains significant after controlling for baseline characteristics and socio-demographic variables.

To investigate the effects of *targeted messages*, we examine if the information



(a) All subjects

(b) Subjects with health concerns

(c) Subjects with high baseline efficacy beliefs

(d) Frequent meat eaters

Figure 2.3: Regression coefficients showing the effect size of the information interventions on intended reduction in number of days of meat consumption per week across different targeted groups. The right hand axis plots the equivalent of reduced meat consumption (days per week) in terms of reduced CO_2 emissions per year. The regression coefficients plotted show the effect size after controlling for baseline beliefs and practices, and socio-demographic characteristics.

interventions had a differential impact on participants depending upon their prior health concerns, baseline efficacy beliefs, and baseline weekly frequency of meat consumption.

We first look at the effect of any prior family history of diseases on intended reduction in meat intake. Table 2.3 reports the results. The positive and significant ($p < 0.01$) interaction term between health information and history of diseases indicates that the health information condition has a significantly bigger effect on Δ Days if the subject has a family history of diseases, compared to when they do not (figure 2.A.4). When the subject has prior health concerns the health information condition is significantly more effective than the control group, where as when the subject has no prior health concerns the control group information is more effective as shown in table 2.A.4. Figure 2.3 (b) and table 2.A.4 show that when the subject has a history of family diseases, along with more scientific information and efficacy salience, the health information condi-

tion also has a significantly larger effect on Δ Days, relative to the control group. Health information increases Δ Days by 0.34 days/week more than the control ($p < 0.05$) i.e. a reduction in CO_2 emissions by 54.78 kg/year per person.

Table 2.3: Impact of health information intervention on planned reduction in number of days of meat consumption per week depending on history of diseases in the family

	Dependent Variable: Δ Days		
	(1)	(2)	(3)
Health information × Disease	0.6685***	0.6786***	0.6913***
	(0.2387)	(0.2356)	(0.2359)
Health information	-0.3511*	-0.3522*	-0.3622*
	(0.1958)	(0.1931)	(0.1923)
Disease	-0.3144**	-0.3371**	-0.3510**
	(0.1457)	(0.1458)	(0.1486)
Prior Ecological Concern		0.0842	0.0617
		(0.0575)	(0.0599)
Pre-intervention Diet		0.1510***	0.1571***
		(0.0560)	(0.0564)
Baseline Efficacy Beliefs		0.1086*	0.0874
		(0.0600)	(0.0583)
Age			-0.0094
			(0.0501)
Female			0.2195*
			(0.1280)
Democrat			0.2017*
			(0.1152)
Education			0.0436
			(0.1285)
N	351	351	351

Standard errors in parentheses. Statistical significance indicated as follows: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Disease is a dummy variable which equals 1 if the subject indicated that they had a history of diseases in the immediate family and 0 otherwise.

We also look at the differential effect of the interventions on subjects with varied baseline efficacy beliefs and pre-intervention diet. In table 2.A.5, the interaction terms of the dummy variable 'High Efficacy Beliefs' (= 1 if subject reported ≥ median baseline efficacy beliefs and 0 otherwise) with more scientific information, efficacy salience, animal welfare and social norms conditions are significant and positive. This implies that, the impact of these interventions on Δ Days is significantly greater for subjects with high baseline efficacy beliefs than those with low efficacy beliefs (depicted in figure 2.A.5). Figure 2.3 (c) and table 2.A.6 show that when the subject has high baseline efficacy beliefs, along with more

scientific information and efficacy salience, social norms has a significantly ($p < 0.10$) larger effect on Δ Days, relative to the control group. Figure 2.3 (d) and table 2.A.8 show that for frequent meat eaters (pre-intervention meat intake $\geq$ the median value i.e. 4 days/week), more scientific information and efficacy salience are the only treatment conditions which have a significantly larger effect on Δ Days, relative to the control group.

2.3.3 Donation behaviour, evidence recall and moral offence

After the intervention, all subjects were offered the opportunity to either donate a bonus sum of $ 0.50 to a climate change charity or to keep it for themselves. A probit model evaluating the effect of the interventions on donation (table 2.4), revealed that after the inclusion of sensible control variables, no treatment had a significant effect on donation behaviour. Prior ecological concern is a strong positive determinant of donation behaviour. An increase in prior ecological concern by 1 standard deviation increases the probability of donation by 12 percentage points. Column 3 shows that, for those with low prior ecological concern (below median), more scientific information and social norms conditions increased the probability of donation by 14.1 ($p<0.05$) and 14.9 ($p<0.05$) percentage points more than control, respectively. Social efficacy also had a small significant ($p <0.10$) impact on donation behaviour for those with low prior concern. The average donation levels across the 7 groups are depicted in figure 2.A.7.[3]

The paper also evaluates the text reported by the subjects when they were asked to recall the evidence presented to them during the course of the study. For each participant we calculated a cosine similarity score between the words they reported and the evidence they were shown which varied by treatment. Cosine similarity was calculated by converting the reported evidence (R) and actual evidence (A) to vector representations (after removing punctuation and stopwords from both) and then by measuring the cosine of the angle between these two vectors. A cosine similarity of 1 would mean that the reported and actual evidence are perfectly similar.

$$Similarity(R, A) = cos\theta = \frac{R \cdot A}{||R||||A||} \tag{2.1}$$

[3]The significant negative effect of baseline efficacy beliefs on donation behaviour is potentially because subjects with high perceived efficacy believe in the power of human actions taken at a personal level in combating climate change. This perhaps makes them less inclined to contribute a portion of their earnings to a climate change organisation. Although, any conclusive inference is beyond the scope of the present study.

Table 2.4: Impact of information interventions on decision to donate bonus earnings

	Dependent Variable: Prob(donate=1)			
	(1) All	(2) All	(3) Low Priors	(4) High Priors
More scientific information	0.0600	0.0335	0.1415**	-0.0471
	(0.0480)	(0.0455)	(0.0594)	(0.0652)
Efficacy salience	0.0586	0.0221	0.0727	-0.0227
	(0.0477)	(0.0446)	(0.0553)	(0.0643)
Health information	0.0044	-0.0149	0.0313	-0.0360
	(0.0466)	(0.0438)	(0.0548)	(0.0637)
Animal welfare	0.0621	0.0203	0.0456	0.0190
	(0.0479)	(0.0444)	(0.0562)	(0.0659)
Social norms	0.1116**	0.0708	0.1489**	0.0179
	(0.0499)	(0.0467)	(0.0632)	(0.0655)
Social efficacy	0.0619	0.0589	0.1101*	0.0362
	(0.0481)	(0.0466)	(0.0626)	(0.0648)
Prior Ecological Concern		0.1166***	0.1363***	-0.0338
		(0.0142)	(0.0276)	(0.0482)
Pre-intervention Diet		-0.0809***	-0.0857***	-0.0651***
		(0.0121)	(0.0164)	(0.0170)
Δ Days		0.0250*	0.0216	0.0225
		(0.0129)	(0.0178)	(0.0173)
Baseline Efficacy Beliefs		-0.1078***	-0.0541***	-0.1433***
		(0.0132)	(0.0178)	(0.0183)
Controls	No	Yes	Yes	Yes
N	1220	1220	545	675

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependant variable is the binary variable donate, which equals 1 if the subject donated their bonus earnings to the climate change charity and 0 otherwise. The table reports the average marginal effects from Probit regressions. Columns 3 and 4 show the impact of the interventions on subjects with low and high prior ecological concern, respectively. The control variables are the subject's age and dummy variables for being female, democrat and having an educational qualification greater than a school degree. Coefficients for the control variables (omitted here for parsimony) are presented in the Appendix in table 2.A.9.

The group which received efficacy salience information had the highest mean cosine similarity score of 0.36. However, the difference compared to the control group score (0.33) was not significant (figure 2.A.9).

After the intervention, we included the question "Is it morally wrong to show people the consequences of their own behaviour?" which was answered on a scale of 1 to 5 where 1 is very morally wrong and 5 is very morally right. We found a significant positive correlation between the answer to the morality question and

Δ Days, with Pearson's correlation coefficient .0998 and $p < 0.01$ (figure 2.A.10). This implies that subjects who have committed to a greater behavioural change are less inclined to take any moral offence from being shown the consequences of own behaviour. However, people who had a higher baseline meat intake were more inclined to report moral offence at being shown the consequences of their own behaviour (Pearson correlation coefficient -0.0655 and $p < 0.05$).

2.4 Discussion

There is shockingly low awareness among individuals about the impact of meat consumption on climate change, despite the United Nation's Intergovernmental Panel on Climate Change releasing new evidence to this effect in 2019 (Schiermeier, 2019). In the present study, only 9% of the subjects at baseline listed diet change as effective in combating climate change. This study is a comparative evaluation of six different information interventions to increase this awareness and encourage a dietary shift away from meat-based foods.

Our work highlights two interventions that may have the most impact on planned behaviour. Compared to a control group with baseline scientific information, the largest influence was seen with an additional statement providing more scientific facts or a statement highlighting response efficacy (*efficacy salience*). If such interventions were effective at scale and maintained over time, then extrapolating our results to a population the size of the U.S., these interventions would reduce the production of CO_2-equivalents by approximately 10 million tonnes per year, which is useful for comparison with other studies in the future.

While the more scientific information intervention supports the deficit model of science communication, the efficacy salience intervention validates the importance of salient and vivid (Chang and Lee, 2010) response efficacy information. The success of the deficit model highlights a pre-existing knowledge gap about the relationship between climate change and diet which can be addressed with more scientific information (Bauer et al., 2007). Also, consistent with Bandura, 1977), we see that it is not enough that individuals know that they can engage in a behaviour (i.e. consume less meat). In fact, it is also necessary that the individuals are aware that consuming less meat will have the desired environmental impact, i.e. their own actions can make a difference, and by how much presented in a currency they can understand.

Our results also suggest that a targeted messaging approach, dependent upon prior beliefs and individual characteristics, can have a significant impact. For subjects with pre-existing health concerns, providing information on the health

benefits of a plant-based diet was effective in encouraging reduced intended meat intake. This is consistent with the notion that health information serves as a form of intrinsic motivation and is therefore more likely to be effective when the subject already has prior health concerns. This may also be consistent with motivated reasoning (Kunda, 1990) since behavioural change in this instance may well be in a participant's own best interest if prior health concerns are a valid predictor of future poor health. A summary of our key results and how they work for different population groups is provided in table 2.5.

Table 2.5: Targetted messaging approach to encourage reduced meat consumption. The effectiveness of different information interventions for different subject groups, relative to the control group.

	More scientific information	**Efficacy salience**	**Health information**	**Animal welfare**	**Social norms**	**Social efficacy**
All	✓	✓				
Prior health concerns	✓	✓	✓			
High efficacy	✓	✓			✓	
Frequent meat eaters	✓	✓				

Even more dramatically, subjects who ate meat more are less inclined to acknowledge the relationship between climate change and diet at baseline and are less inclined to accept new evidence indicating the same. This is consistent with (Tobler et al., 2011) who find that frequent meat eaters perceive the environmental benefits of reducing meat intake as small. This is also consistent with the idea of directional motivated reasoning (Kunda, 1990) or strategic ignorance (Carrillo and Mariotti, 2000) where people are willing to modify their own beliefs to the extent allowed by self-justification. Frequent meat eaters were more likely to ignore the relationship between climate change and food choices, a form of "science denial" (Washburn and Skitka, 2018) that may not be obvious to those who frequently eat meat.

We also elicited the moral offence taken by subjects on being told the consequences of meat consumption to investigate whether or not meat might be seen as a sacred value, one that is morally licensed irrespective of its cost. This would make moral value of meat consumption resistant to trade-offs with climate or animal suffering (Hanselmann and Tanner, 2008). We observe that people who had a higher baseline meat intake were more inclined to report moral offence at

being shown the consequences of their own behaviour. These results are consistent with holding meat as a sacred value and might also explain the reluctance of those who consume particularly high amounts of meat to recognise the link between diet and climate change.

The paper provides further insight into the effectiveness of the information interventions by evaluating the effect of the treatments on decision to donate to a climate change charity and by measuring the recall value of the information provided. It is observed that using messages with more scientific information or social norm frames can encourage donation to climate change charities among those who report low baseline ecological concerns, compared to the control group. On recall, although efficacy salience was observed to have the highest recall value among the treatment groups, the difference compared to the control group was insignificant which leads us to believe that our results cannot be explained with reference to the relative memorability of the treatments.

Overall, in the absence of a comprehensive and comparative evaluation of different communication campaigns to encourage a shift towards plant-based foods, our study for the first time compares the effectiveness of six different information interventions to reduce meat intake to a control group with baseline scientific information. The study recommends the use of clear scientific information and information highlighting that individual actions can make a difference in the path to mitigating climate change. The study also endorses the need for targeted information interventions for specific subgroups and stresses the crucial role played by frequent meat-eaters who are resistant to new information. The results of the study can inform policy related to climate change communication, ultimately resulting in significant environmental benefits.

2.5 Conclusion

There is a broad scientific consensus that reduced meat consumption can have significant positive effects upon the current climate change crisis. Despite this, there is a lack of awareness of the link between diet and climate change among the general population. This creates an opening for information provision to provide positive benefits. However, there are many alternative approaches to providing additional information and there is no existing consensus on the best approach. We provide a first attempt to compare and contrast the leading approaches with a control group which received baseline scientific information.

Our results show that compared to a control group, which was provided baseline scientific information about the environmental impact of meat intake, an addi-

tional statement about more similar scientific facts or a statement highlighting response efficacy of action can reduce an individual's CO_2 emissions by more than 30 kg/year. Extrapolating our results to a population the size of the U.S., our interventions could reduce the production of CO_2-equivalents by approximately 10 million tonnes per year. Our results also support a targeted messaging approach, dependent upon prior beliefs and individual characteristics, to promote behaviour change. For instance, for subjects with pre-existing health concerns, providing information on the health benefits of a plant-based diet was effective in encouraging reduced intended meat intake.

Additionally, the paper finds that using messages with more scientific information or social norms frames can encourage donation to climate change charities among those who display low baseline ecological concern. Furthermore, our results indicate that the more meat a person consumed, the less they acknowledged the relationship between climate and diet, the more resistant they were to new information, and the more likely they were to take moral offence at being told the consequences of their behaviour. This indicates the presence of *motivated reasoning* around meat consumption among frequent meat eaters.

Overall, our findings provide a general taxonomy that should guide future researchers and policy-makers about the most effective communication strategies to promote a meat-free or low-meat diet. Effectively, this could reduce an individual's carbon footprint, ultimately leading to substantial environmental benefits.

References

Allcott, H. (2011). Social norms and energy conservation [Special Issue: The Role of Firms in Tax Systems]. *Journal of Public Economics*, *95*(9), 1082–1095. https://doi.org/10.1016/j.jpubeco.2011.03.003

Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation. *American Economic Review*, *104*(10), 3003–37. https://doi.org/10.1257/aer.104.10.3003

Balch, G. I. (1974). Multiple Indicators in Survey Research: the concept "Sense of Political Efficacy". *Political Methodology*, 1–43.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bauer, M. W., Allum, N., & Miller, S. (2007). What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understanding of Science*, *16*(1), 79–95. https://doi.org/10.1177/0963662506071287

Bertolotti, M., Chirchiglia, G., & Catellani, P. (2016). Promoting change in meat consumption among the elderly: Factual and prefactual framing of health and well-being [Special Issue: Consumer behaviour in a changing world - Selected papers from the AAEA/EAAE joint seminar in Naples, March 25-27, 2015]. *Appetite*, *106*, 37–47. https://doi.org/10.1016/j.appet.2016.02.150

Bose, N., Hills, T., & Sgroi, D. (2019). Climate change and diet: A persuasive message approach to increasing cooperation [November 25]. *AEA RCT Registry*. https://doi.org/10.1257/rct.5069-1.0

Bose, N., Hills, T., & Sgroi, D. (2020). Climate change and diet: A persuasive message approach to increasing cooperation - wave 2 [March 24]. *AEA RCT Registry*. https://doi.org/10.1257/rct.5584-1.0

Bostrom, A., Böhm, G., & O'Connor, R. E. (2013). Targeting and tailoring climate change communications. *WIREs Climate Change*, *4*(5), 447–455. https://doi.org/10.1002/wcc.234

Carrillo, J. D., & Mariotti, T. (2000). Strategic ignorance as a self-disciplining device. *The Review of Economic Studies*, *67*(3), 529–544. https://doi.org/10.1111/1467-937X.00142

Chang, C.-T., & Lee, Y.-K. (2010). Effects of message framing, vividness congruency and statistical framing on responses to charity advertising. *International Journal of Advertising*, *29*(2), 195–220. https://doi.org/10.2501/S0265048710201129

Chen, D. L., Schonger, M., & Wickens, C. (2016). Otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Cialdini, R., Reno, R., & Kallgren, C. (1990). A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places. *Journal of Personality and Social Psychology*, *58*, 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015

de Boer, J., de Witt, A., & Aiking, H. (2016). Help the climate, change your diet: A cross-sectional study on how to involve consumers in a transition to a low-carbon society. *Appetite*, *98*, 19–27. https://doi.org/10.1016/j.appet.2015.12.001

Doherty, K. L., & Webler, T. N. (2016). Social norms and efficacy beliefs drive the Alarmed segment's public-sphere climate actions. *Nature Climate Change*, *6*(9), 879–884. https://doi.org/10.1038/nclimate3025

Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, *9*(2), 111–119. https://doi.org/10.1038/s41558-018-0360-1

Dunlap, R. E., & Van Liere, K. D. (1978). The new environmental paradigm: A proposed measuring instrument and preliminary results. *Journal of Environmental Education*, *9*, 10–19.

FAO. (2006). Livestock a major threat to environment [Accessed May 6, 2021]. http://www.fao.org/newsroom/en/news/2006/1000448/index.html

Fehrenbach, K. S. (2015). *Designing messages to reduce meat consumption: A test of the extended parallel process model* (PhD dissertation) [Arizona State University].

Feinberg, M., Kovacheff, C., Teper, R., & Inbar, Y. (2019). Understanding the process of moralization: How eating meat becomes a moral issue. *Journal of Personality and Social Psychology*, *117*(1), 50–72. https://doi.org/10.1037/pspa0000149

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press.

Goldstein, N., Cialdini, R., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482. https://doi.org/10.1086/586910

Gonzales, M. H., Aronson, E., & Costanzo, M. A. (1988). Using social cognition and persuasion to promote energy conservation: A quasi-experiment 1. *Journal of Applied Social Psychology*, *18*(12), 1049–1066.

Graham, T., & Abrahamse, W. (2017). Communicating the climate impacts of meat consumption: The effect of values and message framing. *Global Environmental Change*, *44*, 98–108. https://doi.org/10.1016/j.gloenvcha.2017.03.004

Hafner, R., Elmes, D., Read, D., & White, M. P. (2019). Exploring the role of normative, financial and environmental information in promoting uptake of energy efficient technologies. *Journal of Environmental Psychology*, *63*, 26–35. https://doi.org/10.1016/j.jenvp.2019.03.004

Hafner, R. J., Elmes, D., & Read, D. (2019). Promoting behavioural change to reduce thermal energy demand in households: A review. *Renewable and Sustainable Energy Reviews*, *102*, 205–214. https://doi.org/10.1016/j.rser.2018.12.004

Hanselmann, M., & Tanner, C. (2008). Taboos and conflicts in decision making: Sacred values, decision difficulty, and emotions. *Judgment and Decision Making*, *3*(1), 51–63. https://doi.org/10.5167/uzh-5948

Harguess, J. M., Crespo, N. C., & Hong, M. Y. (2020). Strategies to reduce meat consumption: A systematic literature review of experimental studies. *Appetite*, *144*, 104478. https://doi.org/10.1016/j.appet.2019.104478

Hart, P. S., & Feldman, L. (2016a). The Impact of Climate Change–Related Imagery and Text on Public Opinion and Behavior Change. *Science Communication*, *38*(4), 415–441. https://doi.org/10.1177/1075547016655357

Hart, P. S., & Feldman, L. (2016b). The influence of climate change efficacy messages and efficacy beliefs on intended political participation. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0157658

Hart, P. S., & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research*, *39*(6), 701–723. https://doi.org/10.1177/0093650211416646

Higgs, S., & Thomas, J. (2016). Social influences on eating. *Current Opinion in Behavioral Sciences*, *9*, 1–6. https://doi.org/10.1016/j.cobeha.2015.10.005

Huang, T., Yang, B., Zheng, J., Li, G., Wahlqvist, M. L., & Li, D. (2012). Cardiovascular disease mortality and cancer incidence in vegetarians: a meta-analysis and systematic review. *Annals of nutrition & metabolism*, *60*(4), 233–240. https://doi.org/10.1159/000337301

Kunda, Z. (1990). The case for motivated reasoning. *Psychological bulletin*, *108*(3), 480.

Kunst, J. R., & Hohle, S. M. (2016). Meat eaters by dissociation: How we present, prepare and talk about meat increases willingness to eat meat by reducing empathy and disgust. *Appetite*, *105*, 758–774. https://doi.org/10.1016/j.appet.2016.07.009

Kunst, J. R., & Palacios Haugestad, C. A. (2018). The effects of dissociation on willingness to eat meat are moderated by exposure to unprocessed meat: A cross-cultural demonstration. *Appetite*, *120*, 356–366. https://doi.org/10.1016/j.appet.2017.09.016

Lacroix, K., & Gifford, R. (2019). Reducing meat consumption: Identifying group-specific inhibitors using latent profile analysis. *Appetite*, *138*, 233–241. https://doi.org/10.1016/j.appet.2019.04.002

Lewandowsky, S., Cook, J., Fay, N., & Gignac, G. E. (2019). Science by social media: Attitudes towards climate change are mediated by perceived social consensus. *Memory & Cognition*, *47*(8), 1445–1456. https://doi.org/10.3758/s13421-019-00948-y

Lewandowsky, S., Gignac, G. E., & Vaughan, S. (2013). The pivotal role of perceived scientific consensus in acceptance of science. *Nature Climate Change*, *3*(4), 399–404. https://doi.org/10.1038/nclimate1720

Lorenzoni, I., Nicholson-Cole, S., & Whitmarsh, L. (2007). Barriers perceived to engaging with climate change among the uk public and their policy implications. *Global Environmental Change*, *17*(3), 445–459. https://doi.org/10.1016/j.gloenvcha.2007.01.004

Loughnan, S., Haslam, N., & Bastian, B. (2010). The role of meat consumption in the denial of moral status and mind to meat animals. *Appetite*, *55*(1), 156–159. https://doi.org/10.1016/j.appet.2010.05.043

Love, H. J., & Sulikowski, D. (2018). Of Meat and Men: Sex Differences in Implicit and Explicit Attitudes Toward Meat. *Frontiers in psychology*, *9*, 559. https://doi.org/10.3389/fpsyg.2018.00559

Macdiarmid, J. I., Douglas, F., & Campbell, J. (2016). Eating like there's no tomorrow: Public awareness of the environmental impact of food and reluctance to eat less meat as part of a sustainable diet. *Appetite*, *96*, 487–493. https://doi.org/10.1016/j.appet.2015.10.011

McCright, A. M. (2010). The effects of gender on climate change knowledge and concern in the American public. *Population and Environment*, *32*(1), 66–87. https://doi.org/10.1007/s11111-010-0113-1

McQuiggan, S. W., Rowe, J. P., Lee, S., & Lester, J. C. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent tutoring systems* (pp. 530–539). Springer Berlin Heidelberg.

Monroe, M. C., Plate, R. R., Oxarart, A., Bowers, A., & Chaves, W. A. (2019). Identifying effective climate change education strategies: A systematic review of the research. *Environmental Education Research*, *25*(6), 791–812. https://doi.org/10.1080/13504622.2017.1360842

NASA. (2020). The causes of climate change [Accessed May 6, 2021]. https://climate.nasa.gov/causes/

Ockwell, D., Whitmarsh, L., & O'Neill, S. (2009). Reorienting Climate Change Communication for Effective Mitigation: Forcing People to be Green or Fostering Grass-Roots Engagement? *Science Communication*, *30*(3), 305–327. https://doi.org/10.1177/1075547008328969

PETA. (2013). How many animals have you saved? find out! [Accessed May 6, 2021]. https://www.peta.org/blog/many-animals-saved/

Salomon, E., Preston, J. L., & Tannenbaum, M. B. (2017). Climate change helplessness and the (de)moralization of individual energy behavior. *Journal of Experimental Psychology: Applied*, *23*, 15–28. https://doi.org/10.1037/xap0000105

Sanchez-Sabate, R., & Sabaté, J. (2019). Consumer Attitudes Towards Environmental Concerns of Meat Consumption: A Systematic Review. *International Journal of Environmental Research and Public Health*, *16*(7). https://doi.org/10.3390/ijerph16071220

Scarborough, P., Appleby, P. N., Mizdrak, A., Briggs, A. D. M., Travis, R. C., Bradbury, K. E., & Key, T. J. (2014). Dietary greenhouse gas emissions of meat-eaters, fish-eaters, vegetarians and vegans in the UK. *Climatic Change*, *125*(2), 179–192. https://doi.org/10.1007/s10584-014-1169-1

Schiermeier, Q. (2019). Eat less meat: UN climate-change report calls for change to human diet. *Nature*, *572*(7769), 291–292. https://doi.org/10.1038/d41586-019-02409-7

Schultz, P. W. (1999). Changing Behavior With Normative Feedback Interventions: A Field Experiment on Curbside Recycling. *Basic and Applied Social Psychology*, *21*(1), 25–36. https://doi.org/10.1207/s15324834basp2101_3

Shiller, R. J. (2017). Narrative Economics. *The American Economic Review*, *107*(4), 967–1004. http://www.jstor.org/stable/44251584

Shoup, M. E. (2018). Innovative plant-based food options outperform traditional staples, nielsen finds [Accessed May 6, 2021]. https://www.foodnavigator-usa.com/Article/2018/06/20/Innovative-plant-based-food-options-outperform-traditional-staples-Nielsen-finds

Slater, M. D., Buller, D. B., Waters, E., Archibeque, M., & LeBlanc, M. (2003). A test of conversational and testimonial messages versus didactic presenta-

tions of nutrition information. *Journal of nutrition education and behavior*, *35*(5), 255–259. https://doi.org/10.1016/s1499-4046(06)60056-0

Sparkman, G., & Walton, G. M. (2017). Dynamic norms promote sustainable behavior, even if it is counternormative [PMID: 28961062]. *Psychological Science*, *28*(11), 1663–1674. https://doi.org/10.1177/0956797617719950

Stedman, R. C. (2004). Risk and climate change: perceptions of key policy actors in Canada. *Risk Analysis: An International Journal*, *24*(5), 1395–1406.

Stehfest, E., Bouwman, L., Van Vuuren, D. P., Den Elzen, M. G., Eickhout, B., & Kabat, P. (2009). Climate benefits of changing diet. *Climatic Change*, *95*(1-2), 83–102. https://doi.org/10.1007/s10584-008-9534-6

Tian, Q., Hilton, D., & Becker, M. (2016). Confronting the meat paradox in different cultural contexts: Reactions among Chinese and French participants. *Appetite*, *96*, 187–194. https://doi.org/10.1016/j.appet.2015.09.009

Tilman, D., & Clark, M. (2015). Food, agriculture & the environment: Can we feed the world & save the earth? *Daedalus*, *144*(4), 8–23. https://doi.org/10.1162/DAED_a_00350

Tobler, C., Visschers, V. H., & Siegrist, M. (2011). Eating green. consumers' willingness to adopt ecological food consumption behaviors. *Appetite*, *57*(3), 674–682. https://doi.org/10.1016/j.appet.2011.08.010

Truelove, H. B., & Parks, C. (2012). Perceptions of behaviors that cause and mitigate global warming and intentions to perform these behaviors. *Journal of Environmental Psychology*, *32*(3), 246–259. https://doi.org/10.1016/j.jenvp.2012.04.002

Vainio, A., Irz, X., & Hartikainen, H. (2018). How effective are messages and their characteristics in changing behavioural intentions to substitute plant-based foods for red meat? The mediating role of prior beliefs. *Appetite*, *125*, 217–224. https://doi.org/10.1016/j.appet.2018.02.002

Vanhonacker, F., Loo, E. J. V., Gellynck, X., & Verbeke, W. (2013). Flemish consumer attitudes towards more sustainable food choices. *Appetite*, *62*, 7–16. https://doi.org/10.1016/j.appet.2012.11.003

Verain, M., Sijtsema, S., Dagevos, H., & Antonides, G. (2017). Attribute Segmentation and Communication Effects on Healthy and Sustainable Consumer Diet Intentions. *Sustainability*, *9*(5), 743. https://doi.org/10.3390/su9050743

Washburn, A. N., & Skitka, L. J. (2018). Science denial across the political divide: Liberals and conservatives are similarly motivated to deny attitude-inconsistent science. *Social Psychological and Personality Science*, *9*(8), 972–980. https://doi.org/10.1177/1948550617731500

Weber, C. L., & Matthews, H. S. (2008). Food-Miles and the Relative Climate Impacts of Food Choices in the United States. *Environmental Science & Technology*, *42*(10), 3508–3513. https://doi.org/10.1021/es702969f

Weeks, B. E., Ardèvol-Abreu, A., & Gil de Zúñiga, H. (2015). Online Influence? Social Media Use, Opinion Leadership, and Political Persuasion. *International Journal of Public Opinion Research*, *29*(2), 214–239. https://doi.org/10.1093/ijpor/edv050

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, *59*(4), 329–349. https://doi.org/10.1080/03637759209376276

Wolske, K. S., Gillingham, K. T., & Schultz, P. W. (2020). Peer influence on household energy behaviours. *Nature Energy*, *5*(3), 202–212. https://doi.org/10.1038/s41560-019-0541-9

Xue, W., Hine, D., Marks, A., Phillips, W., Nunn, P., & Shouying, Z. (2016). Combining threat and efficacy messaging to increase public engagement with climate change in Beijing, China. *Climatic Change*, *137*. https://doi.org/10.1007/s10584-016-1678-1

Appendix

2.A Additional Tables and Figures

Table 2.A.1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Change in meat intake (days/week)	0.597	1.052	-5	7	1220
Donation to charity	0.305	0.461	0	1	1220
Prior Ecological Concern	4.025	0.734	1	5	1220
Pre-intervention diet	4.239	2.05	0	7	1220
Post-intervention diet	3.642	2.147	0	7	1220
Baseline Efficacy Beliefs	3.688	1.024	1	5	1220
Post-intervention Efficacy Beliefs	3.788	1.033	1	5	1220
Disease (history of family diseases)	0.594	0.491	0	1	1220
Age	36.329	10.394	19	78	1220
Female	0.388	0.487	0	1	1220
Democrat	0.555	0.497	0	1	1220
Education	0.774	0.419	0	1	1220
Moral offence taken	3.975	0.84	1	5	1220

The dummy variable 'female' included in the analysis in the paper takes value 1 if the subject listed their gender as female, and 0 otherwise. The 747 subjects for whom the female dummy variable takes value 0 includes 743 males, 3 subjects who chose 'Other' as their gender and 1 subject who chose 'Prefer not to say' when asked about their gender. Similarly, the dummy variable 'democrat' takes value 1 if the subject answered 'Democrat' when asked 'Which party would you prefer to win the next election?', and 0 otherwise. The 543 for whom the democrat dummy variable takes value 0 includes 374 subjects who answered 'Republican' when asked the same question, 104 subjects who chose 'Other' and 65 subjects who answered 'Prefer not to say'.

Table 2.A.2: ANOVA tests to compare means across groups

Variable	Control	More Scientific Information	Efficacy Salience	Health Information	Animal Welfare	Social Norms	Social Efficacy	F-statistic	p-value
Prior Ecological Concern	3.9693 (0.7508)	4.0449 (0.7588)	4.0607 (0.6985)	4.0353 (0.6904)	3.9960 (0.7736)	3.9843 (0.7837)	4.0854 (0.6824)	0.59	0.7366
Pre-intervention Diet	4.4972 (1.9832)	4.0800 (2.0579)	4.2458 (2.0679)	4.3448 (1.9845)	4.0565 (2.2098)	4.0793 (2.0544)	4.3563 (1.9737)	1.22	0.2948
Baseline Efficacy	3.6949 (1.0793)	3.8029 (0.9926)	3.7095 (1.0039)	3.6724 (0.9687)	3.6271 (0.9516)	3.5640 (1.1298)	3.7385 (1.0393)	0.96	0.4489
Age	35.4859 (8.9450)	37.4114 (11.0984)	36.4525 (10.1761)	34.6379 (10.0462)	36.3220 (10.4399)	37.2073 (10.7626)	36.8391 (11.0740)	1.55	0.1580
Female	0.3503 (0.4784)	0.4400 (0.4978)	0.3352 (0.4734)	0.3736 (0.4851)	0.3672 (0.4834)	0.3720 (0.4848)	0.4770 (0.5009)	1.94	0.0709*
Democrat	0.6045 (0.4903)	0.5714 (0.4963)	0.5475 (0.4991)	0.5920 (0.4929)	0.5085 (0.5013)	0.4695 (0.5006)	0.5862 (0.4939)	1.68	0.1229
Education	0.7345 (0.4429)	0.8229 (0.3829)	0.8212 (0.3842)	0.7414 (0.4391)	0.7571 (0.4301)	0.7988 (0.4021)	0.7414 (0.4391)	1.54	0.1613

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.3: Impact of information interventions on proportionate planned reduction in number of days of meat consumption per week

	Dependent Variable: $\frac{\Delta Days}{Pre-intervention\ Diet}$		
	(1)	(2)	(3)
More scientific information	0.0544*	0.0460	0.0411
	(0.0283)	(0.0281)	(0.0283)
Efficacy salience	0.0655**	0.0612**	0.0587**
	(0.0282)	(0.0274)	(0.0274)
Health information	0.0044	0.0028	0.0021
	(0.0295)	(0.0291)	(0.0289)
Animal welfare	-0.0169	-0.0173	-0.0177
	(0.0306)	(0.0300)	(0.0302)
Social norms	-0.0073	-0.0048	-0.0067
	(0.0282)	(0.0276)	(0.0276)
Social efficacy	-0.0365	-0.0419	-0.0446
	(0.0279)	(0.0280)	(0.0286)
Prior Ecological Concern		0.0187**	0.0177**
		(0.0074)	(0.0078)
Pre-intervention Diet		-0.0111	-0.0088
		(0.0093)	(0.0092)
Baseline Efficacy Beliefs		0.0403***	0.0424***
		(0.0087)	(0.0089)
Age			-0.0020
			(0.0080)
Female			0.0229
			(0.0182)
Democrat			0.0004
			(0.0188)
Education			0.0395**
			(0.0154)
N	1220	1220	1220

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.4: Impact of information interventions on planned reduction in number of days of meat consumption per week for those with and without prior history of diseases in the family

	Dependent Variable: Δ Days			
	Prior history		No history	
More scientific information	0.3061**	0.3498***	0.0529	-0.0048
	(0.1195)	(0.1170)	(0.2079)	(0.2020)
Efficacy salience	0.2446*	0.2639**	0.0673	0.0382
	(0.1300)	(0.1275)	(0.1781)	(0.1724)
Health information	0.3174**	0.3423**	-0.3511*	-0.3666*
	(0.1365)	(0.1337)	(0.1960)	(0.1938)
Animal welfare	0.0653	0.1242	-0.2015	-0.1793
	(0.1255)	(0.1240)	(0.1778)	(0.1741)
Social norms	-0.0093	0.0865	-0.1894	-0.1889
	(0.1306)	(0.1330)	(0.1642)	(0.1576)
Social efficacy	0.1108	0.0925	-0.4464**	-0.4825**
	(0.1061)	(0.1109)	(0.1907)	(0.1907)
Prior Ecological Concern		0.0613		0.1294***
		(0.0384)		(0.0486)
Pre-intervention Diet		0.2362***		0.1944***
		(0.0349)		(0.0489)
Baseline Efficacy Beliefs		0.1608***		0.1468***
		(0.0418)		(0.0528)
Age		-0.0064		0.0107
		(0.0317)		(0.0528)
Female		0.0931		0.1790
		(0.0777)		(0.1099)
Democrat		0.1486*		-0.1234
		(0.0783)		(0.1013)
Education		-0.0170		0.2273*
		(0.0797)		(0.1180)
N	725	725	495	495

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.5: Impact of information interventions on planned reduction in number of days of meat consumption per week depending on baseline efficacy beliefs

	Dependent Variable: Δ Days		
	(1)	(2)	(3)
More scientific information × High efficacy beliefs	0.4410**	0.3883*	0.3936*
	(0.2122)	(0.2084)	(0.2089)
Efficacy salience × High efficacy beliefs	0.6858***	0.6268***	0.6309***
	(0.2036)	(0.2027)	(0.2031)
Health information × High efficacy beliefs	0.3549	0.2841	0.2821
	(0.2263)	(0.2236)	(0.2240)
Animal welfare × High efficacy beliefs	0.3920*	0.3559*	0.3475*
	(0.2041)	(0.2024)	(0.2027)
Social norms × High efficacy beliefs	0.4500**	0.4148**	0.3959**
	(0.1956)	(0.1934)	(0.1936)
Social efficacy × High efficacy beliefs	0.1265	0.0359	0.0381
	(0.2005)	(0.1999)	(0.2004)
More scientific information	-0.0329	0.0238	0.0035
	(0.1626)	(0.1611)	(0.1626)
Efficacy salience	-0.1525	-0.1149	-0.1204
	(0.1458)	(0.1467)	(0.1473)
Health information	-0.1150	-0.0691	-0.0702
	(0.1703)	(0.1697)	(0.1711)
Animal welfare	-0.2146	-0.1568	-0.1537
	(0.1557)	(0.1540)	(0.1534)
Social norms	-0.2639*	-0.2017	-0.1974
	(0.1517)	(0.1516)	(0.1520)
Social efficacy	-0.1584	-0.1083	-0.1256
	(0.1634)	(0.1633)	(0.1646)
High Efficacy Beliefs	-0.0344	-0.2456	-0.2535
	(0.1299)	(0.1680)	(0.1684)
Prior Ecological Concern		0.1042***	0.1010***
		(0.0304)	(0.0309)
Pre-intervention Diet		0.2091***	0.2170***
		(0.0285)	(0.0287)
Baseline Efficacy Beliefs		0.1307**	0.1344**
		(0.0583)	(0.0582)
Age			0.0033
			(0.0273)
Female			0.1308**
			(0.0643)
Democrat			0.0170
			(0.0626)
Education			0.0824
			(0.0663)
N	1220	1220	1220

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

High efficacy beliefs is a dummy variable which equals 1 if the subject's baseline efficacy beliefs is greater than or equal to the median value and 0 otherwise.

Table 2.A.6: Impact of information interventions on planned reduction in number of days of meat consumption per week for those with high and low baseline efficacy beliefs

	Dependent Variable: Δ Days			
	High efficacy		Low efficacy	
More scientific information	0.4081***	0.4039***	-0.0329	-0.0117
	(0.1362)	(0.1334)	(0.1627)	(0.1663)
Efficacy salience	0.5333***	0.5128***	-0.1525	-0.1401
	(0.1421)	(0.1419)	(0.1458)	(0.1484)
Health information	0.2398	0.2123	-0.1150	-0.0850
	(0.1489)	(0.1435)	(0.1704)	(0.1738)
Animal welfare	0.1774	0.1981	-0.2146	-0.1764
	(0.1318)	(0.1354)	(0.1558)	(0.1552)
Social norms	0.1860	0.2048*	-0.2639*	-0.2200
	(0.1235)	(0.1223)	(0.1517)	(0.1531)
Social efficacy	-0.0319	-0.0898	-0.1584	-0.1381
	(0.1162)	(0.1182)	(0.1634)	(0.1669)
Prior Ecological Concern		0.0765		0.1099***
		(0.0491)		(0.0401)
Pre-intervention Diet		0.2532***		0.1695***
		(0.0367)		(0.0467)
Baseline Efficacy Beliefs		0.2169**		0.0937
		(0.1054)		(0.0704)
Age		-0.0006		0.0105
		(0.0373)		(0.0416)
Female		0.1596*		0.0987
		(0.0834)		(0.1000)
Democrat		0.0299		0.0020
		(0.0841)		(0.0935)
Education		0.0491		0.1216
		(0.0858)		(0.1042)
N	640	640	580	580

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.7: Impact of information interventions on planned reduction in number of days of meat consumption per week depending on baseline frequency of meat consumption

	Dependent Variable: Δ Days		
	(1)	(2)	(3)
More scientific information × Frequent meat eater	0.3458*	0.3180	0.3050
	(0.2002)	(0.1964)	(0.1966)
Efficacy salience × Frequent meat eater	0.2563	0.1968	0.1842
	(0.1940)	(0.1896)	(0.1903)
Health information × Frequent meat eater	0.3057	0.2172	0.2080
	(0.2138)	(0.2107)	(0.2130)
Animal welfare × Frequent meat eater	0.3955**	0.3567*	0.3545*
	(0.1951)	(0.1952)	(0.1967)
Social norms × Frequent meat eater	0.3199*	0.3232*	0.3135*
	(0.1887)	(0.1875)	(0.1873)
Social efficacy × Frequent meat eater	0.1463	0.0574	0.0752
	(0.1977)	(0.1981)	(0.1990)
More scientific information	0.0233	0.0250	0.0173
	(0.1315)	(0.1300)	(0.1299)
Efficacy salience	0.0407	0.0730	0.0793
	(0.1322)	(0.1302)	(0.1295)
Health information	-0.1239	-0.0685	-0.0649
	(0.1521)	(0.1496)	(0.1507)
Animal welfare	-0.2511*	-0.2052	-0.2034
	(0.1375)	(0.1374)	(0.1377)
Social norms	-0.2259*	-0.1989	-0.1961
	(0.1313)	(0.1307)	(0.1312)
Social efficacy	-0.1791	-0.1432	-0.1699
	(0.1551)	(0.1553)	(0.1571)
Frequent meat eater	0.0820	-0.1349	-0.1441
	(0.1253)	(0.1611)	(0.1610)
Prior Ecological Concern		0.1048***	0.1013***
		(0.0301)	(0.0306)
Pre-intervention Diet		0.1789***	0.1919***
		(0.0536)	(0.0542)
Baseline Efficacy Beliefs		0.1526***	0.1510***
		(0.0324)	(0.0331)
Age			-0.0004
			(0.0275)
Female			0.1276**
			(0.0648)
Democrat			0.0256
			(0.0628)
Education			0.0780
			(0.0668)
N	1220	1220	1220

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Frequent meat eater is a dummy variable which equals 1 if subject reports that their baseline number of days of weekly meat consumption is greater than or equal to the median value (4 days) and 0 otherwise.

Table 2.A.8: Impact of information interventions on planned reduction in number of days of meat consumption per week for frequent and infrequent meat eaters

	Dependent Variable: Δ Days			
	Frequent		Infrequent	
More scientific information	0.3692**	0.3038**	0.0233	0.0293
	(0.1507)	(0.1471)	(0.1317)	(0.1240)
Efficacy salience	0.2970**	0.2567*	0.0407	0.0779
	(0.1418)	(0.1393)	(0.1325)	(0.1261)
Health information	0.1818	0.1355	-0.1239	-0.0775
	(0.1500)	(0.1477)	(0.1524)	(0.1450)
Animal welfare	0.1444	0.1440	-0.2511*	-0.1736
	(0.1383)	(0.1385)	(0.1378)	(0.1358)
Social norms	0.0941	0.1125	-0.2259*	-0.1935
	(0.1354)	(0.1324)	(0.1316)	(0.1276)
Social efficacy	-0.0328	-0.1157	-0.1791	-0.1566
	(0.1223)	(0.1236)	(0.1555)	(0.1552)
Prior Ecological Concern		0.1093***		0.0602
		(0.0384)		(0.0456)
Pre-intervention Diet		0.0789		0.4054***
		(0.0719)		(0.0523)
Baseline Efficacy Beliefs		0.1721***		0.1310***
		(0.0465)		(0.0399)
Age		0.0332		-0.0240
		(0.0422)		(0.0300)
Female		0.2206**		-0.0019
		(0.0889)		(0.0822)
Democrat		0.0186		0.0443
		(0.0890)		(0.0813)
Education		0.0254		0.1407**
		(0.0905)		(0.0700)
N	771	771	449	449

Standard errors in parentheses. Statistical significance indicated as follows:
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2.A.9: Impact of information interventions on decision to donate bonus earnings

	Dependent Variable: Prob(donate=1)			
	(1) All	(2) All	(3) Low Priors	(4) High Priors
More scientific information	0.0600	0.0335	0.1415**	-0.0471
	(0.0480)	(0.0455)	(0.0594)	(0.0652)
Efficacy salience	0.0586	0.0221	0.0727	-0.0227
	(0.0477)	(0.0446)	(0.0553)	(0.0643)
Health information	0.0044	-0.0149	0.0313	-0.0360
	(0.0466)	(0.0438)	(0.0548)	(0.0637)
Animal welfare	0.0621	0.0203	0.0456	0.0190
	(0.0479)	(0.0444)	(0.0562)	(0.0659)
Social norms	0.1116**	0.0708	0.1489**	0.0179
	(0.0499)	(0.0467)	(0.0632)	(0.0655)
Social efficacy	0.0619	0.0589	0.1101*	0.0362
	(0.0481)	(0.0466)	(0.0626)	(0.0648)
Prior Ecological Concern		0.1166***	0.1363***	-0.0338
		(0.0142)	(0.0276)	(0.0482)
Pre-intervention Diet		-0.0809***	-0.0857***	-0.0651***
		(0.0121)	(0.0164)	(0.0170)
Δ Days		0.0250*	0.0216	0.0225
		(0.0129)	(0.0178)	(0.0173)
Baseline Efficacy Beliefs		-0.1078***	-0.0541***	-0.1433***
		(0.0132)	(0.0178)	(0.0183)
Age		-0.0030	-0.0123	0.0311*
		(0.0128)	(0.0174)	(0.0184)
Female		-0.0554**	-0.0342	-0.0670*
		(0.0257)	(0.0344)	(0.0355)
Democrat		-0.0439*	-0.0018	-0.0773**
		(0.0264)	(0.0341)	(0.0358)
Education		0.1236***	0.1700***	0.0692
		(0.0313)	(0.0447)	(0.0450)
N	1220	1220	545	675

Standard errors in parentheses. Statistical significance indicated as follows:

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The dependant variable is the binary variable donate, which equals 1 if the subject donated their bonus earnings to the climate change charity and 0 otherwise. The table reports the average marginal effects from Probit regressions. Columns 3 and 4 show the impact of the interventions on subjects with low and high prior ecological concern, respectively.

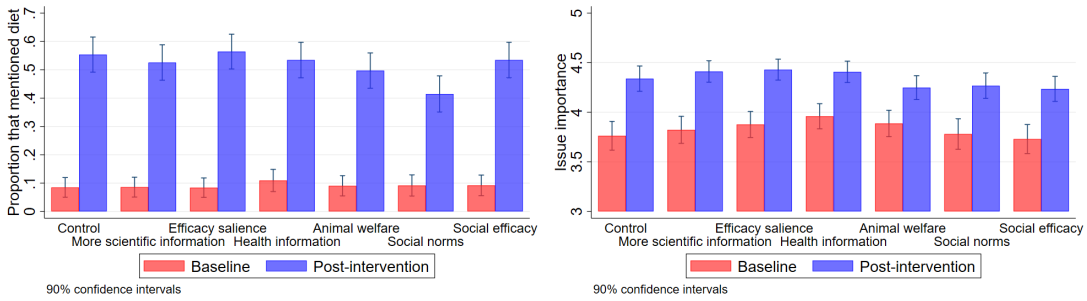
(a) Baseline (b) Control

(c) More scientific information (d) Efficacy salience

(e) Health information (f) Animal welfare

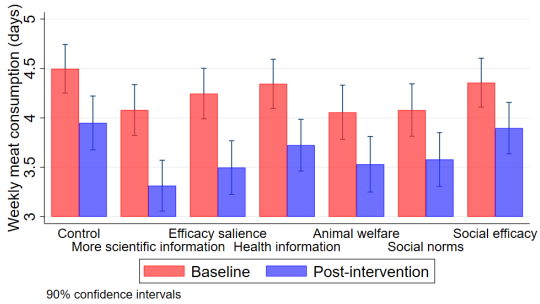
(g) Social norms (h) Social efficacy

Figure 2.A.1: Actions to combat climate change listed at baseline and post-intervention. (a) Wordcloud representing baseline answers to actions you can take personally to combat climate change. (b) to (h) Wordclouds representing post interventions answers to actions you can take personally to combat climate change across the different groups. Meat intake-related actions were identified by the use of words such as *vegetarian, meat, plant-based, vegan* and *beef*.

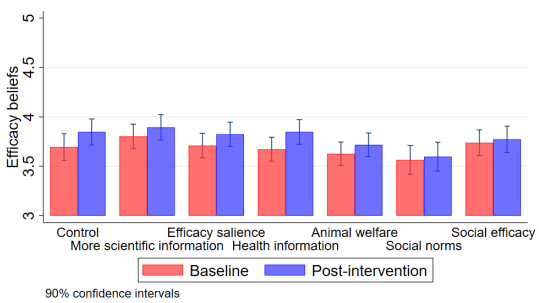


(a) Proportion mentioning diet

(b) Issue importance



(c) Weekly meat consumption (days)



(d) Efficacy beliefs

Figure 2.A.2: (a) Proportion of subjects that mentioned diet as an effective action to combat climate change (b) issue importance (c) weekly meat consumption (days) and (d) efficacy beliefs before and after intervention across different conditions.

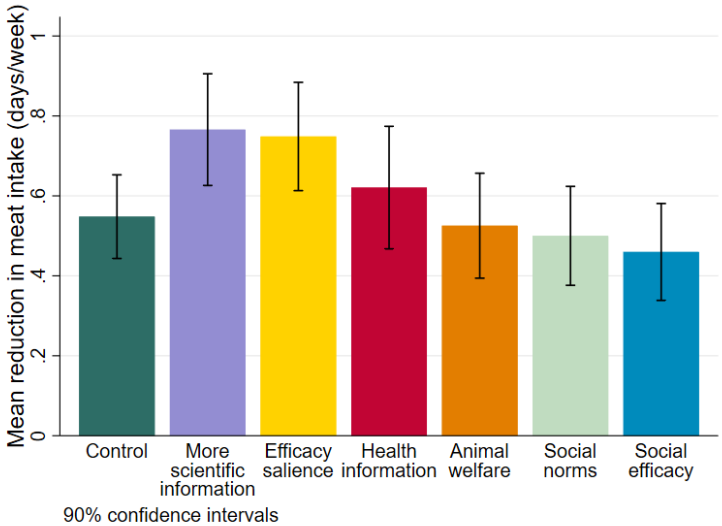
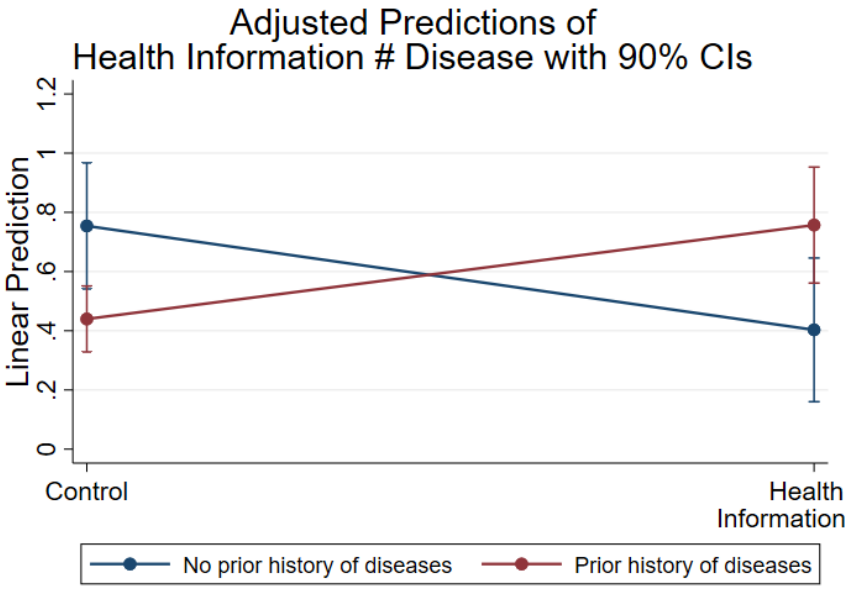
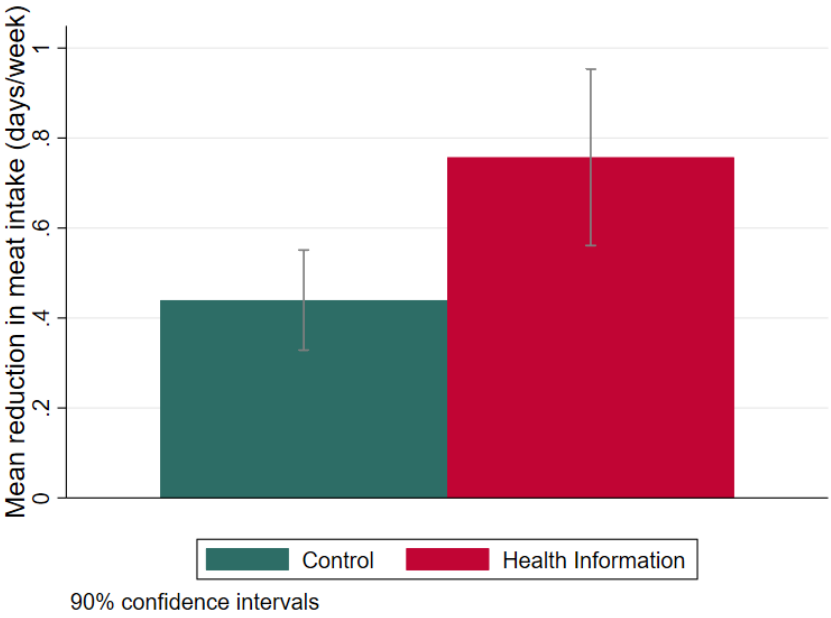


Figure 2.A.3: (a) Average planned reduction in number of days of meat consumption per week across groups. A oneway ANOVA test rejected the null of no significant difference in means between the groups with F-statistic 2.32 and $p = 0.0309$.

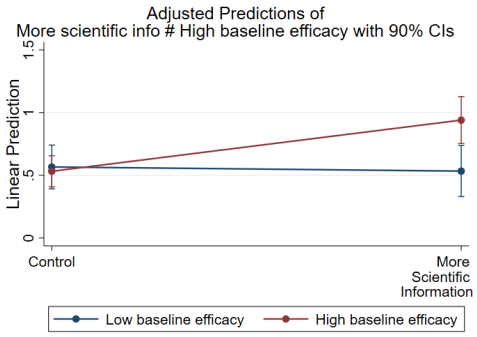


(a) Interaction plot

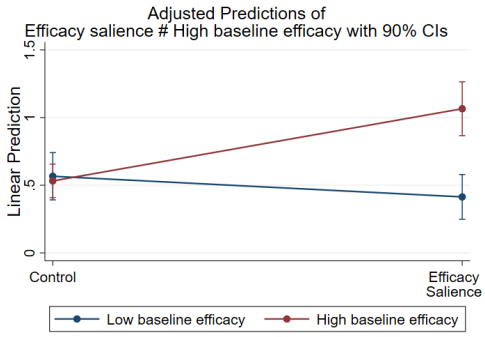


(b) Subjects with prior health concerns

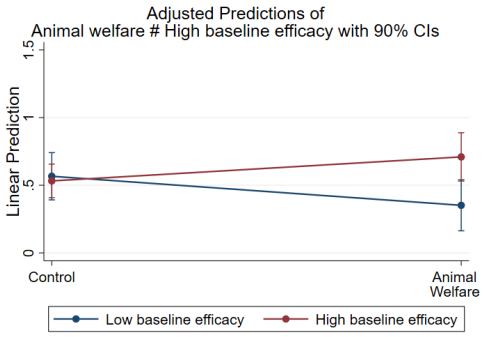
Figure 2.A.4: The differential effect of the health information intervention on intended change in weekly meat consumption depending on whether the subject has prior health concerns. (a) shows that health information is more effective than the Control when the individual has prior health concerns (disease=1). (b) plots the mean change in weekly meat consumption for control and health information groups, only when subjects have prior health concerns. The null of equal means was rejected by a t-test with t-statistic -2.3716 and $p = 0.0186$.



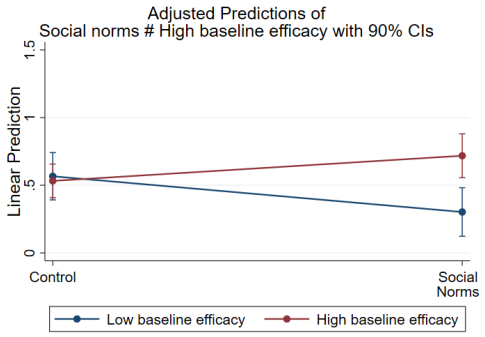
(a) More scientific information



(b) Efficacy salience

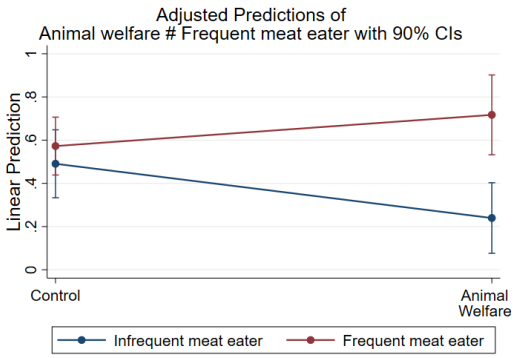


(c) Animal welfare

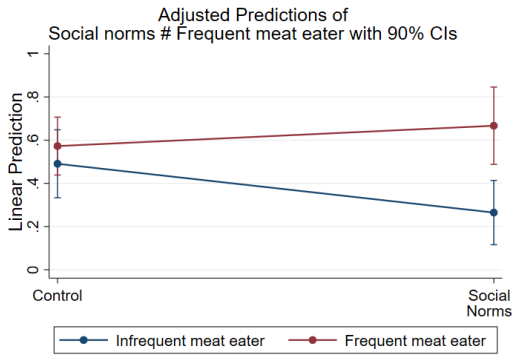


(d) Social norms

Figure 2.A.5: Interaction plots showing effects of (a) More scientific information, (b) Efficacy salience, (c) Animal welfare and (d) Social norms on intended reduction in number of days of meat consumption per week, for different baseline efficacy beliefs values. High efficacy beliefs is a dummy variable which equals 1 if the subject's baseline efficacy beliefs is greater than or equal to the median value and 0 otherwise.

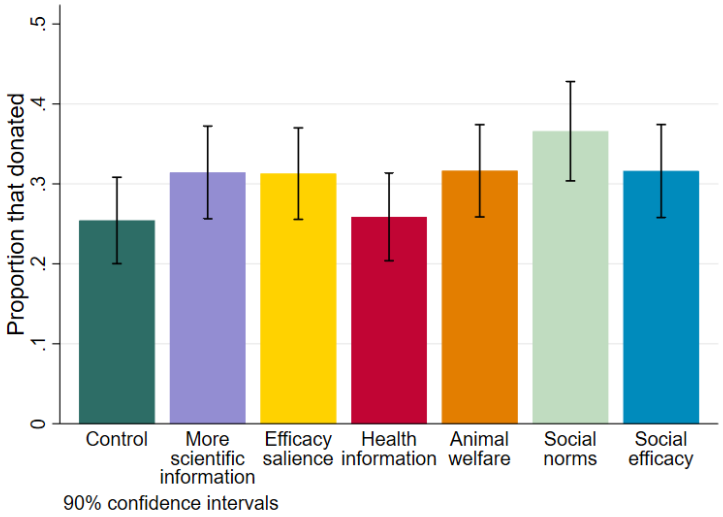


(a) Animal welfare

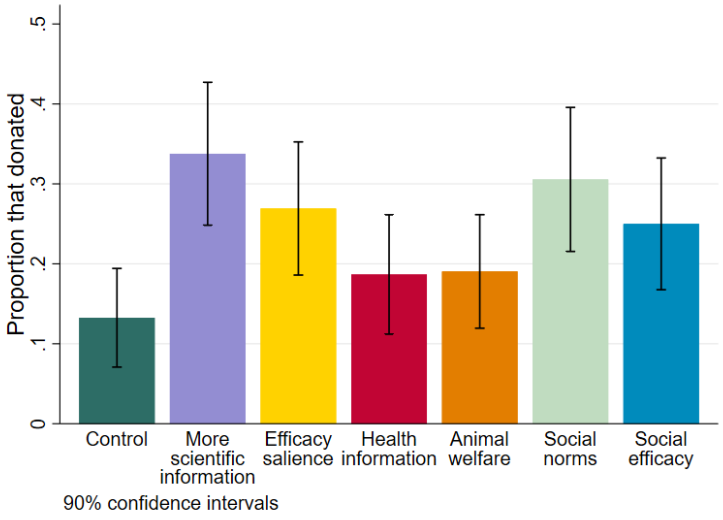


(b) Social norms

Figure 2.A.6: Interaction plot showing effects of (a) Animal welfare and (b) Social norms conditions on planned reduction in number of days of meat consumption per week, for different frequencies of baseline meat consumption. Frequent meat eater is a dummy variable which equals 1 if subject reports that their baseline number of days of weekly meat consumption is greater than or equal to the median value (4 days) and 0 otherwise.

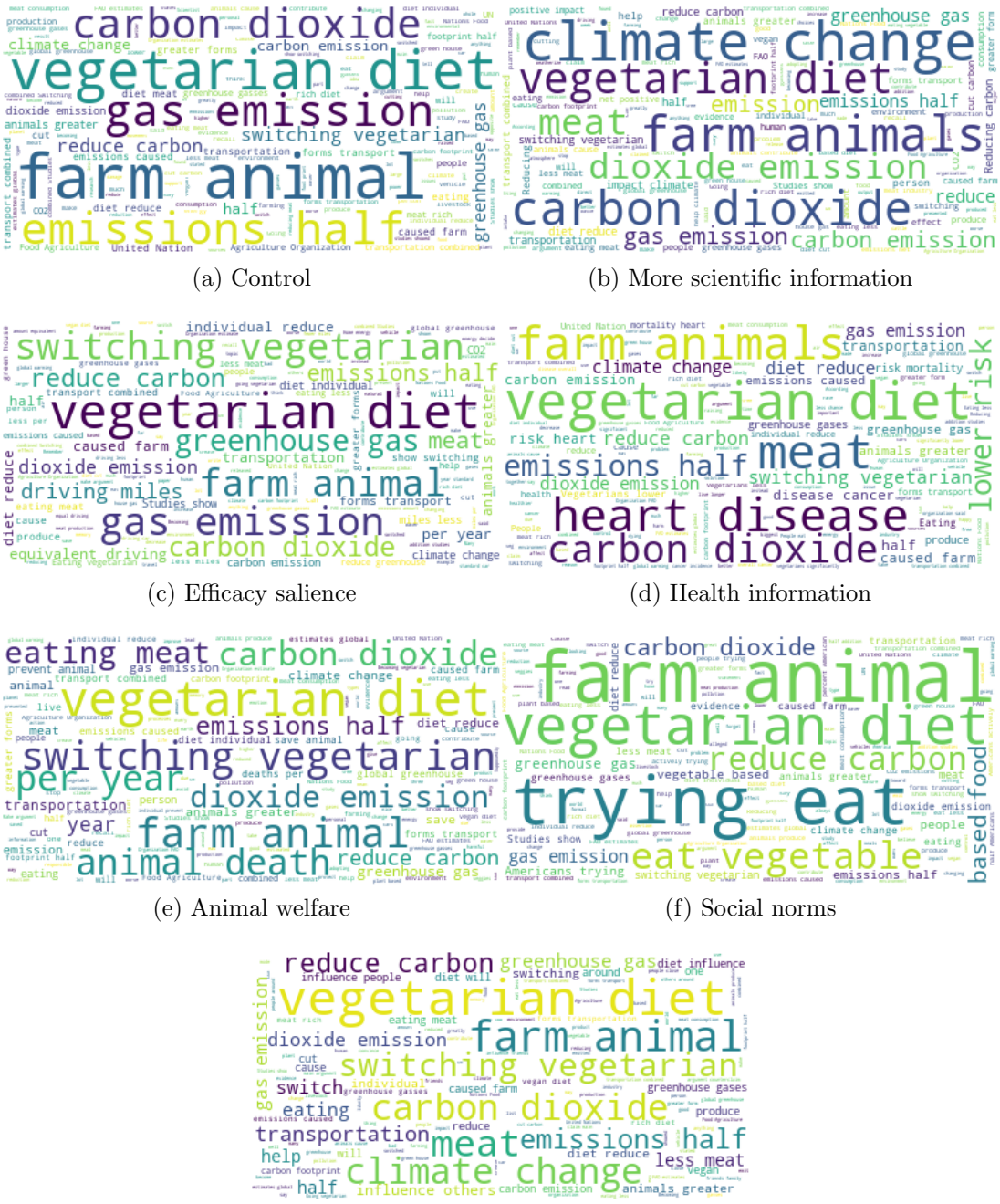


(a) All



(b) Low Priors

Figure 2.A.7: Donation to charity across groups. (a) presents the proportions that donated to charity in all groups. (b) presents the proportions that donated to charity when the subject had low prior ecological concern, across the different groups. A oneway ANOVA test rejected the null of no significant difference in means between the groups for those with low priors with F-statistic 2.33 and $p = 0.0316$.



(a) Control

(b) More scientific information

(c) Efficacy salience

(d) Health information

(e) Animal welfare

(f) Social norms

(g) Social efficacy

Figure 2.A.8: Evidence recall from different treatment conditions

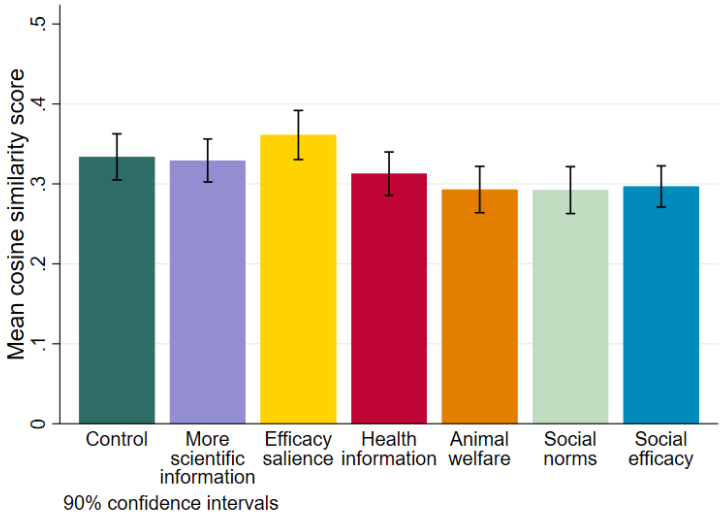


Figure 2.A.9: Cosine similarity scores between the evidence the subjects could recall and the actual evidence given to them across the different groups

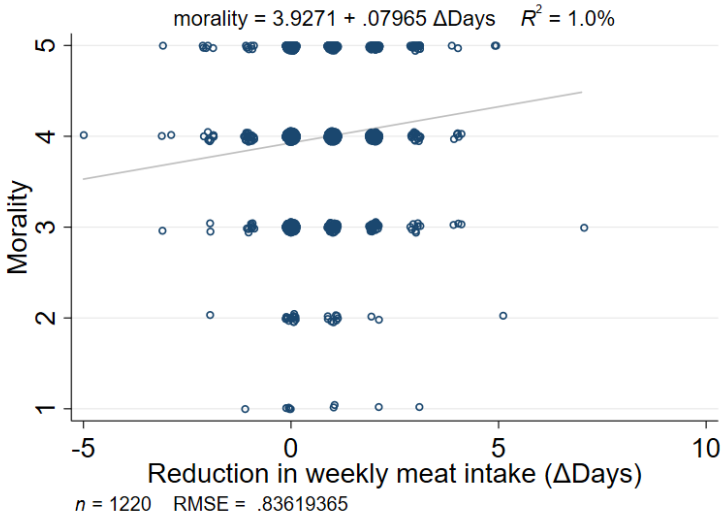


Figure 2.A.10: Relationship between Δ Days and answer to "Is it morally wrong to show people the consequences of their own behaviour?" on a scale of 1 to 5 where 1 is very morally wrong and 5 is very morally right.

2.B Information Interventions

The following are the messages which were used in the interventions. Sources of the information mentioned below were not included as part of the information interventions.

Control

The United Nations Food and Agriculture Organization (FAO) estimates global greenhouse gas emissions caused by farm animals is greater than all forms of transport combined.

Studies show that by switching to a vegetarian diet from a meat rich diet, an individual could reduce their carbon dioxide emissions by half.

Source: FAO, 2006; Tilman and Clark, 2015; Scarborough et al., 2014

More scientific information

Control
+
In addition, studies show that reducing carbon dioxide emissions can have a net positive impact on climate change.

Source: NASA, 2020

Efficacy salience

Control
+
In addition, studies show that by switching to a vegetarian diet an individual can reduce their greenhouse gas emissions by an amount that is equivalent to driving about 3000 miles less per year in a standard car.

Source: Weber and Matthews, 2008.

Health information

Control
+

In addition, studies show that vegetarians have a significantly lower risk of mortality from heart disease and overall cancer incidence.

Source: Huang et al., 2012

Animal welfare

Control
+
In addition, studies show that by switching to a vegetarian diet an individual can prevent more than 100 animal deaths per year.

Source: PETA, 2013

Social norms

Control
+
In addition, studies show that 39% of Americans are actively trying to eat more vegetable-based foods.

Source: Shoup, 2018

Social efficacy

Control
+
In addition, studies show that if you switch to a vegetarian diet from a meat rich diet you could influence people close to you to do the same.

Source: Higgs and Thomas, 2016

2.C Experiment Script

1. Rate the degree to which you agree with the following statements:

   Answer on a 5-point likert scale where 1 is strongly disagree and 5 is strongly agree.

   (a) The balance of nature is very delicate and easily upset by human activities.

(b) Ecological, rather than economic, factors must guide our use of natural resources.

(c) We attach too much importance to economic measures of the well-being of our society.

(d) We are approaching the limit of the number of people the earth can support.

(e) When humans interfere with nature, it often produces disastrous consequences.

(f) Humans must live in harmony with nature in order to survive.

(g) There are limits to growth beyond which our industrialized society cannot expand.

2. How concerned about climate change are you?

   Answer on a 5-point likert scale where 1 is Not at all concerned and 5 is Extremely concerned.

3. How concerned about climate change do your friends think you are

   Answer on a 5-point likert scale where 1 is Not at all concerned and 5 is Extremely concerned.

4. What actions could you take personally to reduce climate change? Please mention 3 actions. Leave empty if you cannot think of any.

5. Rate the degree to which you agree with the following statements:

   Answer on a 5-point likert scale where 1 is Strongly disagree and 5 is strongly agree.

   (a) Individuals can influence climate change.

   (b) Collectively humans have little influence on climate change.

6. How many days in a week do you eat meat? Please indicate a number between 0 to 7 days.

7. Rate the degree to which you agree with the following statement.

   There is a relationship between climate change and people's food choice.

   Answer on a 5-point likert scale where 1 is Strongly disagree and 5 is strongly agree.

Information Intervention

8. Rate the degree to which you agree with the following statement.

   The information I read made me feel that there is a relationship between climate change and people's food choices.

   Answer on a 5-point likert scale where 1 is Strongly disagree and 5 is strongly agree.

9. Rate the degree to which you agree with the following statements.

   Answer on a 5-point likert scale where 1 is Strongly disagree and 5 is strongly agree.

   (a) Individuals can influence climate change.

   (b) Collectively humans have little influence on climate change.

10. After reading the information provided, how many days in a week will you eat meat? Please indicate a number between 0 to 7 days.

11. What actions could you take personally to reduce climate change? Please mention 3 actions. Leave empty if you cannot think of any.

12. Please write down everything you can recall from the evidence presented to you.

13. Is it morally wrong to show people the consequences of their own behaviour?

    Options - very morally wrong, morally wrong, neither morally wrong nor right, morally right, very morally right

14. What is your age?

15. What is your gender?

    Options - Male, Female, Other, Prefer not to say

16. Please indicate the highest academic degree you have completed.

    Options - None, High/Secondary School, Vocational Training, Bachelor, Master, PhD

17. In which part of the US are you currently located?

18. Which party would you prefer to win the next election?

Democrat, Republican, Other, Prefer not to say.

19. How many people would notice if you changed your diet?

20. How many of these people do you think might be influenced to change their diets as well?

21. Have you or any member of your immediate family (father, mother, siblings, and grandparents) had or suffered from heart disease, stroke, cancer, diabetes, high blood pressure or high cholesterol?

    You have now reached the end of the study.

    For your participation today, you will be receiving $1.50.

    Additionally, for your effort, we would like to offer you a **bonus of $.50**.

    You can keep the bonus earnings for yourself or you can give it to the experimenters who will **donate on your behalf to the** Adaptation Fund.

    The Adaptation Fund was set up under the Kyoto protocol of **United Nations Framework Convention on Climate Change**.The fund finances projects and programmes which help vulnerable communities in developing countries adapt to climate change.

    Do you want to donate your bonus earnings to The Adaptation Fund?

    Thank you for participating in the study. We appreciate your help with our research.

    Your earnings today is $x.

    In case of donation to The Adaptation Fund;

    Also, we want to thank you for your donation to The Adaptation Fund.

2.D Carbon dioxide emissions

Weber and Matthews, 2008 estimate that by switching to a plant-based diet just 1 day per week from red meat and dairy, a household can reduce their greenhouse gas (GHG) emissions equivalent to driving 1160 miles less per year. The authors normalise the data to the unit of a household using US census data in 1997 which included around 267 million residents in 101 million households. Hence, the size

of each household is around $\frac{267}{101}$ i.e 2.64. Thus, by switching to a plant-based diet 1 day per week an individual can reduce their GHG emission equivalent to driving $\frac{1160}{2.64}$ or 439.39 miles per year less. Further, the authors estimates that driving 12,000 miles in a standard automobile (25 miles/gallon) produces around 4.4 t CO_2 per year. Which means, that by switching to a plant-based diet 1 day per week an individual can reduce their CO_2 emissions by $\frac{4400}{12000} \times 439.39$ or 161.11 kg CO_2 per year.

3 Attitude towards Immigrants: Evidence from U.S. Congressional Speeches[1]

Immigration and attitudes towards immigration have been key features in economic development and political debate for decades. It can be hard to disentangle true beliefs about immigrants even where we have seemingly strong evidence such as the voting records of politicians. This paper builds an "immigration corpus" consisting of 24,351 U.S. Congressional speeches relevant to immigration issues between 1990-2015. The corpus is used to form two distinct measures of attitude towards immigrants - one based on sentiment (or valence) and one based on the concreteness of language. The linguistic measures, particularly sentiment, show systematic variation over time and across states in a manner consistent with the history and experiences of immigrants in the U.S. The paper also computes a speaker specific measure of sentiment towards immigrants which is found to be a significant positive predictor of voting behaviour with respect to immigration related bills. Applying a Latent Dirichlet Allocation (LDA) topic modelling algorithm provides further insight into how different topics (such as border security or national security) have risen and fallen in importance over time in the face of key events such as 9/11.

3.1 Introduction

The number of international migrants worldwide has increased rapidly in recent years with the total number of migrants reaching 272 million in 2019 (according to the latest World Migration Report by the United Nations; International Organization for Migration, 2020). Political turmoil, conflict, poverty and limited educational or employment opportunities force millions of people to migrate from their countries. The largest number of international migrants were reported to be living in the U.S. in 2019, the total number reaching 51 million (United Nations, 2019). Immigrants contribute significantly to the economy of the host country in terms of filling key labour market gaps, contributing to tax revenues and social security (United Nations, 2017). Particularly in the U.S., immigrants comprise up to a third of the workforce in several industries including farming, fishing

[1]I am grateful to Daniel Sgroi and Thomas Hills for their invaluable feedback and guidance. I thank Giovanni Facchini for his generosity with time and for sharing the voting data used in the paper. I am also grateful to Ying Li, Elliott Ash, Arianna Ornaghi, Clement Imbert, Sharun Mukand and seminar participants at Warwick University for their helpful comments. I declare no conflict of interest.

and forestry, textile manufacturing, food manufacturing, and home health care (Sherman et al., 2019). Immigrants also help the ageing U.S. population, which is crucial for improving the ratio of workers to retirees as 78% of the foreign-born population are of working age compared to 59% of native born; immigrants contribute significantly to the U.S. housing market and immigrants' children tend to gain more education, have higher salaries and work in higher-paying professions compared to their parents (Sherman et al., 2019). Yet, immigrants, as one of the most vulnerable sections of the society, are subject to constant human rights violation and discrimination.

In the U.S., for example, the members of the Congress play a massive role in terms of shaping legislation related to immigrant issues. Congress members vote on bills which either liberalise or restrict immigration. Variables such as party affiliation, regional ideologies and differences (Fetzer, 2006; Goldin, 1994), and key economic and socio-demographic variables (Facchini and Steinhardt, 2011; Milner and Tingley, 2011; Gonzalez and Kamdar, 2000) have been helpful in explaining voting behaviour of Congress legislators on immigration bills. This paper looks at another measure which helps shed light on variation in immigrant outcomes and voting behaviour of legislators on immigration related bills - Congressional speeches. The Congressional speeches record all speeches verbatim delivered by legislators on the floor of the Congress.

This paper uses the United States Congressional Records from 1990 to 2015 which is a record of all speeches delivered on the floor of Congress. From the Congressional records, an "immigration corpus" is formed by identifying speeches which contain *immigration*, *refugees* and related words. This yields a corpus consisting of 24,351 Congressional speeches with 43.6 million total number of words. The corpus comprises of speeches from 1098 unique speakers.

This immigration corpus is used to form two distinct linguistic measures of attitude towards immigrants - *sentiment* or valence and language *concreteness*. Sentiment measure of each immigration speech is computed using the valence ratings proposed by Warriner et al., 2013. Valence of a word refers to the pleasant emotion conveyed by a word, with the rating increasing as it moves from unhappy to happy. Concreteness[2] value of immigration speeches are calculated as a proxy for social distance. In this paper, concreteness measure of each speech is computed using the ratings proposed by Brysbaert et al., 2014. Self-categorisation theory posits that humans are conditioned to view different social groups as

[2]Concreteness can be defined as a word's ability to make specific and definite references to particular objects (Hills et al., 2017).

either in-group (within close social proximity) or out-group (socially distant) (Turner, 1985). Concrete linguistic patterns reflect positive out-group behaviour or lesser perceived social distance as supported by *Linguistic Intergroup Bias model* (Maass et al., 1989) and *Construal Level Theory* (Trope and Liberman, 2010). Use of concrete language declines when one shifts from describing family and friends to neighbours to colleagues to visitors and finally to foreigners (Snefjella and Kuperman, 2015). Thus, concreteness of language used can act as an indicator of intergroup bias which helps create and sustain stereotypes about social groups (Mastro et al., 2014).

The approach in this paper is inspired from Li and Hills, 2020 who form an immigration corpus from a collection of New York Times articles to quantify and compare sentiment and social distance towards different immigrant groups, as expressed in the articles. The present work, however, builds linguistic measures of attitude towards immigrants to study variation in immigrant outcomes over time and predict voting behaviour of politicians on immigration-related bills.

The sentiment and concreteness measures of immigration-related speeches delivered in the U.S. Congress were used to examine variation of attitude towards immigrants of the Congress members over time and by political affiliation and region. It was observed that, the variation of linguistic measures, particularly sentiment, is consistent with the history of immigrant outcomes, with a marked decline following the September 2001 terrorist attacks in the U.S. and the subsequent introduction of several anti-immigration bills. The sentiment measure displayed strict partisan polarisation consistently since early 1990s. There was also a clear regional distinction with conservative southern U.S. states displaying lower pro-immigrant sentiment compared to the north-eastern states.

Analysis of voting behaviour of Congressional members on seminal immigration bills in the decades before and after the 9/11 attacks was also undertaken in the paper. It was observed that speaker sentiment (as computed from speeches delivered on matters related to immigration policy) is a significant positive predictor of speaker's pro-immigration voting pattern even after controlling for a list of speaker characteristics, party affiliation and district level socio-economic variables. Depending on the econometric specification, a 1 standard deviation increase in speaker speech sentiment can increase the probability of a pro-immigration vote by 2 to 5 percentage points. Speaker speech sentiment is an especially strong predictor of voting behaviour in the southern, south-western and western U.S. states.

Lastly, a Latent Dirichlet Allocation or LDA topic modelling algorithm (Blei et

al., 2003) was used to identify the different immigration-related topics discussed in the U.S. Congress. This helps recognise the key topics for immigration legislation and examine the distribution of the different topics over time. The distribution of topics shows a hike in concern towards border security and national security following the 9/11 attacks and a decline in discussions related to educational policies and government aid in speeches addressing immigration. The distribution of topics over time expresses the variation in concern over different immigrant issues and is generally reflected by the nature of the bills introduced in Congress during the time. Applying LDA is an alternative, time saving and efficient way of examining the historical trends of immigration legislation compared to extensive manual studying of Congressional bills.

The present work adds to the small branch of literature which looks at linguistic bias towards immigrants, particularly in news coverage (Li and Hills, 2020; Mastro et al., 2014). This paper expands the literature by evaluating attitude or bias towards immigrants in political speeches. More broadly, the present study contributes to literature deriving measures of implicit bias from language use and how that might affect decision making. Related to this is the work by Rice et al., 2019 which looks at implicit racial bias in written judicial opinions in U.S. state and federal courts, and Ash et al., 2021 who build a similar measure for gender attitudes of U.S. circuit court judges.

This study also adds to the existing branch of literature which examines roll call voting patterns of U.S. Congress members on immigration bills (Facchini and Steinhardt, 2011; Milner and Tingley, 2011; Fetzer, 2006; Gonzalez and Kamdar, 2000). While the existing studies have relied on economic, demographic and political determinants of voting behaviour in the U.S. Congress, this paper uses a language-based measure of inherent attitude towards immigrants to predict pro-immigration voting behaviour. Lastly, this paper contributes to the modest literature which has focussed on analysing U.S. Congressional speeches using Natural Language Processing tools to build measures of partisanship (Gentzkow et al., 2019; Jensen et al., 2012).

Congressional speeches delivered on immigration issues are crucial for the following reasons: (1) They are better able to capture the *intensity* of the speaker's attitude towards immigration since it yields a continuous measure. This is superior to examining a speaker's attitude towards immigration based on past voting behaviour which gives a binary measure since a vote is either in favour of or against liberalising immigration. (2) The immigration corpus, due to its sheer volume, is expected to be more informative of the speaker's attitude towards im-

migrants compared to other measures, such as past voting behaviour. (3) In the absence of past voting behaviour, for example in case the bill does not go to a vote, Congressional speeches can be a useful measure for estimating the speaker's attitude towards the particular issue under debate.

Moreover, the study of political speeches, not only during Congressional sessions, but also during campaign trails and other public appearances, is crucial to form measures of the speaker's attitude towards immigration or other relevant issues. The benefit of the method proposed in this paper is that it can be applied to study politicians' attitude on other critical issues such as gender, race, climate change, etc. Developing a language-based measure of politicians' attitudes towards any key issue can be helpful in predicting legislative outcomes related to the topic.

The rest of the paper is structured as follows: sections 3.2 and 3.3 describe the data and the methodology used in the paper, respectively. Section 3.4 presents the results obtained. Section 3.5 concludes.

3.2 Data

3.2.1 Congressional Speeches and Speaker Data

The paper uses United States Congressional Records from 1990 to 2015. This is a record of all speeches on the floor of Congress (for both chambers of the Congress). The records were digitised using optical character recognition or OCR on scanned print volumes (Gentzkow et al., 2019). The source of the data from 1990 to 1993 is HeinOnline. The source of the data from 1994 to 2015 is the United States Government Publishing Office. The records were issued in bound editions compiling data from an entire Congressional session. This yields a total of 1,767,180 speeches. The speakers and the state they belong to are identified at the beginning of speeches. The speaker is matched with their biographical data obtained from Github repository on Congress legislators[3]. The repository data contains information on current and historical U.S. legislators such as party affiliation, year of birth, gender, race, military service and ICPSR code[4].

3.2.2 Voting Data

For voting records on immigration bills, voting dataset from Facchini and Steinhardt, 2011 is utilised, which provides data for 1265 immigration votes cast by 445 Congress members in the U.S. House of Representatives (speaker character-

[3] https://github.com/unitedstates/congress-legislators/tree/1473ea983d5538c25f5d315626445ab038d8141b
[4] ICPSR code is an identification number assigned to members of the U.S. Congress.

istics provided in Table 3.A.1) between 1990-2006. The dataset identifies 8 bills from 1990 to 2006, 3 before and 5 after the 9/11 attacks, which affect outcomes for immigrants in the U.S. The list only includes final votes and excludes intermediate votes (such as floor amendments). The list of these 8 bills and whether the bills were for (*pro*) or against (*contra*) liberalising immigration are presented in Table 3.1. Votes on immigration related issues increased substantially after the 9/11 attacks in 2001, although the effect was not immediate. Particularly, votes on immigrant rights and benefits peaked around 1995-1996 in the 104th Congress, only to be followed by a stark increase in votes on border security, after 9/11, in the 109th Congress in 2005-2006 (Fennelly et al., 2015).

Table 3.1: Final passage votes on immigration issues in the House of Representatives 1990-2006

Congress	Date	Bill	Topic	Keyword	Direction
101	3.10.1990	H.R.4300*	Immigration	The 1990 Immigration Act (IMMACT)	Pro
104	21.3.1996	H.R.2202*	Illegal Immigration	Illegal Immigration Reform and Immigrant Responsibility Act	Con
105	24.9.1998	H.R.3736	Skilled Immigration	Temporary Access to Skilled Workers and H-1B Nonimmigrant Program Improvement Act	Pro
109	10.2.2005	H.R.418	Illegal Immigration	Real ID Act	Con
109	16.12.2005	H.R.4437*	Illegal Immigration	Border Protection, Anti-terrorism and Illegal Immigration Control	Con
109	14.9.2006	H.R.6061	Illegal Immigration	Secure Fence Act	Con
109	21.9.2006	H.R.6094	Illegal Immigration	Community Protection Act of 2006	Con
109	21.9.2006	H.R.6095	Illegal Immigration	Immigration Law Enforcement Act of 2006	Con

This table is adapted from Facchini and Steinhardt, 2011. Major legislations on immigration during this period are marked using an asterisk (*).

Of the 8 bills, Facchini and Steinhardt, 2011 classified 3 as major immigration legislations. The first major bill was the *1990 Immigration Act* or *IMMACT*. Among other provisions, the *IMMACT* raised the annual cap for legal permanent residents to 675,000 from 530,000 (Leiden and Neal, 1990). This was followed by the *Illegal Immigration Reform and Immigrant Responsibility Act* of 1996. The bill's provisions included expedited removal procedure of undocumented immigrants and criminal penalty for those immigrants who re-enter or attempt to re-enter the U.S. within a specific time period after removal (Grable, 1998).

Of the bills post 9/11, a major bill was the *Border Protection, Anti-terrorism and Illegal Immigration Control Act* of 2005. The bill's provisions included the construction of a 700 miles long fence along the US-Mexico border and elimination of the visa lottery programme (Fetzer, 2006). More critically, if approved, the

bill would have made illegal presence in the U.S. an aggravated felony instead of a civil offence and made any individual found to have aided an undocumented immigrant to remain in the U.S. liable to up to five years in prison (Fetzer, 2006). Although the bill was passed in the house, it was not passed in the Senate.

3.3 Methodology

3.3.1 Immigration Corpus

All speeches were first cleaned by removing punctuation and upper case. The speeches were then lemmatized to reduce words to their base forms, while maintaining context, using NLTK WordNet lemmatizer (Bird et al., 2009).

From the Congressional records, speeches containing the words *immigration* or *refugee* or related words were identified to form the "immigration corpus" (similar to Li and Hills, 2020)[5]. The immigration corpus is what is referred to throughout the rest of the paper as a collection of all immigration policy-related speeches made on the floor of the Congress. All procedural phrases were removed from the immigration speeches by using the list of procedural bigrams provided by Gentzkow et al., 2019. Following Gentzkow et al., 2019, all speakers' last names and names of states and months were also removed from the speeches. The paper excludes speeches where the speaker is identified by office and not by name (such as 'chairman'). This yields a total of 24,351 immigration related speeches from 1098 distinct speakers between 1990 to 2015. The immigration corpus contains a total of 43,564,960 words and the average number of words per speech is 1789[6].

3.3.2 Valence and Concreteness

For each speech the paper computes the valence (or sentiment) and concreteness (or proxy for social distance) ratings. For the former, the paper uses valence norms proposed by Warriner et al., 2013. Warriner et al., 2013 provides valence norms for approximately 14,000 words, each rated on a scale of 1 to 9.[7]. The valence rating of a word refers to the pleasant emotion conveyed by it, with the rating increasing as it moves from unhappy to happy. Valence of a speech was computed by taking the mean valence rating of all words in the speech. For

[5]The full list of words is provided in Appendix 3.B

[6]It should be noted that after removing speeches where the speaker was identified by office and not by name, there were no immigration speeches identified in 2015. However, these speeches were re-incorporated in the corpus when training the LDA topic modelling algorithm as detailed in section 3.4.3

[7]Warriner et al., 2013 computed the valence of each word by taking the average rating of approximately 20 participants.

concreteness, the paper uses concreteness norms proposed by Brysbaert et al., 2014. Brysbaert et al., 2014 offers the concreteness ratings of nearly 37,000 English words, each rated on a scale of 1 to 5, going from abstract to concrete.[8] Concreteness of a speech was calculated by taking the mean concreteness ratings of all the words in the speech.

3.3.3 Topic Model

The paper uses a Latent Dirichlet Allocation or LDA model (Blei et al., 2003) to extract the different topics discussed in the immigration corpus[9]. The LDA model assumes that there is a fixed number of topics or patterns which explains the structure of a corpus. LDA is a Bayesian model in which each document of a corpus is modelled as a finite mixture over an underlying set of topics. The LDA model reduces each document in a corpus to a "bag-of-words" which is a representation of text that describes the number of occurrences of each word in the document, ignoring the sequence in which the words appear. The model assumes that a set of documents can be explained by a fixed number of underlying patterns or topics. Each document in the corpus is a distribution over these topics and each topic is a distribution over the words which form the corpus vocabulary.

For interpreting the topics extracted from the immigration corpus the paper examines the most relevant words pertinent to each topic. For this the paper uses the *relevance* measure proposed by Sievert and Shirley, 2014. Let V denote the number of words in the vocabulary of the immigration corpus, w be a word $\in 1, ..., V$, ϕ_{kw} be the probability that w belongs to topic $k \in 1, ...K$ and p_w be the marginal probability of word w in the immigration corpus. Then the relevance of word w to topic k is defined as:

$$r(w, k|\lambda) = \lambda log(\phi_{kw}) + (1 - \lambda) log(\frac{\phi_{kw}}{p_w}) \tag{3.1}$$

where λ $(0 \leq \lambda \leq 1)$ is the weight given to the probability that word w is assigned to topic k relative to the *lift* or $\frac{\phi_{kw}}{p_w}$.[10]

The paper aims to track trends in topics over different years, similar to Li, Hills,

[8] Brysbaert et al., 2014 computed the concreteness rating of each word by taking the average rating across roughly 30 participants.

[9] The LDA model was implemented using *gensim* package in Python.

[10] The *lift* of a word (Taddy, 2012) is the ratio of a word's probability within a topic to the word's marginal probability of being in the corpus. The lift measure is used to rank terms within topics. This paper calculated the top relevant words for each topic by trying different values of k, between 5 to 15 and different values of λ between 0.5 to 1, to find the most contextually appropriate topics and keywords, presented in Table 3.5.

and Hertwig, 2020. This was done to understand the key topics associated with immigration over time. All speeches that occurred in the same year were considered as 1 document. For each document the paper calculates the probability that document d belongs to topic k as:

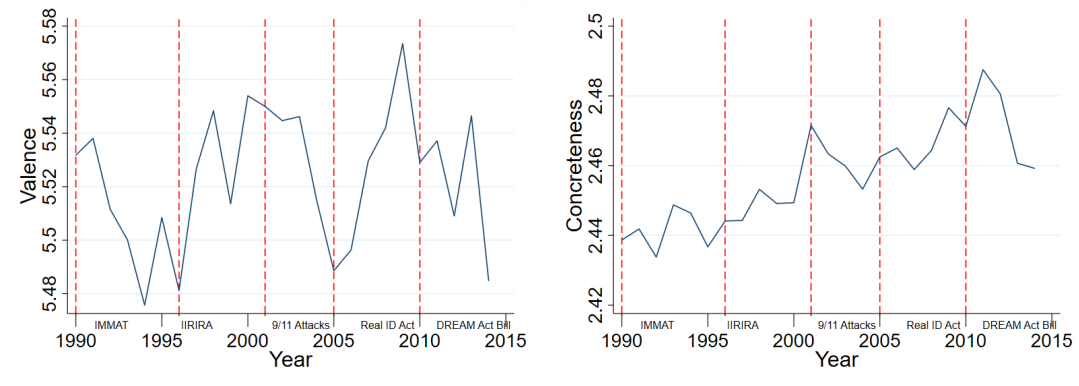
$$\phi_d(k) = \frac{|w \in d : topic(w) = k|}{|d|} \tag{3.2}$$

where the numerator is the number of words in document d that are assigned to topic k and the denominator is the total number of words in document d.

3.4 Results

The results are divided into 3 sections. Section 3.4.1 presents descriptive results which examine the variation of immigrant speech valence and concreteness over time, while also evaluating partisan and regional variation. Section 3.4.2 examines if linguistic measures of attitude towards immigrants can predict voting behaviour on immigration bills in the House of Representatives. Lastly, section 3.4.3 presents the results obtained from applying an LDA topic modelling algorithm to the immigration corpus.

3.4.1 Descriptive Results



(a) Valence (b) Concreteness

Figure 3.1: Variation in (a) valence and (b) concreteness of immigration policy-related speeches over time

This section begins by looking at the variation in valence and concreteness measures of Congressional speeches (in both chambers of the Congress) over time. Figures 3.1 (a) and (b) plots the mean valence and concreteness of immigration speeches over time, respectively. Figure 3.1 (a) shows a small rise in the valence measure following the *Immigration Act (IMMACT) of 1990*. The bill, which was

signed into law by President George H. W. Bush, raised the annual cap on immigrant admissions (Leiden and Neal, 1990), a positive step towards liberalising immigration policy. The amount of attention paid to immigration policy varies by year. The number of votes on immigration policy (including votes on amendments and procedure) declined during the 103rd Congress (1993-1994) and rose in the 104th Congress (1995-1996) (Fennelly et al., 2015). The Illegal *Immigration Reform and Immigrant Responsibility Act (IIRARA) of 1996*, signed by President Clinton on 30th September 1996, aimed primarily at curbing the flow of illegal immigration into the United States (Grable, 1998). This is expressed by a steep drop in the valence measure leading up to the introduction of IIRARA. This is followed by an increase in valence measure leading to the House passing the *Temporary Access to Skilled Workers and H-1B Nonimmigrant Program Improvement Act of 1998*, temporarily increasing the annual number of H1B visas[11].

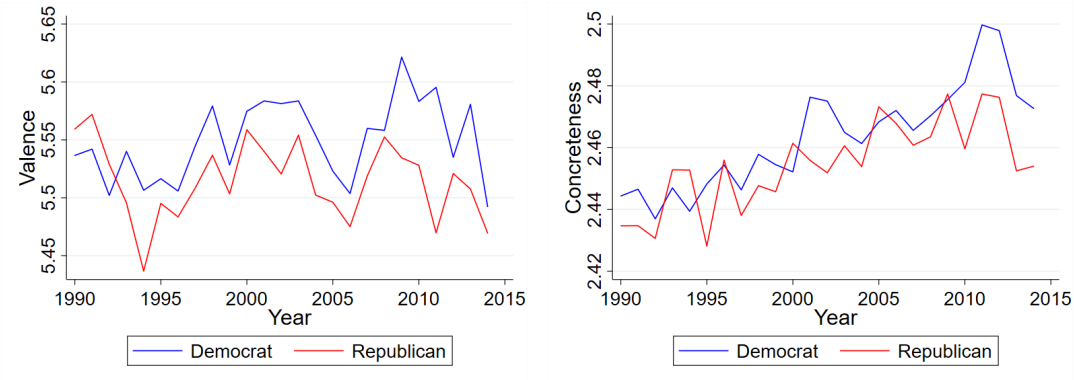
There was an increase in valence towards late 1990s leading to the Congress enabling around 400,000 illegal immigrants, who had been deemed ineligible for earlier amnesties, to obtain green cards in December 2000 (Schuck, 2005). Further, in March 2001, 150,000 undocumented Salvadorans acquired temporary legal status (Schuck, 2005). However, such positive immigrant measures came to an end with the attack on the World Trade Centre on September 11, 2001. There was a drastic decline in the valence measure after 2001. Following the attacks, there was a major shift in immigration policies in the U.S. Immigrant policy debates became centred on national security issues (Rodriguez, 2008; Hing, 2006; Schuck, 2005). The post 9/11 era saw the enactment of the *USA Patriot Act* in 2001, increasing the surveillance of immigrants, and the creation of the Department of Homeland Security in 2002, aimed at securing national borders and identifying perceived threat posed by non-citizens residing in the U.S. (Rodriguez, 2008).

From 2005 onwards, votes on immigration policy became centred on issues of border security (Fennelly et al., 2015). This included the *Real ID Act of 2005*, which, among other provisions, tightened the eligibility criteria for asylum seekers (Cianciarulo, 2006). This period also included other restrictive bills, adversely affecting immigrants, such as *Border Protection, Anti-terrorism and Illegal Immigration Control Act of 2005*, *Secure Fence Act of 2006* and *Immigration Law Enforcement Act of 2006*. The valence measure improved after this under the Democratic control of Congress (both House and Senate) from 2007-2011. Around this time, there were several attempts to pass different versions of the *Development, Relief, and Education for Alien Minors Act* or *DREAM Act*, aiming to grant legal status to undocumented individuals who came to the U.S. as minors. It came closest

[11]The H1B visa allows U.S. organizations to hire skilled foreign workers for specialised jobs.

to being passed in 2010 when the bill was passed in the House, but fell of short of the votes required to proceed in the Senate (American Immigration Council, 2021). Further, there was an increase in the valence measure around the time of President Obama's announcement of the *Deferred Action for Childhood Arrivals (DACA) in 2012*, by executive Branch memorandum. The DACA program entitled individuals who had travelled to the U.S. as minors to apply for deferred action from deportation (American Immigration Council, 2012).

The concreteness measure, which is a proxy for social distance (with higher values indicating close social proximity), has overall displayed an increasing trend during the period under consideration. This is consistent with Hills et al., 2017 who find that American English usage has become more concrete in nature with time. Although, Figure 3.1 (b) shows a fall in the measure following the 9/11 attacks in 2001, indicating an increase in perceived social distance towards immigrants.



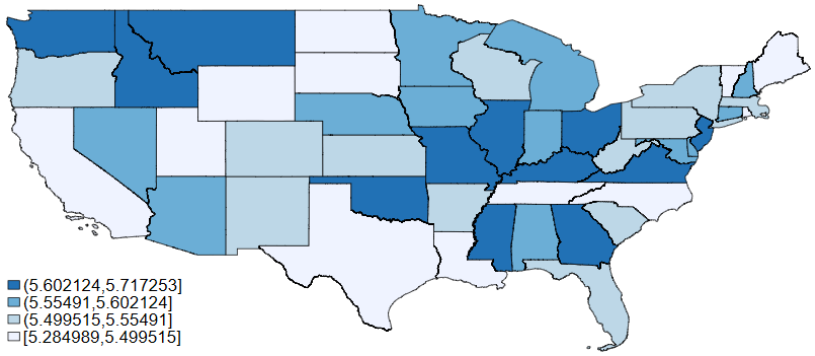
(a) Valence (b) Concreteness

Figure 3.2: Partisan variation in (a) valence and (b) concreteness of immigration policy-related speeches over time

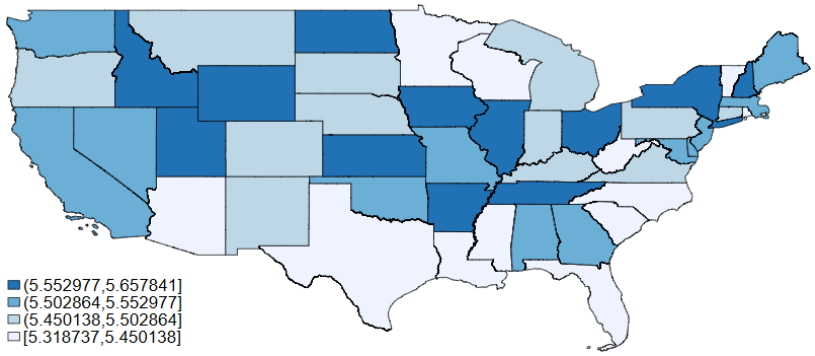
Next, the speakers were matched with data on individual characteristics. There were 1026 speakers, from both chambers of the Congress, for whom biographical data was available. These were grouped according to their party affiliation to examine the partisan variation of the valence and concreteness measures over time. There were 510 Republicans, 512 Democrat and 4 Independent legislators in the dataset. Figure 3.2 plots the results. Figure 3.2 (a) shows that the valence measure for Democrats and Independents (grouped together and labelled as Democrat) has been consistently higher than that for Republicans since early 1990s. The concreteness measure (Figure 3.2 (b)) has been higher for Democrats and Independents (grouped together and labelled as Democrat) than Republicans for most of the period under consideration, although the distinction is less clear. This is indicative of polarised attitudes of the Congress members towards immigrants, with Democrats showing more pro-immigrant attitude. This result

aligns with a vast literature which has documented the positive association between liberal ideology and support for immigration (Brooks et al., 2016; Kugler et al., 2014; Hainmueller and Hiscox, 2007; Hix and Noury, 2007; Fetzer, 2006).

Further, the paper examines valence and concreteness across the different states in the U.S. Figure 3.3 shows the mean valence of immigration speeches across states in years 1990 (Figure 3.3 (a)) and 2006 (Figure 3.3 (b)). The lighter shades indicate states with lower valence or sentiment. States with the lowest sentiment towards immigrants in 1990 were the southern states of Texas, Louisiana, Tennessee and North Carolina. Other low sentiment states were California, Utah and Wyoming in the west and North Dakota and South Dakota in the mid-west. Figure 3.3 (b) shows an increase in the number of southern states with low sentiment towards immigrants. In 2006 the southern states of Texas, Louisiana, Mississippi, Florida, South Carolina and North Carolina showed some of the lowest average valence measures. The anti-immigrant attitude in the south is consistent with Fetzer, 2006 who found that people from the southern U.S. states oppose immigration owing to their conservative political culture. Figure 3.A.1 shows the mean concreteness ratings of immigration-related speeches across states in 1990 and 2006 where the variation is less distinct compared to the valence measure.



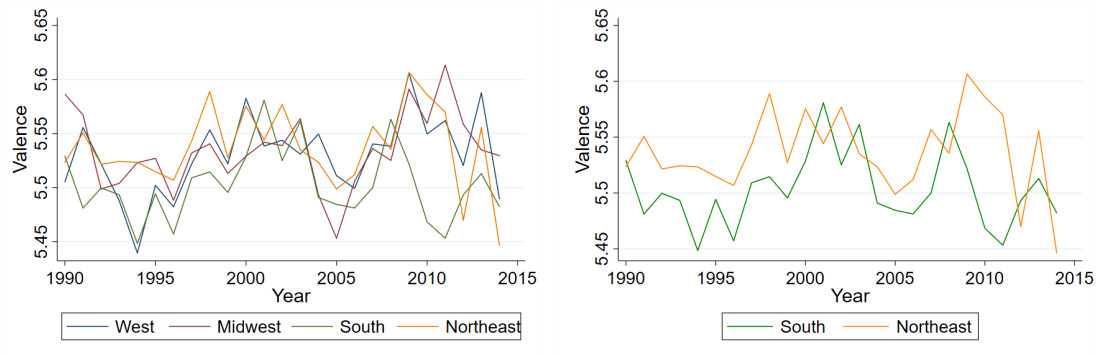
(a) Speech valence in 1990



(b) Speech Valence in 2006

Figure 3.3: Valence of immigration policy-related speeches across different states in the U.S. in (a) 1990 and (b) 2006

Figure 3.4 shows the average valence or sentiment of immigration related speeches over time in the different regions in the U.S. The Census Bureau of the United States divides the country into 4 regions - North-east, Midwest, South and West.[12] While there are clear overlaps, the north-eastern states appear to have had the most positive or highest sentiment towards immigrants over time. This is likely because of the strong democratic support in the North-eastern states, especially since the early 1990s (Harris, 2014). The southern states on the other hand have consistently displayed a lower speech sentiment. Figure 3.4 (b) shows the sentiment for just the north-eastern and southern states for clearer comparison. Regional variation of the concreteness measure over time, presented in Figure 3.A.2, was not as distinct as the sentiment measure.



(a) All regions

(b) South and Northeast

Figure 3.4: Variation in valence of immigration policy-related speeches over regions in the US. (a) shows valence over time for all regions (b) shows valence over time for the southern and north-eastern states.

3.4.2 Voting Behaviour

Having evaluated the variation in linguistic measures of attitude towards immigrants of Congress members over time, by party and by region, the paper next studies their voting behaviour. The paper seeks to examine if the linguistic measures computed for each speaker can predict their voting behaviour on bills related to immigrant issues. With regards to voting behaviour, the paper makes the following predictions:

Prediction 1: Speakers with higher speech valence (i.e. greater pro-immigrant

[12]The North-eastern States are Connecticut, Maine, New Hampshire, Massachusetts, New Jersey, New York, Pennsylvania, Rhode Island, Vermont. The Mid-western states are Illinois, Indiana, Iowa, Kansas, Michigan, Missouri, Minnesota, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin. The southern States include Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Oklahoma, North Carolina, South Carolina, Tennessee, Texas, Virginia, West Virginia. Lastly, the Western States are Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming.

sentiment) are more likely to vote in favour of bills liberalising immigration.

Speech sentiment is hypothesised to be a significant predictor of future voting behaviour on immigration bills. This result is expected to hold even after controlling for the speaker's individual characteristics, party affiliation and Congressional district characteristics. This is because speech valence captures the inherent opinion that Congress members have about immigrant issues, above and beyond their party's position and their district's demographic and socio-economic features. Further, due to the high frequency of speeches made and the length of these speeches, a measure computed from speeches is more informative of the speaker's voting intentions than other potential measures of attitude towards immigrants.

Prediction 2: Speaker speech valence is hypothesised to be a strong predictor of pro-immigration voting behaviour even after limiting the dataset to include only southern, western or south-western states or states with high foreign born population.

Members of the southern and western states in the U.S., owing to their conservative political ideologies or anti-immigration stance, have been observed to oppose immigration (Fetzer, 2006). In such states, where the sentiment towards immigrants is low (Figure 3.4), valence measure of speeches is expected to be a predictor of voting behaviour of Congress members. Moreover, prediction 2 posits that in states that are majorly opposed to immigration, the sentiment towards immigrants is more polarised and hence the valence measure is expected to be a good predictor of voting behaviour on immigration-related bills.

Hawley, 2011 noted that in the U.S., higher the local foreign born population in the state, greater the partisan influence on support for restricting immigration. This is consistent with the *group threat theory* which posits that competition for limited resources leads to conflict among different demographic groups i.e. larger the size of the out-group, greater the threat perceived by the in-group (Blumer, 1958; Blalock, 1967; LeVine and Campbell, 1972). Accordingly, it is hypothesised that in such polarised states, with large foreign born population, the polarised sentiment towards immigrants is a good predictor of Congress member's support towards immigration policy.

Prediction 3: Concreteness measure of immigration-related speeches is not a good predictor of voting behaviour of Congress members on immigration bills.

Concreteness of immigration-related speeches has been observed to grow over time during the period considered (Figure 3.1). This could imply that with

the growth in immigrant population over time in the U.S., the perceived social distance towards immigrants has declined. This is consistent with *contact theory* (Allport, 1954) which suggests that biases towards outgroups are lessened though more interaction. However, the rise in the concreteness measure over time could also be because American English usage has become more concrete in nature with time (Hills et al., 2017). Further, due to the unsystematic regional variation of the concreteness measure over time (Figures 3.A.1 and 3.A.2), concreteness is hypothesised to not be a good predictor of voting behaviour on immigrant issues.

The list of bills provided in Table 3.1 is used to examine the relationship between speaker speech sentiment and voting behaviour on immigration bills in the House of Representatives. The paper uses the probit model presented in equation (3.3) to evaluate the effect of speech sentiment on speaker's voting behaviour.

$$Prob(Vote_{idt} = 1) = \Phi(\gamma_1 Valence_{idt} + \gamma_2 Z_i + \gamma_3 Y_{dt} + \delta_s + Post_t + \delta_s \times Post_t) \quad (3.3)$$

where the binary dependent variable $Vote_{idt}$ takes the value 1 if the speaker i from district d votes in favour of liberalising immigration at time t. $\phi()$ is the cumulative distribution function. The key explanatory variable $Valence_{idt}$ is the average sentiment or valence of all immigrant related speeches delivered by the speaker a year leading up to the vote. For example, if a vote occurs in March 1996, the value of $Valence_{idt}$ will be the average valence value of all the immigration speeches delivered by the speaker between March 1995 and February 1996. Z_i accounts for the speaker $i's$ individual characteristics. Y_{dt} includes district level characteristics of district d in year t. The dummy variable $Post_t$ takes the value 1 if the vote occurred after the 9/11 attacks and 0 otherwise. δ_s are state dummies. The model also accounts for state specific shocks post 9/11 attacks by including $\delta_s \times Post_t$ interaction terms. The district level variables examined, along with their definitions, are provided in Table 3.2.

The results of equation (3.3) are presented in table 3.3 and the summary statistics for the variables used are provided in Table 3.A.2. The regressions use standardised values of the valence variable. The table reports the average marginal effects from probit regressions. Column 1 shows that an increase in 1 standard deviation in the valence or sentiment of the speaker's speeches increases the probability of a pro-immigration vote by 5 percentage points ($p < 0.01$). Column 2 also accounts for the speakers' personal characteristics namely, whether the speaker was non-white or female and the speaker's age. The valence measure still remains significant ($p < 0.05$). Further, it is observed that being non-white significantly ($p < 0.01$) increases the probability of a pro-immigration vote. The variables age

Table 3.2: District level variables

Variable	Definition
Foreign Born	The share of foreign born people in the total population in the district
Foreign Born Growth	The increase in the foreign born population compared to last period
African American	The share of African American people in the total population in the district
Hispanic	The share of people with Hispanic origin in the total population in the district
Democratic Vote Share	The Democratic share of the votes at the last House election.
Unemployment	The share of unemployed people in the total labour force in the district.
Highly Skilled ratio	The percentage of the population in the district over 16 who are employed in executive, administrative, managerial and professional speciality occupations
Unskilled ratio	The percentage of the population in the district over 25 years of age with less than 4 years of High School
Farm worker	The share of farm workers in the total labour force in the district.

The data for the district level variables used in the paper are obtained from Facchini and Steinhardt, 2011.

and female, though positive and significant in column 2, become insignificant after including additional control variables in columns 3 and 4.

Column 3, along with including the dummy variables Democrat (which equals 1 if the speaker is from a democratic party) and No military service (which equals 1 if the speaker has never performed any military service), also includes district level characteristics. Further, column 4 includes the contributions from labour and corporate political action committees (PACs), expressed as a share of total PAC contributions, as control variables. This is included in the regressions because interest groups have been observed to shape migration policy (Facchini, Mayda, and Mishra, 2011).

Columns 3 and 4 of table 3.3 show that even after controlling for speaker specific characteristics, district level variables and interest group contributions, speaker speech sentiment remains a significant positive predictor of voting behaviour. Column 4 shows that an increase in 1 standard deviation in speaker speech valence increases the probability of a pro-immigration vote by 2 percentage points ($p < 0.05$). Further, among the control variables, it was observed that being a democrat significantly increases the Congress member's probability of a pro-immigration vote. The share of farm workers has a significant negative effect. Also, the skill level of the population significantly impacts voting behaviour. An

Table 3.3: Impact of speech sentiment on Immigration votes

	All bills				Major bills
	(1)	(2)	(3)	(4)	(5)
Valence	0.0476***	0.0283**	0.0167*	0.0242**	0.0284*
	(0.0177)	(0.0140)	(0.0097)	(0.0100)	(0.0167)
Non-white		0.4714***	0.1262***	0.1575***	0.1685**
		(0.0492)	(0.0475)	(0.0457)	(0.0762)
Age		0.0037**	-0.0001	-0.0013	-0.0049**
		(0.0018)	(0.0013)	(0.0014)	(0.0019)
Female		0.0957**	0.0125	0.0059	-0.0286
		(0.0483)	(0.0335)	(0.0359)	(0.0544)
Democrat			0.2564***	0.2883***	0.2939***
			(0.0407)	(0.0608)	(0.0854)
No Military Service			-0.0404	-0.0447	-0.0562
			(0.0309)	(0.0347)	(0.0448)
Foreign Born			-0.0851	-0.4402*	-0.2257
			(0.2306)	(0.2354)	(0.4578)
Foreign Born Growth			-0.0442**	-0.0239	-0.0490**
			(0.0206)	(0.0187)	(0.0233)
African American			0.1714	0.1180	-0.1603
			(0.1539)	(0.1512)	(0.2166)
Hispanic			-0.1403	0.1088	-0.0265
			(0.1633)	(0.1802)	(0.2232)
Democratic Vote Share			0.1980*	0.1565	0.1192
			(0.1032)	(0.1047)	(0.1484)
Unemployment			-1.2480	-2.4244**	-2.1876
			(1.0413)	(1.1118)	(1.5581)
Highly Skilled Ratio			0.6117**	1.0202***	0.8588
			(0.2805)	(0.3377)	(0.5407)
Unskilled Ratio			1.3138***	1.7441***	1.9765***
			(0.3236)	(0.3760)	(0.4972)
Farm worker			-1.6431**	-2.5525***	-2.1445**
			(0.7112)	(0.7411)	(0.9935)
Labour PAC				-0.0900	-0.0067
				(0.1534)	(0.2179)
Corporate PAC				-0.1671	-0.1817
				(0.1139)	(0.1573)
Post 9/11	-0.0784	-0.1018	-0.1129	-0.1395	0.1137
	(0.1511)	(0.1271)	(0.0940)	(0.1442)	(0.2035)
N	1137	1082	1082	844	316

Standard errors, clustered at the speaker level, in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All regressions control for state fixed effects and state interacted with Post 9/11 dummy fixed effects.

increase in the percentage of highly skilled individuals in the district has a significant positive effect on pro-immigration vote, consistent with the results of Facchini and Steinhardt, 2011 and Milner and Tingley, 2011. However, contradictory to Facchini and Steinhardt, 2011, this paper also finds that an increase in the percentage of unskilled individuals in the district has a positive effect on the legislator's probability of voting in favour of liberalising immigration.[13] Column 5 replicates column 4 by only considering the major votes as highlighted in Table 3.1. Speaker speech sentiment remains a positive and significant ($p < 0.10$) predictor of pro-immigration voting behaviour. Thus, consistent with prediction 1, the paper finds evidence of the following:

Result 1: Speaker speech sentiment is a significant positive predictor of pro-immigration voting behaviour, even after controlling for speaker's individual characteristics, party affiliation and the district's socio-economic features.

Next, the paper evaluates the ability of the valence measure to predict pro-immigration voting behaviour in the southern, south-western and western states. The results are presented in Table 3.4. The table also looks at the voting behaviour of representatives in high immigration states i.e. states with a large share of foreign born individuals[14]. Speaker speech valence is a positive and significant predictor of voting in favour of liberalising immigration in the south, south-west, west and the high immigration states.

In the southern states, an increase in 1 standard deviation in speaker speech valence significantly ($p < 0.05$) increases the probability of a pro-immigration vote by 3 percentage points. The effect is even stronger in the south-western and western states. An increase in 1 standard deviation in speaker speech valence

[13]Literature exploring the relation between native skill levels and attitude towards immigration is divided and largely focusses on two perspectives - labour market competition and fiscal burden on public services (Hainmueller and Hiscox, 2010; Facchini and Steinhardt, 2011; Milner and Tingley, 2011; Facchini and Mayda, 2009). The labour market competition argument states that natives will oppose immigrants with same skill levels (because increased supply in labour will decrease wages), while supporting immigrants with contrasting skills (due to increase in real wages) (Hainmueller and Hiscox, 2010). On the other hand, the public finance argument posits that high-skilled natives will oppose low-skilled immigration, especially in states with high redistribution where immigrants have greater access to public services (Milner and Tingley, 2011). However, several researchers concur that there is a need for exploring variables which affect attitude towards immigration, beyond economic concerns (Burns and Gimpel, 2000; Hainmueller and Hiscox, 2010). Accordingly, this paper shows the impact of politicians' sentiment towards immigrants in predicting pro-immigration voting behaviour.

[14]Similar to Facchini and Steinhardt, 2011, this paper also evaluates the effect on pro-immigration voting behaviour in south-western states (Arizona, California, Colorado, Kansas, Nevada, New Mexico, Oklahoma, Texas, and Utah) and high immigration states i.e. states with the highest share of foreign born people. In the current dataset, the high immigration states were Arizona, California, Connecticut, Florida, Georgia, Illinois, Massachusetts, Maryland, New Jersey, Nevada, New York, Rhode Island, Texas, Virginia and Washington.

Table 3.4: Impact of speech sentiment on Immigration votes in different regions

	South	Southwest	West	High Immigration
	(1)	(2)	(3)	(4)
Valence	0.0331**	0.0419***	0.0499**	0.0211*
	(0.0156)	(0.0158)	(0.0207)	(0.0113)
Non-white	0.1464***	-0.0201	-0.0209	0.0835*
	(0.0359)	(0.0597)	(0.0763)	(0.0505)
Age	-0.0005	-0.0021	-0.0009	0.0001
	(0.0020)	(0.0022)	(0.0027)	(0.0015)
Female	0.1265*	-0.0041	-0.0691	0.0214
	(0.0674)	(0.0641)	(0.0721)	(0.0403)
Democrat	0.2062***	0.3338***	0.3452***	0.2391***
	(0.0532)	(0.0741)	(0.1159)	(0.0484)
No Military Service	0.0093	-0.0071	0.0015	-0.0339
	(0.0458)	(0.0526)	(0.0575)	(0.0351)
Foreign Born	1.2495***	-1.0623***	-1.9541***	-0.0794
	(0.3710)	(0.3401)	(0.4693)	(0.2506)
Foreign Born Growth	0.0399**	-0.1120***	-0.1200***	-0.0802***
	(0.0175)	(0.0406)	(0.0462)	(0.0274)
African American	-0.0332	-0.1835	-0.3169	0.1845
	(0.1992)	(0.3402)	(0.4909)	(0.1885)
Hispanic	-0.3181*	0.1793	0.4698	-0.0691
	(0.1894)	(0.2496)	(0.4147)	(0.1901)
Democratic Vote Share	0.0152	0.2849	0.2943	0.1784
	(0.1154)	(0.2137)	(0.3122)	(0.1192)
Unemployment	1.3779	-1.7474	-2.9285	-1.1945
	(1.2407)	(1.3157)	(1.9384)	(1.3109)
Highly Skilled Ratio	-0.7411*	0.9259	2.4481***	0.6858**
	(0.3906)	(0.5781)	(0.7374)	(0.3310)
Unskilled Ratio	0.0351	1.7023**	3.6025***	1.4614***
	(0.3271)	(0.6787)	(1.1269)	(0.3996)
Farm worker	-0.9956	-2.5328***	-4.6462***	-2.0781**
	(1.2290)	(0.9347)	(1.3900)	(0.8940)
Post 9/11	-0.5713***	0.0147	-0.1451	-0.0417
	(0.0955)	(0.0962)	(0.2772)	(0.2177)
N	369	325	260	734

Standard errors, clustered at the speaker level, in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All regressions control for state fixed effects and state interacted with Post 9/11 dummy fixed effects.

significantly increases the probability of a pro-immigration vote by 4 percentage points in the south-west ($p < 0.01$) and by 5 percentage points in the west ($p < 0.05$). In the high immigration states an increase in 1 standard deviation in valence significantly ($p < 0.10$) increases the probability of a pro-immigration vote by 2 percentage points. Further, being a non-white significantly increases the legislator's probability of a pro-immigration vote in the south and in the high immigration states. Being a democrat has a significant effect on the probability of a pro-immigration vote across all 4 specifications considered.[15] Thus, consistent with prediction 2, the following result is observed:

Result 2: Even after limiting the dataset to southern, western or south-western states, or high immigration states, speaker speech sentiment remains a significant positive predictor of pro-immigration voting behaviour.

Lastly, the ability of the concreteness measure to predict pro-immigration voting behaviour was tested for all votes, as well as for votes in southern, south-western, western and high immigration states. The results are presented in tables 3.A.3 and 3.A.4. For all regions, as well as for the selected geographic specifications, concreteness was not found to be a good predictor of voting behaviour. Thus, consistent with prediction 3, the following is observed:

Results 3: Concreteness was observed not to be a good predictor of the speaker's voting behaviour on immigration bills.

3.4.3 Immigrant Topics

This section presents the results from applying a Latent Dirichlet Allocation (LDA) model to the immigration corpus. The LDA algorithm helps to understand the key topics for immigration legislation, discussed in the U.S. Congress between 1990 to 2015. Table 3.5 presents 10 relevant words describing each of the 8 topics that were identified. Each topic was labelled by careful consideration of the keywords[16]. The topics reflect the possible categories of immigration legislation. The 8 topics are immigrant benefits or assistance, border security, national security, defense, educational policy, health care, policy amendment and

[15]The district level economic variables have different impacts on pro-immigration voting behaviour, depending on the geographic specification. This could potentially be due to cultural or ideological differences across the regions. However, further investigation with respect to the economic variables is beyond the scope of the present work.

[16]To maximise the dataset for training the LDA model the paper uses all speeches containing immigration, refugee and related words, even those where the speaker is identified by office and not by name. The LDA model was applied to 26,467 immigration related speeches in both chambers of the Congress.

U.S. Government support.[17]

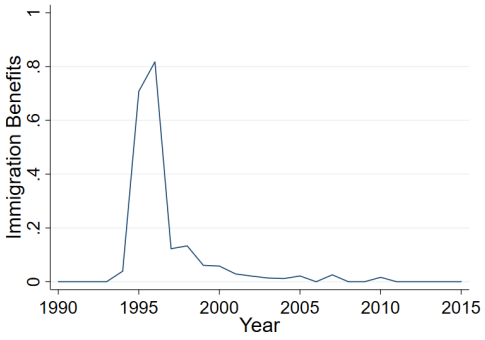
Table 3.5: Immigrant Topics and keywords

	Topic	Keywords
1	Immigrant Benefits	child, program, amendment, law, fund, assistance, public, government, health, provision
2	Border security	security, act, law, border, national, department, amendment, bill, work, right
3	National Security	security, national, intelligence, report, agency, service, government, federal, program, fund
4	Defense	law, fund, national, defense, service, bill, fund, department, American, act
5	Educational Policy	program, school, child, federal, fund, educational, education, law, assistance, amount
6	Health care	health, care, act, fund, provided, law, amount, plan, available, security
7	Policy Amendment	act, program, service, fund, amount, amended, law, fiscal, available, section
8	U.S. Government Support	refugee, law, war, government, American, support, need, right, world, program

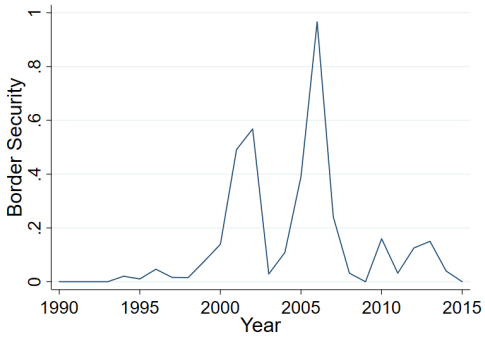
Next, the paper examines the distribution of these topics over time. Figure 3.5 plots the time variation of the 8 topics. The distribution of topics over time expresses the variation in concern over different immigrant-related issues and is generally reflected by the nature of the bills introduced during the time. Figure 3.5 (a) shows that discussion on benefits and rights of immigrants peaked in 1995-96, consistent with the findings of Fennelly et al., 2015. Figure 3.5 (b) shows that border security concerns first peaked in 2002 with the formation of the Department of Homeland Security in 2002, post the 9/11 attacks. This period also included the USA Patriot Act of 2001, which among other provisions, increased the surveillance of immigrants living in the U.S. (Rodriguez, 2008). The second peak in Border security topics was in 2005-2006 with bills such as Border Protection, Anti-terrorism and Illegal Immigration Control Act of 2005, Secure Fence Act of 2006 and Immigration Law Enforcement Act of 2006. Similarly, Figure 3.5 (c) shows a peak in national security-related discussions around mid-2000s.

For discussions related to defense, Figure 3.5 (d) shows a peak in 2000 (which could be due to the National Defense Authorisation Act of 1999, which had immigration provisions (Fennelly et al., 2015)). This is followed by another peak

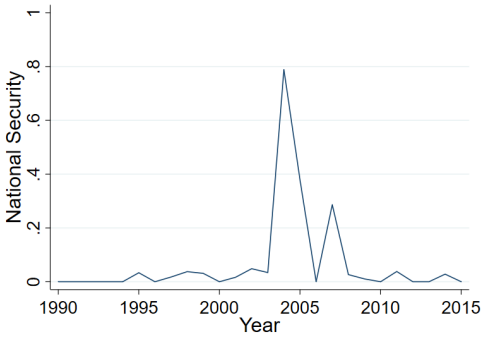
[17]The value of λ (defined in section 3.3.3) was set at 1 and the top 20 keywords were evaluated for each topic. Of the top 20 words, the 10 most interpretive words (and excluding words with no semantic interpretation such as would, year, subsection) are presented in the paper. Experimenting with different values of λ between 0.5 to 1 yielded very similar results.



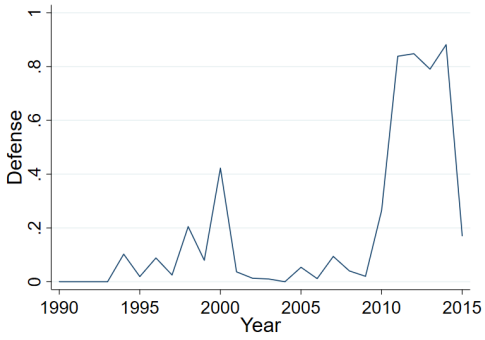
(a) Immigration Benefits



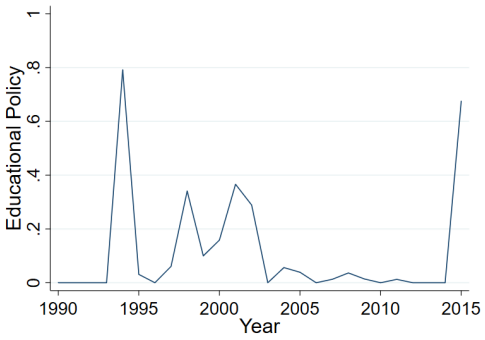
(b) Border Security



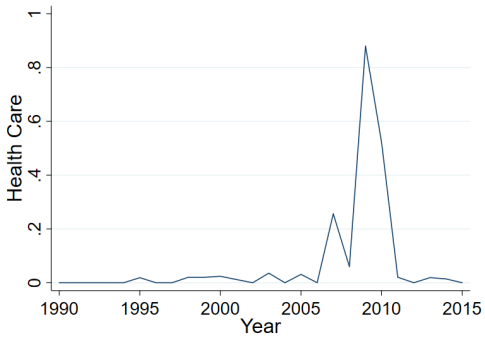
(c) National Security



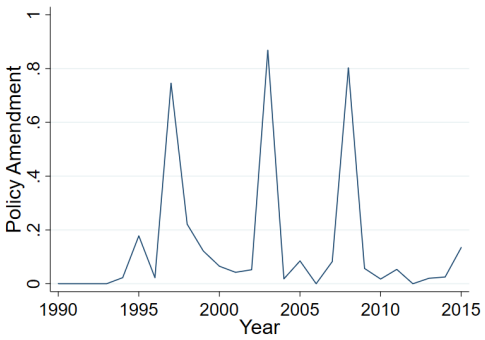
(d) Defense



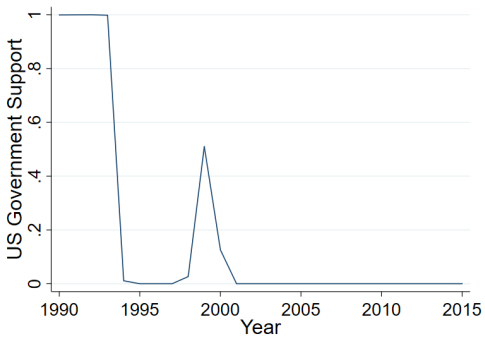
(e) Educational Policy



(f) Health Care



(g) Policy Amendment



(h) US Government Support

Figure 3.5: Immigrant Topics over Time

between 2011-2014. This period marked the U.S. military involvement in the Syrian Civil War. Figure 3.5 (e) shows that educational policy topics peaked in 1994, likely with the Departments of Labor, Health and Human Services, and Education, and Related Agencies Appropriations Act, which involved immigration provisions (Fennelly et al., 2015), and had scattered attention after that, with a prominent decline following 2001. Figure 3.5 (f) shows that health care topics peaked in 2009 (potentially with the Children's Health Insurance Program Reauthorisation Act of 2009, which included immigration provisions (Fennelly et al., 2015)). Not surprisingly, discussion over policy amendments in Congress have been scattered all over the period under consideration. Finally, discussions related to U.S. government support or aid was maximum between 1990-1993, a period which included the Gulf War, and tailed off since early 2000s.

3.5 Conclusion

Anti-immigrant narratives have taken the centre stage in various recent political events, be it the Brexit referendum in the United Kingdom in 2016 or Donald Trump's 2016 presidential campaign in the United States. With anti-immigrant sentiment on the rise, it is crucial to scrutinise politicians' inherent attitude towards immigrants, which plays a substantial role in shaping immigration policy in western countries.

This paper creates an "immigration corpus" which consists of 24,351 Congressional speeches containing *immigration*, *refugee* and related words, delivered on the floor of the U.S. Congress between 1990 and 2015. The corpus is used to build two distinct measures of attitude towards immigrants - *sentiment* or valence and language *concreteness*. Out of these two, sentiment showed systematic variation over time, consistent with the history of immigrant outcomes in the U.S., with a marked decline following the September 2001 terrorist attacks in the U.S. and the subsequent introduction of several anti-immigration bills. The sentiment measure was strictly polarised with democrats displaying a consistently higher sentiment than republicans in the period considered. The sentiment measure also displayed orderly variation over regions, with the conservative states in the south showing consistently low sentiment towards immigrants. The concreteness measure, barring fluctuations around major immigrant events such as the 9/11 attacks, displayed an overall increasing trend during the period considered.

The paper also computed speaker specific values of speech sentiment to show that sentiment is a significant positive predictor of speakers' pro-immigration voting behaviour on immigration policy-related bills. This result holds even after

controlling for speaker characteristics, party affiliation and district level socio-economic variables.

Further, the paper applies a Latent Dirichlet Allocation or LDA topic modelling algorithm to the immigration corpus to identify the different immigration policy-related topics discussed in the U.S. Congress and also to study the distribution of such topics over time. The topic distributions show a hike in concern over border security and national security following the 9/11 attacks and a decline in discussion on topics like educational policies and government aid.

Congressional speeches delivered on issues related to immigration are vital for predicting immigrant outcomes in the U.S. A linguistic measure of attitude towards immigrants developed using speeches could be more helpful for predicting immigrant outcomes compared to other measures like past voting behaviour of speakers on immigration bills. This is because the former provides a continuous measure thus capturing the *intensity* of the speaker's attitude compared to the latter which is a binary measure. Further, the U.S. Congress members play an important role in shaping legislation related to immigrant issues by voting either in favour of or against bills which liberalise or restrict immigration. Thus, a linguistic measure of attitude towards immigrants is useful for predicting the speaker's voting behaviour.

Overall, the study of speeches delivered by Congress members either during Congressional sessions or during campaign trails and other public appearances can be critical for building measures of the politician's attitude towards immigration or other key issues. Future researchers can use a similar methodology as employed in the paper to analyse politicians' attitudes on other crucial matters, such as gender, race and climate change, and use it to predict the politician's voting behaviour on bills related to such topics.

References

Allport, G. W. (1954). *The nature of prejudice*. New York: Doubleday.

American Immigration Council. (2012). Deferred Action for Childhood Arrivals: A Q&A Guide [Accessed 19 May 2021]. https://www.americanimmigrationcouncil.org/research/deferred-action-childhood-arrivals-qa-guide

American Immigration Council. (2021). The Dream Act: An Overview [Accessed 19 May 2021]. https://www.americanimmigrationcouncil.org/research/dream-act-overview

Ash, E., Chen, D. L., & Ornaghi, A. (2021). Gender attitudes in the judiciary: Evidence from us circuit courts. *Center for Law & Economics Working Paper Series*, *2019*(02). https://doi.org/10.3929/ethz-b-000472797

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. O'Reilly Media.

Blalock, H. M. (1967). *Toward a Theory of Minority-Group Relations* (Vol. 325). New York: Wiley.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of machine Learning research*, *3*, 993–1022.

Blumer, H. (1958). Race prejudice as a sense of group position. *Pacific sociological review*, *1*(1), 3–7.

Brooks, C., Manza, J., & Cohen, E. D. (2016). Political ideology and immigrant acceptance. *Socius*, *2*, 1–12. https://doi.org/10.1177/2378023116668881

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Burns, P., & Gimpel, J. G. (2000). Economic Insecurity, Prejudicial Stereotypes, and Public Opinion on Immigration Policy. *Political Science Quarterly*, *115*(2), 201–225. http://www.jstor.org/stable/2657900

Cianciarulo, M. S. (2006). Terrorism and asylum seekers: Why the real ID act is a false promise. *Harvard Journal on Legislation*, *43*, 101.

Facchini, G., & Mayda, A. M. (2009). Does the Welfare State Affect Individual Attitudes toward Immigrants? Evidence across Countries. *The Review of Economics and Statistics*, *91*(2), 295–314. https://doi.org/10.1162/rest.91.2.295

Facchini, G., Mayda, A. M., & Mishra, P. (2011). Do interest groups affect us immigration policy? *Journal of International Economics*, *85*(1), 114–128. https://doi.org/10.1016/j.jinteco.2011.05.006

Facchini, G., & Steinhardt, M. F. (2011). What drives U.S. immigration policy? Evidence from congressional roll call votes. *Journal of Public Economics*, *95*(7-8), 734–743. https://doi.org/10.1016/j.jpubeco.2011.02.008

Fennelly, K., Pearson, K., & Hackett, S. (2015). The US Congressional Immigration Agenda: Partisan Politics, Policy Stalemate and Political Posturing. *Journal of Ethnic and Migration Studies*, *41*(9), 1412–1432. https://doi.org/10.1080/1369183X.2015.1021586

Fetzer, J. S. (2006). Why Did House Members Vote for H.R. 4437? *International Migration Review*, *40*(3), 698–706. https://doi.org/10.1111/j.1747-7379.2006.00041.x

Gentzkow, M., Shapiro, J. M., & Taddy, M. (2019). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*, *87*(4), 1307–1340. https://doi.org/10.3982/ECTA16566

Goldin, C. (1994). The political economy of immigration restriction in the United States, 1890 to 1921. *The regulated economy: A historical approach to political economy* (pp. 223–258). University of Chicago Press.

Gonzalez, J. G., & Kamdar, N. (2000). Do Not Give Me Your Tired, Your Poor! Determinants of Legislator Voting on Immigration Issues. *Eastern Economic Journal*, *26*(2), 127–143.

Grable, D. M. (1998). Personhood Under the Due Process Clause: A Constitutional Analysis of the Illegal Immigration Reform and Immigrant Responsibility Act of 1996. *Cornell Law Review*, *83*, 820.

Hainmueller, J., & Hiscox, M. J. (2007). Educated Preferences: Explaining Attitudes Toward Immigration in Europe. *International Organization*, *61*(2), 399–442. https://doi.org/10.1017/S0020818307070142

Hainmueller, J., & Hiscox, M. J. (2010). Attitudes toward highly skilled and low-skilled immigration: Evidence from a survey experiment. *American Political Science Review*, *104*(1), 61–84. https://doi.org/10.1017/S0003055409990372

Harris, W. (2014). Essays on the Political Geography of The United States of America. *All dissertations*. https://tigerprints.clemson.edu/all_dissertations/1318

Hawley, G. (2011). Political Threat and Immigration: Party Identification, Demographic Context, and Immigration Policy Preference. *Social Science Quarterly*, *92*(2), 404–422. https://doi.org/10.1111/j.1540-6237.2011.00775.x

Hills, T. T., Adelman, J. S., & Noguchi, T. (2017). Attention economies, information crowding, and language change. *Big data in cognitive science.* (pp. 270–293). New York, NY: Routledge.

Hing, B. O. (2006). Misusing immigration policies in the name of homeland security. *CR: The New Centennial Review*, *6*(1), 195–224. http://www.jstor.org/stable/41949511

Hix, S., & Noury, A. (2007). Politics, not economic interests: Determinants of migration policies in the european union. *International Migration Review*, *41*(1), 182–205. https://doi.org/10.1111/j.1747-7379.2007.00061.x

International Organization for Migration. (2020). *World migration report 2020*. United Nations. https://doi.org/10.18356/b1710e30-en

Jensen, J., Naidu, S., Kaplan, E., Wilse-Samson, L., Gergen, D., Zuckerman, M., & Spirling, A. (2012). Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech. *Brookings Papers on Economic Activity*, 1–81.

Kugler, M., Jost, J. T., & Noorbaloochi, S. (2014). Another look at moral foundations theory: Do authoritarianism and social dominance orientation explain liberal-conservative differences in "moral" intuitions? *Social Justice Research*, *27*(4), 413–431. https://doi.org/10.1007/s11211-014-0223-5

Leiden, W. R., & Neal, D. L. (1990). Highlights of the US Immigration Act of 1990. *Fordham International Law Journal*, *14*, 328.

LeVine, R. A., & Campbell, D. T. (1972). *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior*. John Wiley & Sons.

Li, Y., & Hills, T. (2020). Language of Outgroup Prejudice: Revealing Perceptions towards 60 U.S Immigrant Groups from Language Patterns. https://doi.org/10.31219/osf.io/zxhde

Li, Y., Hills, T., & Hertwig, R. (2020). A brief history of risk. *Cognition*, *203*, 104344. https://doi.org/10.1016/j.cognition.2020.104344

Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of personality and social psychology*, *57*(6), 981.

Mastro, D., Tukachinsky, R., Behm-Morawitz, E., & Blecha, E. (2014). News Coverage of Immigration: The Influence of Exposure to Linguistic Bias in the News on Consumer's Racial/Ethnic Cognitions. *Communication Quarterly*, *62*(2), 135–154. https://doi.org/10.1080/01463373.2014.890115

Milner, H., & Tingley, D. (2011). The economic and political influences on different dimensions of united states immigration policy. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2182086

Rice, D., Rhodes, J. H., & Nteta, T. (2019). Racial bias in legal language. *Research & Politics*, *6*(2). https://doi.org/10.1177/2053168019848930

Rodriguez, R. M. (2008). (Dis)unity and Diversity in Post-9/11 America. *Sociological Forum*, *23*(2), 379–389. http://www.jstor.org/stable/20110274

Schuck, P. H. (2005). Immigrants' Political and Legal Incorporation in the United States after 9/11: Two Steps Forward, One Step Back. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.709783

Sherman, A., Trisi, D., Stone, C., Gonazales, S., & Parrott, S. (2019). *Immigrants Contribute Greatly to U.S. Economy, Despite Administration's "Public Charge" Rule Rationale* [Accessed 18 May 2021]. Washington, DC: Centre on Budget and Policy Priorities. https://www.cbpp.org/research/poverty-and-inequality/immigrants-contribute-greatly-to-us-economy-despite-administrations

Sievert, C., & Shirley, K. (2014). Ldavis: A method for visualizing and interpreting topics. *Proceedings of the workshop on interactive language learning, visualization, and interfaces*, 63–70.

Snefjella, B., & Kuperman, V. (2015). Concreteness and psychological distance in natural language use [PMID: 26239108]. *Psychological Science*, *26*(9), 1449–1460. https://doi.org/10.1177/0956797615591771

Taddy, M. (2012). On Estimation and Selection for Topic Models. In N. D. Lawrence & M. Girolami (Eds.), *Proceedings of the fifteenth international conference on artificial intelligence and statistics* (pp. 1184–1193). PMLR.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological review*, *117*(2), 440–463. https://doi.org/10.1037/a0018963

Turner, J. (1985). Social Categorization and Self-Concept: A Social Cognitive Theory of Group Behavior. In E. Lawler (Ed.), *Advances in group process: Theory and research* (pp. 77–121). JAI Press.

United Nations. (2017). *International Migration Report 2017: Highlights* (ST/ESA/SER.A/404). United Nations, Department of Economic and Social Affairs, Population Division. https://doi.org/10.18356/5e2626a2-en

United Nations. (2019). *International Migration 2019: Report* (ST/ESA/SER.A/438). United Nations, Department of Economic and Social Affairs, Population Division. https://www.un.org/en/development/desa/population/migration/publications/migrationreport/docs/InternationalMigration2019_Report.pdf

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45 VN - r*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Appendix

3.A Additional Tables and Figures

Table 3.A.1: Characteristics of Voters

Characteristic	Number of Speakers
Total Speakers	445
Republican	235
Democrat	210
Female	48
Non-White	77
No Military Service	304

Table 3.A.2: Summary Statistics of Voting Data

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
Vote	0.391	0.488	0	1	1265
Valence	5.526	0.194	4.635	6.262	1265
Age	56.211	9.481	30	81	1213
Foreign Born	0.112	0.108	0.002	0.585	1265
Foreign Born growth	0.507	0.810	-0.718	5.254	1265
African American	0.123	0.161	0.001	0.801	1265
Hispanic	0.118	0.151	0.003	0.745	1264
Democratic Vote Share	0.51	0.259	0	1	1260
Unemployment	0.059	0.025	0.018	0.204	1264
High Skilled Ratio	0.316	0.087	0.116	0.582	1264
Unskilled ratio	0.22	0.093	0.044	0.623	1265
Farm Worker	0.019	0.023	0	0.204	1264
Labour PAC	0.195	0.191	0	0.888	1010
Corporate PAC	0.369	0.15	0	1	1010

Table 3.A.3: Impact of speech concreteness on Immigration votes

	All bills				Major bills
	(1)	(2)	(3)	(4)	(5)
Concreteness	-0.0044	0.0084	0.0064	0.0130	0.0244
	(0.0172)	(0.0162)	(0.0120)	(0.0125)	(0.0188)
Non-white		0.4805***	0.1299***	0.1663***	0.1740**
		(0.0498)	(0.0477)	(0.0470)	(0.0795)
Age		0.0037**	-0.0001	-0.0011	-0.0046**
		(0.0019)	(0.0013)	(0.0014)	(0.0020)
Female		0.1001**	0.0119	0.0067	-0.0322
		(0.0486)	(0.0336)	(0.0371)	(0.0559)
Democrat			0.2562***	0.2845***	0.2968***
			(0.0411)	(0.0599)	(0.0852)
No Military Service			-0.0365	-0.0363	-0.0447
			(0.0309)	(0.0338)	(0.0457)
Foreign Born			-0.0614	-0.3835	-0.2031
			(0.2309)	(0.2372)	(0.4596)
Foreign Born Growth			-0.0434**	-0.0207	-0.0453*
			(0.0211)	(0.0194)	(0.0246)
African American			0.1636	0.1074	-0.1952
			(0.1556)	(0.1555)	(0.2190)
Hispanic			-0.1565	0.0763	-0.0580
			(0.1619)	(0.1775)	(0.2217)
Democratic Vote Share			0.1993*	0.1642	0.1067
			(0.1042)	(0.1057)	(0.1449)
Unemployment			-1.1203	-2.2365**	-1.8928
			(1.0496)	(1.1219)	(1.5861)
Highly skilled Ratio			0.6235**	1.0326***	0.9385*
			(0.2846)	(0.3517)	(0.5669)
Unskilled Ratio			1.3087***	1.7287***	1.9529***
			(0.3263)	(0.3798)	(0.5045)
Farm worker			-1.6471**	-2.5004***	-2.0986**
			(0.7024)	(0.7321)	(0.9878)
Labour Pac				-0.0805	0.0091
				(0.1515)	(0.2261)
Corporate PAC				-0.1418	-0.1485
				(0.1093)	(0.1565)
Post 9/11	-0.1175	-0.1294	-0.1308	-0.1694	0.0600
	(0.1590)	(0.1311)	(0.0965)	(0.1465)	(0.2097)
N	1137	1082	1082	844	316

Standard errors, clustered at the speaker level, in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All regressions control for state fixed effects and state interacted with Post 9/11 dummy fixed effects.

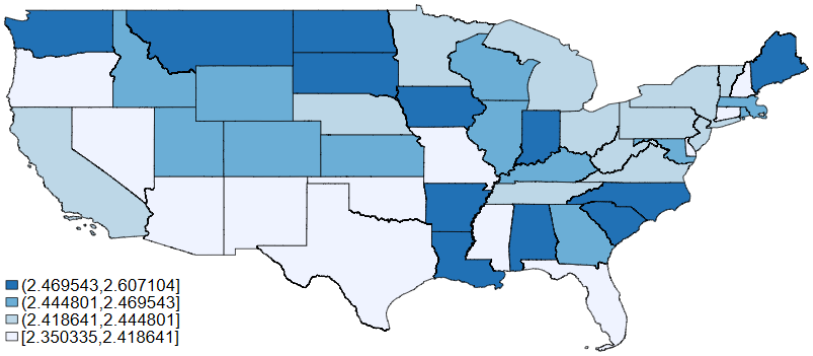
Table 3.A.4: Impact of speech concreteness on Immigration votes in different regions

	South	Southwest	West	High Immigration
	(1)	(2)	(3)	(4)
Concreteness	0.0113	0.0093	-0.0007	0.0214
	(0.0169)	(0.0226)	(0.0280)	(0.0152)
Non-white	0.1572***	-0.0105	-0.0051	0.0872*
	(0.0338)	(0.0599)	(0.0758)	(0.0512)
Age	-0.0003	-0.0017	-0.0009	0.0003
	(0.0021)	(0.0021)	(0.0027)	(0.0015)
Female	0.1105	0.0112	-0.0353	0.0197
	(0.0693)	(0.0630)	(0.0691)	(0.0405)
Democrat	0.1999***	0.3297***	0.3316***	0.2361***
	(0.0504)	(0.0721)	(0.1101)	(0.0494)
No Military Service	0.0212	-0.0091	-0.0089	-0.0275
	(0.0451)	(0.0551)	(0.0616)	(0.0354)
Foreign Born	1.2688***	-0.9215***	-1.7137***	-0.0463
	(0.3703)	(0.3166)	(0.4446)	(0.2483)
Foreign Born Growth	0.0500***	-0.1195***	-0.1252***	-0.0817***
	(0.0165)	(0.0398)	(0.0462)	(0.0276)
African American	-0.0503	-0.2020	-0.3737	0.1773
	(0.2012)	(0.3521)	(0.5058)	(0.1896)
Hispanic	-0.2971	0.1410	0.4187	-0.0767
	(0.1900)	(0.2535)	(0.4327)	(0.1855)
Democratic Vote Share	0.0290	0.2751	0.3273	0.1856
	(0.1079)	(0.2018)	(0.3017)	(0.1224)
Unemployment	1.7573	-1.4930	-2.2908	-1.0070
	(1.2729)	(1.2786)	(1.8888)	(1.3029)
Highly Skilled Ratio	-0.8132**	0.9388*	2.2346***	0.6885**
	(0.3934)	(0.5441)	(0.7128)	(0.3363)
Unskilled Ratio	-0.1527	1.6153**	3.1452***	1.4216***
	(0.3171)	(0.6278)	(1.0886)	(0.3982)
Farm worker	-1.0110	-2.2773***	-4.1269***	-2.0721**
	(1.1748)	(0.8558)	(1.2882)	(0.8502)
Post 9/11	-0.5912***	-0.0038	-0.1020	-0.0133
	(0.0960)	(0.0980)	(0.2401)	(0.1921)
N	369	325	260	734

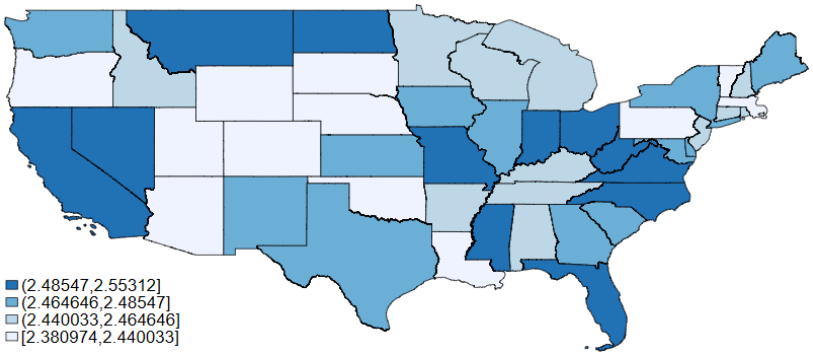
Standard errors, clustered at the speaker level, in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All regressions control for state fixed effects and state interacted with Post 9/11 dummy fixed effects.

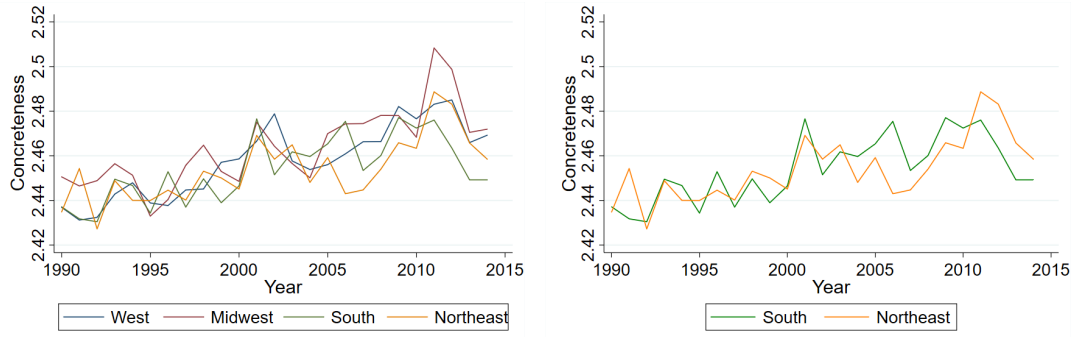


(a) Speech concreteness in 1990



(b) Speech concreteness in 2006

Figure 3.A.1: Concreteness of immigration policy-related speeches across different states in the U.S. in (a) 1990 and (b) 2006



(a) All regions

(b) South and Northeast

Figure 3.A.2: Variation in concreteness of immigration policy-related speeches over regions in the US. (a) shows concreteness over time for all regions (b) shows concreteness over time for the southern and north-eastern states.

3.B List of words used to form the "Immigration Corpus"

Following is the list of words used to identify speeches related to immigration:

Note that, since all speeches were lemmatized, to reduce words to their base forms, speeches containing any other variations of the words 'immigrant' and 'refugee'

(aside from the ones mentioned below) would have also been included in the corpus.

immigrants, immigrant, immigration, immigrations, immigrate, immigrates, immigrated, immigrating, refugee, refugees

Conclusion

This thesis contributes to the growing trend of using language analysis within the Economic discipline by using text analysis tools in a novel evaluation of three distinct forms of communication - small talk interaction, climate change communication and political speeches.

The first chapter makes a novel contribution to literature using personality theory to explain strategic decision making, by exploring the impact of *personality impressions* or *beliefs* developed about others on strategic behaviour. In a laboratory setting, it was observed that after engaging in a brief *small talk* chat with a stranger, the subject developed beliefs about the stranger's personality, particularly trait extraversion. Beliefs about the stranger's extraversion were observed to affect decision making in ensuing strategic interactions involving cooperation and level-k reasoning. Analysis of the language used during the small talk chat revealed that more talkative partners are believed to be more extraverted - a mechanism which indeed provides a reliable measure of trait extraversion. Overall, as the first study of its kind, the findings of this research can pave the way for further work exploring the link between personality impressions and strategic behaviours, potentially involving longer communication times and real world settings.

The second chapter contributes to research examining climate change communication strategies to promote a reduced meat diet. A comparative evaluation of 6 information interventions, formulated using six supporting theories, against a control group with baseline scientific information, was conducted using an online randomised control trial. It was observed that the most effective interventions in reducing planned meat intake were centred on providing more scientific knowledge and presenting the consequences of one's actions in an easily understandable unit i.e. *efficacy salience*. The study also found support for a *targeted messaging approach* and evidence of *motivated reasoning* around meat intake. Evaluating donations to a climate change charity and analysing the text in the evidence recalled at the end of the study generated further insight into the interventions. Overall, the findings of this study can offer guidance to future researchers and policy-makers about effective strategies for large-scale communication campaigns to promote a plant-based diet.

Finally, the third chapter contributes to research developing linguistic measures of attitude or bias against vulnerable groups. The chapter compiled an "immi-

gration corpus" containing immigration policy-related speeches delivered in the U.S. Congress between 1990 and 2015. The corpus was used to develop two distinct measures of attitude towards immigrants - *sentiment* (or valence) and *concreteness*. Of the two measures, sentiment, in addition to displaying systematic variation across time and states consistent with the history of immigrant outcomes, was a significant predictor of the speaker's voting behaviour on immigration bills. Applying a Latent Dirichlet Allocation (LDA) topic modelling algorithm to the immigration corpus demonstrated trends in the different topics related to immigration policy (such as national security) that have been debated in the Congress over time. Moreover, a similar methodology to this study can be adopted by future researchers to build linguistic measures of politicians' attitudes towards other key issues such as gender, race and climate change, to predict legislative outcomes.