Cognitive Abilities and the Demand for Bad Policy *

Salvatore Nunnari† Eugenio Proto‡ Aldo Rustichini§

January 22, 2025

Abstract

Rational choice theories posit that citizens accurately evaluate policy options. However, many policies—such as price controls and Pigouvian taxes—produce *equilibrium effects* that citizens may underestimate, potentially leading to support for detrimental policies or opposition to beneficial ones. This under-appreciation may be linked to cognitive abilities of citizens, raising fundamental research questions: Do cognitive abilities influence citizens' preferences regarding policies, particularly untested reforms? If so, what mechanisms underlie this influence? Furthermore, do citizens take into account how others citizens' cognitive abilities impact their choices? In this study, we employ a theoretical framework and conduct an experiment involving a UK-representative sample to demonstrate that enhanced cognitive abilities can lead to improved policy preferences. We also highlight the significant role of beliefs about the cognitive abilities of *other citizens*: individuals with high cognitive skills may withhold support for beneficial policies if they are skeptical about others' ability to accurately assess positive effects. We substantiate these findings through a textual analysis of open-ended survey responses, where participants forecast the consequences of labor and immigration reforms, and provide further evidence from the UK Household Longitudinal Study linking cognitive abilities to political attitudes. These findings carry important policy implications: educational programs aimed at developing cognitive skills and interventions designed to enhance trust in others' understanding could improve the quality of democratic decision-making.

Keywords: Equilibrium Effects, Policy Reform, Voting, Cognition, Experiment

JEL Codes: C90, D72, D91

*We thank David Gill, Nicola Lacetera, Kai Ou, Davide Pace and Victoria Prowse for helpful comments and suggestions. We are grateful to audiences at St Andrews, CEU, NYU, NYU-AD, TSE, Columbia, Amsterdam, Durham, the LSE Behavioral Political Economy Conference, the RISL$\alpha\beta$ Workshop in Cognitive Economics, the 2024 PoleconUK Annual Conference at King's College, the 2024 BSE Summer Forum, the 2024 Wallis Institute Conference in Political Economy, the 2024 CESifo Annual Conference in Behavioral Economics, and the 2025 ASSA Meeting in San Francisco for helpful comments and discussions. Bakdaulet Baitan, Eleonora Guseletova, and Massimiliano Pozzi provided excellent research assistantship. Nunnari acknowledges financial support from the European Research Council (Grant No. 852526). The pre-registration is available at https://osf.io/mjpsu/?view_only=13201a5b01324d969a030279506c3d88

†Bocconi University, Department of Economics; CEPR; and CESifo.

‡University of Glasgow, Adam Smith Business School; CEPR; CESifo; and IZA.

§University of Minnesota, Department of Economics.

1 Introduction

Public policy decisions are the litmus test of the effectiveness of democratic societies. Citizens are expected to choose effectively among various options, yet this reliance on democratic choice presumes voters can effectively discern the relative merits of policy alternatives. Widely accepted rational choice theories assume that voters are capable of such evaluations (Black, 1948; Downs, 1957; Feddersen and Pesendorfer, 1997; Besley and Coate, 1997; Persson and Tabellini, 2002; Börgers, 2004; Myatt, 2007; Fearon, 2011; McMurray, 2013).[1] However, many policies are inherently complex and produce outcomes not only through their immediate effects but also through indirect or *equilibrium effects* that may elude the average voter.

The equilibrium effects of policies arise from changes in incentives and subsequent adjustments in citizens' behavior, which take time to materialize and are challenging to anticipate because they require expectations on own and others' future behavior. Policies that generate such effects include measures that can have profound implications on the economy and society, such as regulating prices (e.g., minimum wage and rent control), imposing Pigouvian taxes, monetizing fiscal deficits, expanding road infrastructure, and introducing Universal Basic Income. As noted by Smith (1776), North (1990), Romer (2003), Caplan (2011), Beilharz and Gersbach (2016), and Achen and Bartels (2017), citizens often struggle to appreciate equilibrium effects, leading to potential misjudgments about the efficacy of such policies. Recent research further highlights how voters may oppose policies that, despite imposing immediate or direct costs, would ultimately resolve social dilemmas and enhance welfare (Dal Bó, Dal Bó and Eyster, 2018; Vora, 2023).

These challenges raise critical questions: How do voters' cognitive abilities influence their understanding of complex policies? Through what mechanisms do these abilities shape policy

[1] In contrast, expressive voting theories argue that individuals vote primarily as an expression of their values, identity, or beliefs, deriving intrinsic satisfaction from the act of voting regardless of the outcome (Carter and Guerette, 1992; Brennan and Lomasky, 1997; Brennan and Hamlin, 1998; Tyran, 2004; Pons and Tricaud, 2018). Although we acknowledge the empirical relevance of this perspective, our study focuses on (boundedly) rational choice theory.

preferences? How does the confidence in the rational behavior of others affect behavior? This study addresses these questions by proposing a simple theoretical framework and an experimental design explicitly tailored to investigate one potential channel: the hypothesis that higher cognitive abilities enhance understanding of the equilibrium effects of untested policies.

In our theoretical framework, a policy reform changes the game played by citizens, preserving the action space but altering payoffs and incentives. This impact is decomposed into two components: (1) the "direct" effect, reflecting changes in payoffs due to the policy itself while keeping actions fixed, and (2) the "indirect" effect, capturing changes in payoffs following the adjustment of players' actions. Voters assess these two effects together to form their policy preferences. Our key assumption is that players do not necessarily choose the optimal action and that greater cognitive abilities lead to a greater chance of choosing the optimal action in a given policy environment. The main hypothesis is that for voters to support reforms with direct costs and indirect benefits, they must possess both the cognitive abilities to understand the policies' equilibrium consequences and the confidence that other citizens have the cognitive capacity to do the same.

To test this hypothesis, we recruit a large and representative sample of UK nationals and ask them to participate in experimental games. We focus on the simplest setting in which meaningful equilibrium effects can arise, namely, 2×2 complete information games with a unique equilibrium in strictly dominant strategies. After earning experience with one game, we give participants the opportunity to enact a policy reform. The resulting game has lower payoffs across all action profiles but yields higher equilibrium payoffs. We find that reforms that yield benefits through equilibrium effects are unpopular and, as predicted, supported primarily by citizens with high cognitive abilities and high confidence in others' cognition. Since we exogenously manipulate beliefs in others' cognition, the latter effect is causal. Interestingly, our results also indicate that educational attainments have an impact on support for efficient policy reform comparable to and independent from cognitive abilities

— measured through a Raven Progressive Matrices quiz and a Cognitive Reflection Test (CRT).

We complement this experimental analysis with a textual examination of responses to open-ended survey questions on two policy reforms central to the political debate in the UK. Specifically, we asked participants about their perceived consequences of raising the minimum wage and implementing policies to reduce immigration. We compiled a list of indirect effects identified by the respondents and found that the ability to list a greater number of indirect effects correlates with both their cognitive abilities and their political leanings.[2]

An interesting insight from our analysis is that performance on the Raven's Progressive Matrices test—which measures the ability to solve abstract problems—and performance in the CRT—which assesses the tendency to engage in reflective thinking—have distinct and independent effects. While Raven's scores are strongly associated with the ability to comprehend general equilibrium effects in the experimental games, CRT scores are the only predictor of the ability to identify the indirect effects of real-world policies in the survey questions. We argue that these distinct effects arise because individuals are likely already aware of possible indirect effects stemming from changes in minimum wage or immigration thresholds. Thus, the critical factor when answering these questions is the capacity for systematic and reflective thinking.

We believe that the results from our study have significant implications for real-world democracies and policymaking. First, they suggest an additional benefit of enhancing cognitive abilities and education: not only could individual economic prospects improve, but so might the quality of societal decision-making.[3] Second, they highlight a fundamental distinction between policies where indirect effects depend primarily on market forces—such as increasing the minimum wage or imposing rental price ceilings—and those where indirect

[2]This latter correlation is arguably driven by a process of motivated reasoning, often considered a key factor in belief formation. Motivated reasoning refers to the tendency of individuals to shape their assessments of information based on goals or ends extrinsic to accuracy (e.g., Kahan, 2013).

[3]In the Conclusion, we provide more details about the malleability of cognitive abilities.

effects depend primarily on others' individual behavior, as is often the case with environmental policies. In the former, preferences for reforms rely on individuals' ability to predict these effects, while in the latter, cognitive abilities interact with beliefs about others' abilities to shape policy preferences.

Recent contributions analyzing environmental policy reforms suggest that support for stricter carbon taxes or other environmental regulations often depends on individuals' trust in others' compliance and, arguably, their understanding of the policies' effects. If citizens believe others will not change their behavior to reduce pollution or consumption, their own sacrifices may seem futile, leading to opposition. Conversely, trust in broad compliance increases the likelihood of support, as individuals perceive collective action as effective. Dechezleprêtre et al. (2022) explore global perceptions of climate change. They show that beliefs about the effectiveness of policies are the most significant predictor of policy support and provide evidence consistent with the idea that these beliefs causally affect policy preferences. Further reinforcing this point, Andre et al. (2024) document widespread misperception about pro-climate behaviors and show that correcting these misperceptions significantly increases pro-climate behavior, such as donations to environmental causes.

These examples illustrate how beliefs about others' rationality and behavior are central to understanding public support for some reforms and how a deeper understanding of citizens' cognitive abilities and their role in forming these beliefs is critical for understanding the intricacies of voting behavior. Our experiment provides participants with explicit signals about others' cognitive abilities through test scores. Similarly, outside the laboratory, individuals form such beliefs on the basis of indirect signals such as education, profession, language use, or observed decision-making. These beliefs, though possibly imprecise, shape social interactions and influence judgments about others in everyday life.

Our experimental paradigm builds on Dal Bó, Dal Bó and Eyster (2018), but introduces several key innovations. First, we implement a fine-grained measurement of subjects' cognitive abilities. Second, we experimentally manipulate beliefs about others' cognitive abilities, a novel approach that allows us to isolate their role in shaping preferences. Third, we conduct our study with a non-student population broadly representative of the UK, addressing concerns about external validity. Finally, we analyze responses to questions about the perceived consequences of real-world policies, providing richer and more policy-relevant insights.

The growing interest in analyzing the link between cognitive ability and political preferences is underscored by an expanding body of empirical literature that documents a relationship between cognitive ability and political behavior (Deary et al., 2008b; Morton et al., 2011; Oskarsson et al., 2015; Zmigrod et al., 2020; Durante et al., 2019).[4] This body of work provides important real-world context for our theoretical and experimental approach. However, much of the existing evidence is based on observational data and does not elucidate the specific mechanisms by which cognitive abilities shape political behavior. By contrast, our study seeks to uncover the causal pathways through which cognitive abilities influence political behavior, addressing a critical gap in the literature.

Our work contributes also to a blossoming literature analyzing how bounded rationality and cognitive abilities influence behavior in games (Burks et al. 2009, Alaoui and Penta 2016, Gill and Prowse 2016, Proto, Rustichini and Sofianos 2019, 2022, Fe, Gill and Prowse 2022, Lambrecht et al. 2024).[5] By integrating theoretical and experimental approaches, our study advances this frontier, investigating how cognitive abilities influence voters' understanding of complex policy trade-offs and the formation of their policy preferences, with implications for democratic decision-making and policy design.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework that informs our study, while Section 3 outlines the experimental methodology.

[4] In addition, Hillygus (2005), Deary, Batty and Gale (2008a), Denny and Doyle (2008), and Elinder and Erixson (2022) investigate the association between cognitive abilities (often imperfectly proxied by education) and voter turnout. Dal Bó et al. (2017) examine the relationship between cognitive abilities and selection into political careers, while Ottinger and Voigtländer (Forthcoming) provide historical evidence from European monarchs, showing that rulers' cognitive ability significantly influenced state performance, particularly in contexts with limited institutional constraints.

[5] Agranov et al. (2012) and Halevy, Hoelzemann and Kneeland (2023) study how behavior in games depends on opponents' observed characteristics that may be correlated with their strategic sophistication (e.g., whether the opponent is a Ph.D. student or an undergraduate student).

Section 4 analyzes participants' behavior in our experimental games, and Section 5 conducts a textual analysis of participants' perceived consequences of complex economic and social reforms. In Section 6, we show evidence from Understanding Society — a socioeconomic survey on a large UK-representative sample — that is broadly consistent with the assumptions and the predictions of our theoretical framework. Section 7 concludes and discusses the policy implications of our findings.[6]

2 Conceptual Framework and Testable Hypotheses

2.1 The Games

We build on the conceptual framework from Dal Bó, Dal Bó and Eyster (2018) (DBDBE henceforth) to formulate specific hypotheses. Consider the two games presented in Table 1. In both cases, two players must simultaneously choose between Cooperate (C) and Defect (D). In the top panel, Cooperation results in a cost c for the player and a benefit b for the opponent, with $b > c > 0$. This is a classical Prisoner's Dilemma (PD henceforth) where defection is a dominant strategy, so that the unique Nash Equilibrium of this game is (D,D).[7]

Now we consider a policy proposal that would impose on each player taxes t_C and t_D, respectively, on cooperation and defection. If this proposal passed and was implemented, it would transform the payoffs of the game as illustrated in the bottom of Table 1. Taxes are set to satisfy the following conditions: $b > t_D > t_C + c$. The dominant strategy of the resulting game, labeled as Harmony Game by DBDBE (HG henceforth), becomes cooperation by both players, leading to a unique and efficient Nash Equilibrium (C,C).

Given the choice between playing the PD or the HG with a rational opponent, a player

[6]The pre-registration is at https://osf.io/mjpsu/?view_only=13201a5b01324d969a030279506c3d88.

[7]In what follows, we assume that players' objective function is to maximize their payoff from the game (or that game payoffs represent players' utilities and their objective function is to maximize the utility from the game). A necessary condition for this assumption to be satisfied is that players are perfectly selfish and do not care about others' outcomes. As we detail in Section 3, our experiment is carefully designed to remove social preferences.

Table 1: **The Games**

Prisoner's Dilemma

	C	D
C	$b-c, b-c$	$-c, b$
D	$b, -c$	$0, 0$

Harmony Game

	C	D
C	$b-c-t_C, b-c-t_C$	$-c-t_C, b-t_D$
D	$b-t_D, -c-t_C$	$-t_D, -t_D$

who anticipates equilibrium behavior in both games would prefer to play the HG: the equilibrium payoff in HG is $b-c-t_C > 0$, while the equilibrium payoff in PD is 0. Hence, if asked to express a preference over the games ex-ante, perfectly rational players who believe in the opponent's rationality would choose the HG.

2.2 Cognitive Abilities and Strategic Sophistication

We depart from this concept of perfect rationality and common knowledge of rationality and we assume instead that rationality depends on players' strategic sophistication — which, in turn, depends on their cognitive abilities — and on beliefs about other players' strategic sophistication — which, in turn, depends on beliefs about other players' cognitive abilities. Accordingly, let each player have a true cognitive ability, $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$.[8] We make the following assumptions:

Assumption 1 *There is a function $f(\cdot)$ that, for any $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$, assigns a probability that a player with cognitive ability θ understands that there is a dominant action in each game.*

Assumption 2 *When a player understands that there is a dominant action in each game, he chooses that action. When a player does not understand that there is a dominant action in each game, he randomizes uniformly among the available actions.*

[8]Each player has a subjective estimate of his own θ but this subjective estimate does not play any role in the analysis below, so we do not label it.

Note that $f(\theta)$ is an objective function, not a subjective belief about one's own or any other player's cognition or behavior. These assumptions are meant to capture the following decision-making process: a player with cognitive ability θ examines a game, goes through some reasoning, and with probability $f(\theta)$ reaches the conclusion that an action is dominant and that therefore he should choose that action regardless of any belief about the opponent. With the complementary probability, $1 - f(\theta)$, the reasoning process ends with no useful conclusion and the player then tosses a coin to determine what action to choose (as a Level-0 or naïve player in a Level-k Thinking or Cognitive Hierarchy model).

We impose more structure on the function introduced in Assumption 1 by assuming that greater cognitive ability is (stochastically) associated with greater strategic sophistication:

Assumption 3 *A player with greater cognitive ability is more likely to understand that there is a dominant action in each game, that is, $f(\theta)$ is continuous and strictly increasing in θ.*

2.3 Predictions on Preferences over Games

To derive predictions on players' preferences over games (or "policies"), we need to make assumptions on players' beliefs about their opponent's behavior. In our experiment, we give participants explicit information about their opponents' cognitive abilities, and we discuss how our design allows us to test our theory in the following subsection. As mentioned in the Introduction, outside of the laboratory, people form these beliefs on the basis of indirect signals. Here, we assume that players know their opponent's cognitive abilities and make the following assumption.

Assumption 4 *If a player understands the game (in the sense of Assumption 1) and believes the opponent has cognitive ability θ, he believes that the opponent understands the game with probability $f(\theta)$. If a player does not understand the game, he believes that opponents randomize uniformly among the available actions regardless of their cognitive ability θ.*

When ranking games, players prefer the game that maximizes their expected payoffs.[9] Consider a player with cognitive ability θ_s and an opponent with cognitive ability θ_o. There are two cases. With probability $f(\theta_s)$, the player chooses the dominant action in each game and attributes a probability $f(\theta_o)$ to the opponent choosing the dominant action in each game. Then, the player's expected payoffs from the HG and the PD are, respectively

$$\begin{aligned} \text{EU(HG)} &= f(\theta_o)(b - c - t_C) + (1 - f(\theta_o))(-c - t_C) \\ \text{EU(PD)} &= f(\theta_o)0 + (1 - f(\theta_o))b \end{aligned}$$

The player prefers the HG as long as he believes the opponent is sufficiently likely to choose the dominant action in each game, that is, as long as the opponent is sufficiently more likely to cooperate in the HG than in the PD. Since $f(\theta_o)$ is continuously increasing in θ_o, the player prefers the HG if and only if $f(\theta_o) > (b+c+t_C)/2b \in [0,1]$. With the complementary probability, $1 - f(\theta_s)$, the player chooses an action randomly in each game and expects the opponent to do the same. Therefore, he prefers the PD, which has a greater expected payoff under the uniform distribution. This discussion leads to the following prediction:

Proposition 1 *Consider a player with cognitive ability θ_s facing an opponent with cognitive ability θ_o. With probability $f(\theta_s)$, the player prefers the HG iff $\theta_o > f^{-1}\left(\frac{b+c+t_C}{2b}\right)$. With probability $1 - f(\theta_s)$, the player prefers the PD regardless of θ_o.*

Proposition 1 implies that a participant with greater cognitive ability is more likely to prefer the HG if and only if he is paired with an opponent with a sufficiently high cognitive ability. We bring this prediction to the laboratory.

[9] As discussed in Section 3, in our experiment, we measure participants' preferences between the PD and the HG with a design that excludes any role of signaling or selection motives.

2.4 Testable Hypotheses

Assumption 3 already delivers a testable hypothesis on the relationship between cognitive abilities and choice in the games. However, as detailed in Section 3 below, in our experiment, we do not observe participants' true cognitive abilities and use a test score as a noisy measurement. We denote this score with $x \in [\underline{x}, \overline{x}]$. Moreover, we denote with $\mu(\theta)$ the objective prior distribution of cognitive abilities in the population, with $p(x|\theta)$ the conditional distribution of scores given the true cognitive ability, and with $\hat{\mu}(\theta|x)$ the Bayesian posterior distribution of cognitive abilities given score x.

From Assumptions 1 and 2, there is a function $g(\cdot)$ that, for any $x \in [\underline{x}, \overline{x}]$ assigns a probability that a player with score x understands that there is a dominant action in each game and chooses that action:

$$g(x) \equiv \int_{\underline{\theta}}^{\overline{\theta}} \hat{\mu}(\tau|x) f(\tau) d\tau. \tag{1}$$

where

$$\hat{\mu}(\theta|x) \equiv \frac{p(x|\theta)\mu(\theta)}{\int_{\Theta} p(x|\tau) d\mu(\tau)}. \tag{2}$$

We note that, under Assumption 3 and reasonable assumptions on $p(x|\theta)$ (e.g., that it satisfies the Monotone Likelihood Ratio Property), $g(x)$ is increasing in x. This leads to our first testable hypothesis.

Hypothesis 1 *A participant with a greater test score is more likely to choose the dominant action in each game.*

Similarly, in our experiment, we manipulate participants' beliefs about their opponent's cognitive abilities informing them about the opponent's test score (and explaining what the test is supposed to measure). To derive a testable hypothesis on players' preferences over games that relies on knowledge of the opponent's score (as opposed to the opponent's true cognitive ability), we replace Assumption 4 with the assumption below.

Assumption 5 *If a player understands the game (in the sense of Assumption 1), (a) he believes there is a function $h(x)$ that assigns to each opponent with score x a probability that the opponent understands the game, and (b) he believes that $h(x)$ is continuous and strictly increasing in x.*

While $g(x)$ in equation (1) is an objective function computed by the econometrician, $h(x)$ is a subjective belief about the opponent. With additional assumptions on the cognition of the individual forming this belief, the two functions coincide, and, in this case, participants hold correct beliefs about others' behavior.[10] At the same time, it is not necessary that the subjective $h(\cdot)$ function coincides with the objective $g(\cdot)$ function or that it is the same for all players. The essence of Assumption 5 is simple: players are required to understand that test scores are informative about the true cognitive ability and that players with a greater score are more likely to understand the game and to choose its dominant action, but are not required to hold correct beliefs. We believe that this is a plausible assumption, especially insofar as it only applies to the subset of players who are sufficiently sophisticated to understand the game themselves. This leads to our second testable hypothesis:

Hypothesis 2 *A participant with a greater test score is more likely to choose the HG if and only if he is paired with an opponent with a sufficiently high test score.*

3 Experimental Design

Our experimental design and hypotheses were pre-registered on the Open Science Framework (OSF).[11] We recruited a total of 701 participants from the database of volunteers managed by Prolific. Care was taken to ensure that the sample was demographically representative of the United Kingdom's population with respect to age, gender, and ethnicity. This demographic

[10]This requires knowledge of $\mu(\theta)$, $p(x|\theta)$, and $f(\theta)$, and the ability to perform Bayesian updating.

[11]See https://osf.io/mjpsu/?view_only=13201a5b01324d969a030279506c3d88.

representation strengthens the external validity of our findings and enables us to infer more general conclusions.

Each subject participated in two separate sessions, conducted one day apart. In the first session, we measured participants' cognitive abilities with three tasks. First, participants had up to 15 minutes to complete 18 Raven's Advanced Progressive Matrices (Part 1).[12] Raven's matrices are commonly used to evaluate fluid intelligence, that is, the ability to engage in non-verbal abstract reasoning and to solve a problem independently of previously acquired knowledge. This was followed by a 6-question numerical sequences task from the UK Household Longitudinal Study (Part 2), testing again fluid intelligence through numerical reasoning and pattern recognition. The first session ended with a 6-question Cognitive Reflection Test (Part 3) aimed at assessing participants' capacity to override an intuitive response with one that requires more deliberation. We provide more details of these tests in the next section, where we illustrate our main descriptive statistics. As pre-registered, our preferred proxies of cognitive abilities are the Raven's Advanced Progressive Matrices and the Cognitive Reflection Test. We included the numerical sequences task to facilitate comparison with the UK Household Longitudinal Study.

The second session had three parts. In Part 4, participants played five rounds of a one-shot Prisoner's Dilemma (PD) and one round of a one-shot Harmony Game (HG). In each of these rounds, participants were matched with a different participant from a previous study. Since the PD represents the status quo institution citizens have earned experience with, participants were given feedback on the opponent's action and the corresponding earnings after each round of the 5 PDs. On the other hand, since the HG is the untried reform, subjects are not given any feedback after the single round of the HG.[13]

In Part 5, participants played two additional rounds. In each round, they were offered

[12]The Raven's Advanced Progressive Matrices set includes 36 matrices of increasing degree of difficulty. We used the odd-numbered ones.

[13]Note that, in DBDBE, participants only play either 5 rounds of the PD or 5 rounds of the HG with feedback before being asked to choose between institutions. We decided to have our subjects play at least one round of both games to measure their propensity to identify and choose the dominant action in both the PD and the HG, as this is an important element of our theoretical framework.

the choice between playing a PD or an HG with an opponent from a previous study (different from any opponent met in any previous round). Before choosing the game, subjects were informed of their opponent's score in the Raven's matrices task from Part 1. We used a within-subject manipulation: each subject met exactly one opponent with a low score (i.e., 3 Raven's matrices out of 18 solved correctly) and exactly one opponent with a high score (i.e., 15 Raven's matrices out of 18 solved correctly), in random order. Subjects also learned that participants' performance ranged from 0 to 18 matrices solved correctly with an average of 9. In the instructions for Part 5, subjects were told that the opponent's performance in Part 1 could be regarded as a measure of "general intelligence and abstract reasoning" (see Figure A.3 in Appendix A.2). After choosing the game, subjects chose their action in the game selected for that round but did not get any feedback on the opponent's choice and on their earnings until the end of Part 5.

Finally, in Part 6, we collected some socio-demographic data and information on participants' political preferences. This segment included open-ended questions (which we report below) designed to gauge participants' opinions on the potential consequences of real-world untried policies, providing insight into their ability to evaluate and predict policy outcomes. Appendix A.1 reports an overview of the timing of both sessions.

Following the pre-registered design, subjects who failed an attention check at the beginning of the first session were not invited to the second session and excluded from the dataset; subjects who failed a comprehension quiz at the beginning of the second session were excluded from the survey and the dataset. The median completion time for the first session was 22 minutes, and participants were rewarded with £3 for completing it. To reduce cheating and to keep our measurements comparable with those from the UK Household Longitudinal Study and other studies meant to gauge IQ in representative samples of the UK, the tasks measuring cognitive abilities were not incentivized with monetary rewards. Subjects were told that they would learn their score alongside information about the other participants' performance during the second session (and they learned their scores at the end of the study). The median completion time for the second session was 16 minutes. Participants were rewarded with £2 for completing it plus a bonus determined by a random round from Part 4 and a random round from Part 5. The bonus ranged between £0.66 and £11.33 averaging around £4.

Matching Protocol. Since the experiment was conducted online (as a means to reach a large sample representative of the UK general population), we faced some constraints when implementing games, especially given the need for our subjects to be matched with 8 different opponents and to earn experience with the PD(thus, requiring feedback on others' behavior in this game) before choosing between institutions. Here, we detail how we solved this issue.[14] In Part 4, our participants were matched with 6 different participants in the laboratory experiments conducted at Brown and UC Berkeley by DBDBE. This means that our participants' opponents played the Prisoner's Dilemmans and the Harmony Game simultaneously with other participants in the same laboratory experiment.[15]

In Part 5, our subjects were matched with participants who were recruited from the same subject pool, who participated in an identical first session measuring cognitive abilities, and who answered correctly to either 3 (low cognitive abilities opponents) or 15 (high cognitive abilities opponents) Raven's matrices in Part 1 of the first session. These opponents participated in a different version of the second session (that took place before the standard version of the second session). In this session, subjects played 5 rounds of the PD with feedback and 1 round of the HG without feedback against participants from a previous study (as in Part 4 of the standard version of the second session). Then, they played an additional round of the PD and an additional round of the HG (in random order) against participants from a previous study (the laboratory experiments from DBDBE) without knowing the opponent's

[14]Our matching protocol is similar to the "replacement method" used in Alaoui and Penta (2016) and Alaoui, Janezic and Penta (2020).

[15]For the PDs in rounds 1 through 5, we matched our subjects with participants to the corresponding round in Part 1 of the Control, Random Dictator, Majority Once, and Majority Repeated treatments in DBDBE. For the HG in round 6, we matched our subjects with participants to round 1 in Part 1 of the Reverse Control and Reverse Random Dictator treatments in DBDBE. This ensures that our subjects and their opponents have the same degree of experience with each game.

cognitive abilities and without feedback on the opponent's choice. This guarantees that our subjects and their opponents had the same degree of experience with the PD and the HG before Part 5 and allows our subjects to choose unilaterally what game to play in the two rounds of Part 5.

This matching protocol was clearly explained to subjects (see screenshots in Appendix A.2). We note that, beyond enabling online implementation, this design allows for a cleaner identification of the effect of cognitive abilities by minimizing the potential confounding influence of social preferences: participants' actions (and their opponents' actions) affect only their own earnings, as others played the same game with different opponents, and their earnings were determined solely by their own and their opponents' actions. This is an important feature of our design as it reduces the chance that game payoffs misrepresent subjects' preferences for outcomes and thus allows us a tighter control on what actions are dominant or dominated. This is an important element of our theoretical framework, which posits greater cognitive abilities lead to a greater chance of choosing dominant actions. This design also allows us to measure preferences between the two games without any confounding role of signaling or selection motives since each participant decides unilaterally what games to play in Part 5 and their opponents in these games did not choose to participate in one game or the other.

4 Results

4.1 Descriptive Statistics

Socio-Demographics. Table 2 shows descriptive statistics for our sample. Prolific sampled participants to be representative of UK population, in what follows we verify to which extend this corresponds to reality. Participant's average age is 47.2 years, close to the median of the UK adult population (49), as reported by the Office for National Statistics (ONS) in 2022. The gender composition, with a female proportion of 51.1%, mirrors the UK general population according to the 2021 Census data. Socioeconomically, our participants report a higher average household income before taxes than the UK median, which was £38,100 in 2022 (ONS data); and they declare greater educational attainments: 67% holds at least a higher education degree compared to 40.6% in the UK adult population, as reported by Eurostat in 2022. A possible explanation is that our sample more closely represents the UK population with internet access.[16] This group tends to be more educated and have higher incomes, which may partly account for these differences. It is worth noting that this sample is considerably more representative of the general population in these characteristics compared to traditional laboratory experiments.

Finally, we note that in our sample there are only 8 respondents declaring to have only Primary Education. For this reason, in the analyses below, we pool together respondents with Primary and Secondary Education and consider this category as the baseline.[17]

Table 2: **Individual Characteristics**

Variable	Mean	Std. Dev.	Min.	Max.	N
Age	47.257	15.64	20	84	701
Female	0.511	0.5	0	1	701
Raven	9.016	3.313	0	18	701
CRT	3.25	1.702	0	6	701
Numerical Series	541.422	22.966	427.1	579.6	701
Household Income	50.424	33.803	0	150	701
Post-Graduate Degree	0.227	0.419	0	1	701
Graduate Degree	0.441	0.497	0	1	701
Further Secondary	0.17	0.376	0	1	701
Secondary	0.151	0.359	0	1	701
Primary	0.011	0.106	0	1	701

Cognitive Abilities. As stated in the pre-registration, we consider two indicators of cognitive abilities: the Raven Advanced Progressive Matrices (Raven henceforth) and the Cog-

[16] According to Ofcom (2021), approximately 6% of UK households did not have internet access in 2021.

[17] Considering them separately and changing the baseline does not change qualitatively our results. The category Further Secondary indicates the second part of secondary education studies that in the UK is separated from the first part because it is more specific to the choices planned by a student after secondary education.

nitive Reflection Test (CRT henceforth). Additionally, as shown in Table 2, we collected data on a Numerical Series test and Educational Attainments. According to the pre-registration, we included the Numerical Series test to enable comparisons with the UK Household Longitudinal Study (commonly referred to as "Understanding Society"), a large, representative socioeconomic survey of the UK population that includes political preferences.[18] Meanwhile, Educational Attainments were introduced primarily as a control variable. Given the nature of our study, it is essential to clarify what the Raven and CRT and scores specifically measure, describe these data, and examine their relationship. This is the focus of what follows in this section.

The Raven score measures fluid intelligence, abstract reasoning and the ability to solve new problems without drawing on prior knowledge (Cattell, 1963). This score is derived from logic-based graphical puzzles and is not contingent on numerical or verbal aptitude.[19] To gauge how our sample fares in terms of cognitive abilities with respect to the UK general population, we computed our participants' IQ using their Raven score. We benchmarked these scores against the UK general population using the 2022 age distribution (ONS data) and age-specific performance data from the user manual for Raven's matrices (Raven, 2003). The benchmarks are based on the completion of 36 matrices without a time limit.[20] To avoid excessively long participation times, we set a 15-minute time limit and reduced the number of matrices to 18, selecting only the odd-numbered ones. To estimate the IQ, we then doubled our participants' Raven scores based on the assumption that solving one matrix implies the ability to solve the next. This method may slightly inflate the scores. Nonetheless, the 15-minute (or 900 seconds) time limit, with subjects averaging 672.33 seconds and a median of 690.5 seconds, likely mitigated any overestimation. As it is customary, for calculating the

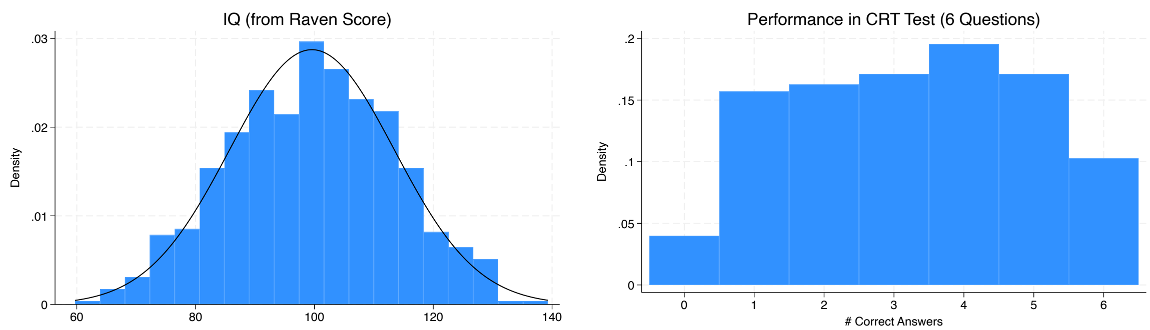
[18]As discussed in Section 6, the score in the Numerical Series test is strongly, although not perfectly, correlated with the Raven score and our results remain qualitatively unchanged when using the Numerical Series score instead of the Raven score.

[19]For this reason, Raven is usually preferred to, for example, the Numerical Series test, which inherently depends on numeracy levels (Fisher et al., 2013).

[20]Using the scores from the Raven's matrices user manual and the age distribution, we estimated an average of 18.23 matrices solved correctly with a standard deviation of 7.6.

IQ, we normalized the UK population mean to 100 and its standard deviation to 15.[21] We display the IQ distribution of our sample in the left panel of Figure 1. Reassuringly, the median IQ in our sample is 100, and approximately 80% of our subjects scored between 82 and 116, with 5% exceeding a score of 124, paralleling the IQ distribution expected in the UK general population.

Figure 1: **Cognitive Abilities' Distributions.**



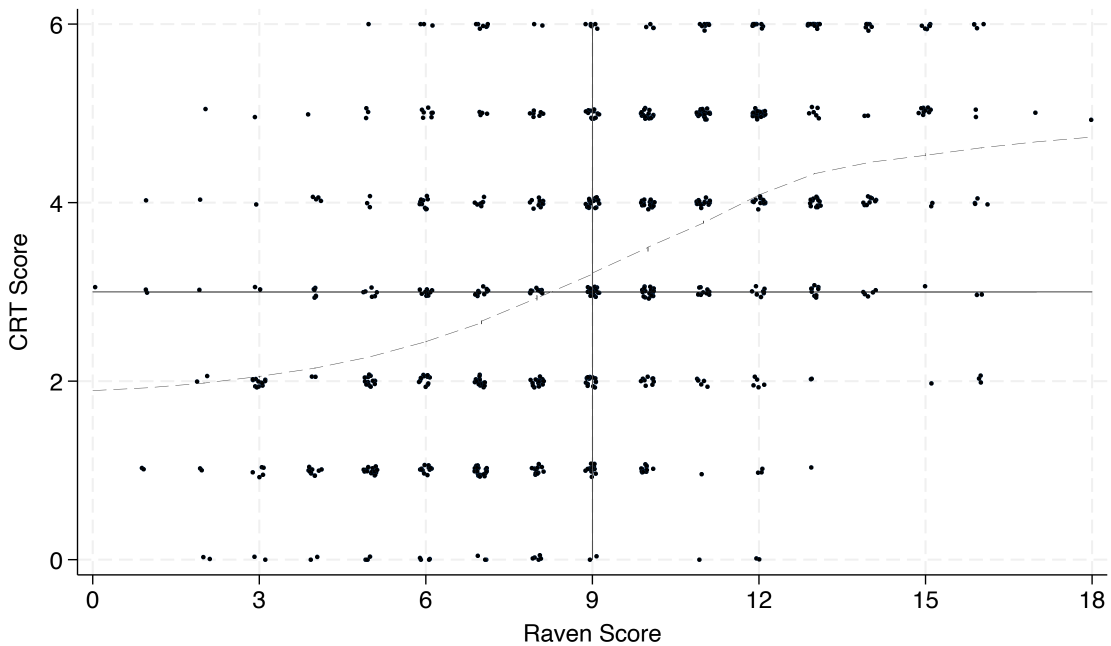
On the other hand, the CRT measures an individual's ability to suppress instinctive but incorrect responses and engage in reflective thought (Frederick, 2005). Unlike the Raven score, the CRT score arguably comprises two distinct components: it evaluates the ability to solve rather simple abstract problems (generally simpler than the ones presented in the Raven test), and it assesses cognitive control, that is, the tendency to think carefully and deliberately. The right panel of Figure 1 displays the distribution of CRT scores in our sample.

Table A.1 in the Appendix reports the correlation between the different measures of cognitive abilities and the correlation between these measures and educational attainment, which we use in some robustness analyses below. Figure 2 is a 'jittered' scatter plot illustrating the Raven and CRT scores for each subject. These scores are significantly correlated.[22] However,

[21]Therefore, we calculated the IQ using this formula $IQ = 2 * RavenScore - 18.23) * (15/7.16) + 100$, where 18.23 and 7.16 are the estimated average and standard deviation within the UK general population.

[22]The correlation between Raven and CRT scores in our sample is 0.472 (0.402 with a 6 items version of Raven). This exceeds the commonly reported correlation of approximately 0.3 (e.g., Willadsen et al., 2024). This divergence could be due to our employment of an extended 6-question CRT, in contrast to the traditional 3-question version.

Figure 2: **Relationship Between Raven and CRT Scores:** Each point represents the Raven and the CRT score of a subject. The location of each point has been perturbed (or "jittered") using random noise for the sake of visualization. The two crossing lines parallel to the axes represent the median of the two scores, the dashed line is the lowess smoothed line.



it is important to highlight the presence of independent variation. A non-negligible number of subjects perform above the median on the CRT but below the median on the Raven test, and vice versa. Specifically, subjects in the first quadrant of the figure demonstrate relatively high cognitive control but low problem-solving capacity, while those in the fourth quadrant (a less frequent outcome) exhibit high problem-solving capacity but relatively low cognitive control. The dashed line ('lowess') indicates a positively sloped but nonlinear relationship between the two measures. Notably, at the extremes of the Raven score (say, less than 4 and greater than 11), the CRT score increases very gradually. This again shows the existence of subjects with low problem-solving capacity but high cognitive control, as well as those with high problem-solving capacity but relatively low cognitive control.

4.2 Behavior in Games

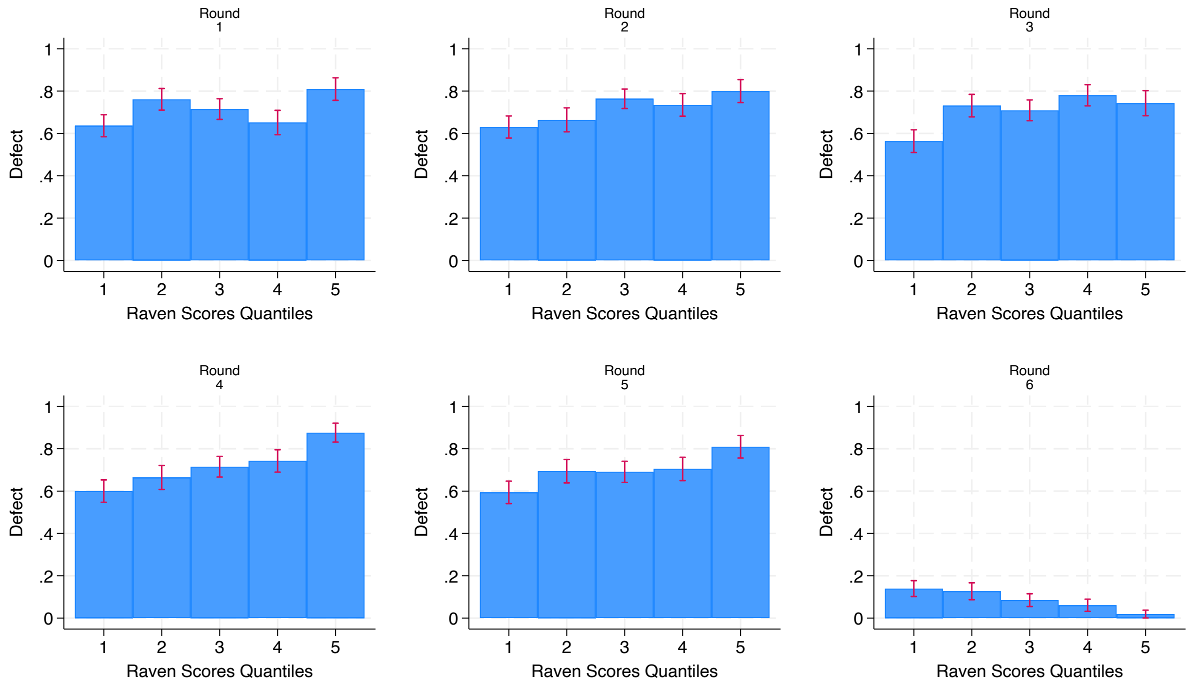
Our analysis begins by testing Hypothesis 1, which predicts that subjects with greater cognitive abilities are more inclined to choose dominant actions. In Table 3, we report the estimates of a Logit model that examines the correlation between a subject's own cognitive abilities and the choice to Defect, the dominant action in the PD and the dominated action in the HG. In the five rounds of the PD, there is a positive correlation between the Raven score and the selection of the dominant action, Defect, with this correlation appearing to strengthen over time (increasing from about 3% to 9% per unit standard deviation in round 4 and decreasing to about 7% in round 5). When both measures of cognitive abilities are included, as in this table, the CRT score does not show a significant effect.

Table 3: **Effect of Own Cognitive Abilities on Behavior in Games:** The binary dependent variable is the action Defect. Raven and CRT scores are standardized (that is, transformed to have mean 0 and standard deviation 1). Socio-Demographics controls include: Age, Gender, Education and Income. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Raven Score (std)	0.0371*	0.0595***	0.0721***	0.0968***	0.0725***	–0.0254**
	(0.0605)	(0.0019)	(0.0003)	(0.0000)	(0.0002)	(0.0269)
CRT Score (std)	–0.0136	–0.0348*	0.0147	–0.0162	–0.0228	–0.0369***
	(0.5026)	(0.0804)	(0.4617)	(0.4112)	(0.2671)	(0.0039)
Socio-Demographics	Yes	Yes	Yes	Yes	Yes	Yes
N	701	701	701	701	701	701

A similar trend is evident in the single round of the HG: subjects with a greater Raven score are more likely to choose the dominant action, Cooperate. Moreover, in this second game, the coefficient of the CRT score is statistically significant, and its magnitude is larger than the magnitude of the Raven score coefficient (noting that both measures are expressed in standard deviations). This result has a natural interpretation: subjects that suppress an instinctive answer and think before making a decision are able to form a better understanding of a novel strategic interaction, that is, a game they play for the first time. Figure 3 visually represents these findings. This discussion can be summarized with the following insight:

Figure 3: **Raven Scores and Behavior in Games:** Defection rates by Raven score quantiles in the 5 rounds of PD and in the sole round of HG from Part 4. The red bar represents the 95% confidence interval.



Result 1. *Participants with higher cognitive abilities are more adept at identifying the dominant action through a process of learning in the PD, and through introspection and reflection in the HG.*

Additional Analyses and Robustness Checks. In Section B of the Appendix, we provide supplementary analyses. Table A.2 shows the coefficients that are hidden in Table 3. Table A.3 reports the estimates for similar regressions without the socioeconomic and demographic controls. In Table A.4, we use a coarser Raven score obtained using only 6 matrices which is more comparable with the score obtained with the 6-items CRT.[23] Here, the Raven score coefficients become smaller, as is to be expected given the noisier measurement, but the results are qualitatively identical to those from Table 3. In Table A.5, we use dummies characterizing the joint distribution of the CRT and Raven scores instead of their respective

[23]These are the matrices numbered 1, 7, 13, 19, 25, and 31, selected to be evenly spaced.

scores, suggesting an independent effect of the CRT in the PDs. Table A.6 shows that the CRT score is typically positively correlated with choosing to Defect in the PD and negatively correlated with choosing to Defect in the HG when it is the sole measure of cognitive skills in the regression. For completeness, Table A.7 reports the estimates of similar regressions including only the Raven score.

4.3 Preferences over Games (or Policies)

Moving on to preferences between games (or policies), we specifically examine Hypothesis 2, which asserts that a preference for the HG over the PD occurs only among players who have sufficiently high cognitive abilities and believe their opponent to be sufficiently capable.

Table 4 shows that a participant's Raven score is positively correlated with the choice of HG, increasing the probability by about 6% per standard deviation when paired with a high-cognitive-abilities opponent, as shown in column 1. However, this correlation is absent when facing an opponent with low cognitive abilities, as indicated in column 2. Columns 3 and 4, which consider both treatments together, demonstrate that the interaction between Raven score and the treatment is both positive and significant.

Figure 4, where we group our sample into five quantiles based on their Raven score and calculate the proportion of participants who choose the HG over the PD for each quantile and for each treatment (i.e., high Raven score opponent and low Raven score opponent), confirms the finding from Table 4 and highlights additional details. Participants in the highest Raven score quantile exhibit a stronger preference for the HG when paired with a high-cognitive-skills opponent. On the contrary, participants in the lowest quantiles are more likely to choose the HG when paired with a low-cognitive-skills opponent.[24] Finally, we note that there seems to be a non-monotonic relationship between cognitive skills and preferences over games, with subjects at the two extremes of the distribution showing a stronger preference

[24]This is reminiscent of the behavior of subjects with low cognitive abilities in Gill and Prowse (2016) where, in a beauty contest, they guess a lower number when matched with other low ability subjects compared to when they are matched with high ability subjects.

Table 4: **Effect of Raven Score, CRT score, Education and Opponent's Cognitive Abilities on Preferences over Games:** The binary dependent variable is choice of HG in Part 5. Raven and CRT scores are standardized. The omitted education category is secondary or primary education. Socio-Demographics controls include: Age, Gender, and Income. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors (clustered at the individual level in columns 3 and 4); $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

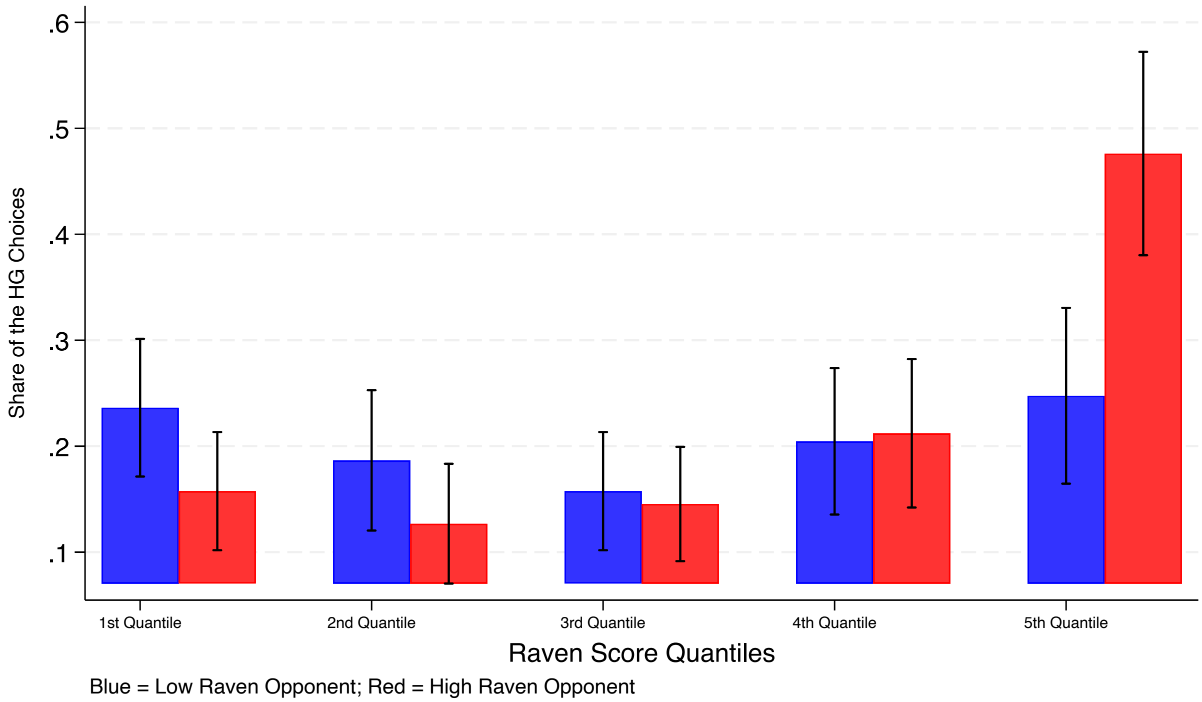
	High Rav.Opp. Vote HG b/p	Low Rav.Opp. Vote HG b/p	All Vote HG b/p	All Vote HG b/p
Raven Score (std)	0.06132***	–0.01780	–0.02133	–0.01678
	(0.0007)	(0.3119)	(0.2085)	(0.3279)
CRT Score (std)	0.01767	0.02834*	0.02566	0.02890*
	(0.3148)	(0.0938)	(0.1152)	(0.0785)
High Raven Opp.			–0.01141	–0.11907**
			(0.5553)	(0.0192)
High Raven Opp.× Raven Score			0.09257***	0.08318***
			(0.0000)	(0.0003)
High Raven Opp.× CRT Score			–0.00600	–0.01312
			(0.7758)	(0.5379)
High Raven Opp.× Graduate				0.12883**
				(0.0285)
High Raven Opp.× Post-Graduate				0.14930**
				(0.0204)
High Raven Opp.× Further Secondary				0.08923
				(0.2514)
Post-Graduate	0.09033	–0.06328	0.00963	–0.05642
	(0.1014)	(0.1874)	(0.8051)	(0.2301)
Graduate	0.05351	–0.07757*	–0.02154	–0.07643*
	(0.2901)	(0.0723)	(0.5433)	(0.0718)
Further Secondary	–0.02796	–0.11781**	–0.07852*	–0.11361**
	(0.6637)	(0.0283)	(0.0725)	(0.0322)
Socio-Demographics	Yes	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	701	701	1402	1402

for the HG than those in the middle.[25] Figure 5 delves deeper into the relationship between cognitive abilities and preferences between games: subjects in the highest quantile are those who (a) more often choose the HG only when paired with a high ability opponent and (b) less often choose always the PD, regardless of the opponent's cognitive abilities.

Table 4 also reveals interesting correlations with the educational variables. In particular,

[25]This non-monotonicity is investigated with a parametric statistical test in Table A.16 in the Appendix. This is a non-preregistered analysis and, since we did not design the study to explore this relationship, we lack the statistical power for a non-parametric test.

Figure 4: **Raven Scores and Preferences over Games by Treatment:** Each rectangle of the histogram represents the share of participants choosing the HG by quantiles of Raven score in each of the two treatments (in blue and red). The vertical bars represent the 95% confidence intervals.

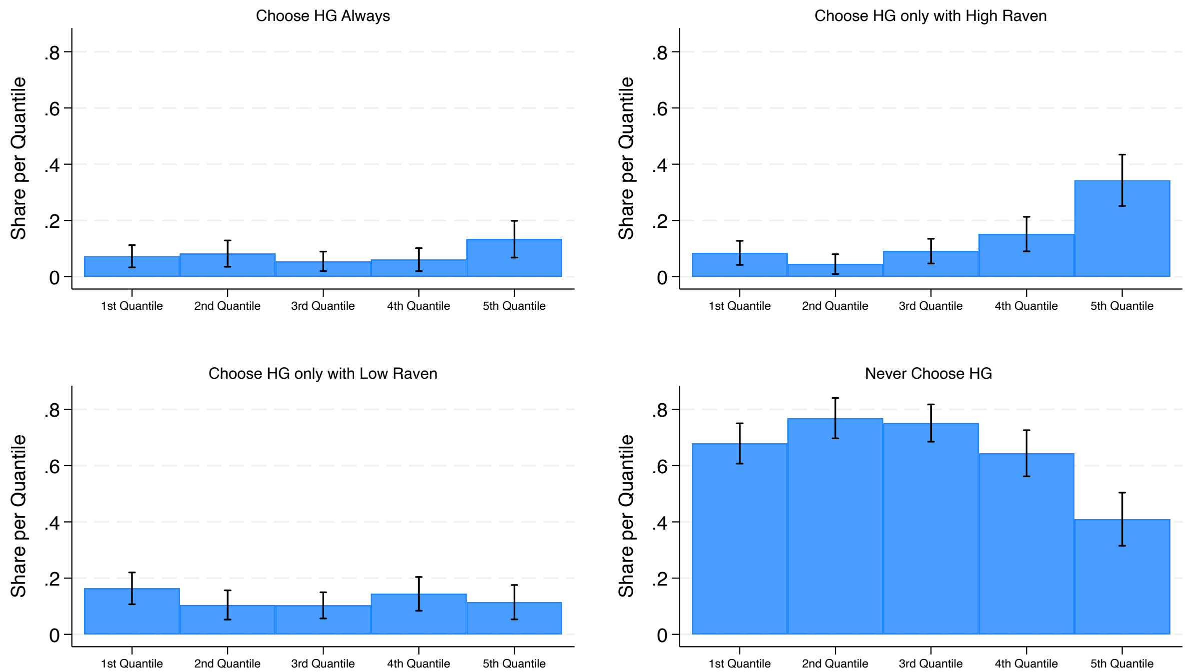


column 4 shows that the treatment (that is, the opponent's Raven score) has a different effect on participants' behavior depending on their education: participants with a graduate or post-graduate degree choose the HG at greater rates than participants with fewer years of education when facing a high ability opponent. This suggests that the participants' education has an effect similar to their Raven score and independent from it in terms of predicting behavior.[26] This effect is interesting and policy relevant: a possibility is that education proxies or develops cognitive abilities not measured by the Raven test or the CRT; another possibility is that education can train individuals to think about indirect effects.

Table 5 provides a more detailed examination of the combined effect of our two measures of cognitive abilities. In these regressions, we replace the raw Raven and CRT scores with categorical indicators for their joint distribution as the main independent variables. Column

[26]For more details on the UK educational system and how to interpret these variables, see footnote 17.

Figure 5: **Raven Scores and Opponent-Dependent Preferences over Games**



1 reveals evidence of complementarity between Raven and CRT scores: only participants who score at or above the median on both tests show a preference for the HG when paired with a high-Raven opponent. Consistently, column 3 indicates that a significant difference in behavior between high-Raven and low-Raven opponents is observed exclusively among subjects who perform well on both tests. These results suggest that, while the Raven score emerges as the stronger determinant in a direct comparison, the CRT score also exerts an independent influence. These observations lead us to the following conclusion:

Result 2. *Participants with greater cognitive abilities are more inclined to select the Harmony Game over the Prisoner's Dilemma when their opponent is perceived as being better cognitively able.*

Additional Analyses and Robustness Checks. Table A.9 shows the estimates of models similar to the ones from Table 4 without socioeconomic and demographic controls. Table A.10 analyzes the determinants of choice between games without interactions. Table A.11

Table 5: **Joint Effects of Raven and CRT Scores and Opponent's Cognitive Abilities on Preferences over Games:** The binary dependent variable is choice of HG in Part 5. The baseline (omitted) is Low Raven and Low CRT scores, defined as scores below the respective medians. The omitted education category is secondary or primary education. Socio-Demographics controls include: Age, Gender, and Income. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors (clustered at the individual level in columns 3 and 4); $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	High Rav.Opp.	Low Rav.Opp.	All
	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p
Low Raven, High CRT	–0.01508	0.02714	0.01127
	(0.7421)	(0.5853)	(0.8211)
High Raven, Low CRT	–0.02231	–0.07577	–0.09113*
	(0.6534)	(0.1306)	(0.0766)
High Raven, High CRT	0.09003**	–0.00923	–0.02573
	(0.0205)	(0.8160)	(0.5114)
High Raven Opp.			–0.07006*
			(0.0601)
High Raven Opp.× Low Raven, High CRT			–0.02551
			(0.6577)
High Raven Opp.× High Raven, Low CRT			0.08142
			(0.3311)
High Raven Opp.× High Raven, High CRT			0.13193***
			(0.0039)
Socio-Demographics	Yes	Yes	Yes
Education	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes
N	701	701	1402

shows that the Raven score is the only cognitive ability measure whose interaction with the treatment is statistically significant also when performing this analysis with a coarser Raven score measure, calculated using only 6 matrices and thus more comparable with the CRT score. Table A.12 shows that results are unchanged if we exclude the 45 participants who declared taking a game theory course. Table A.14 demonstrates that our findings remain robust when using only the Raven score as the measure of cognitive ability. Finally, Table A.15 shows that, when included as the sole measure of cognitive ability, the CRT score is a positive predictor of choosing the HG, with a heterogeneous treatment effect: the treatment is more pronounced among participants with higher CRT scores.

4.4 Optimal Policy Based on Observed Behavior

As discussed in Section 2, when individuals are fully rational and rationality is common knowledge, the optimal behavior is to prefer the HG. However, does this preference still hold when players are boundedly rational? In what follows, we address this question empirically. We calculated the expected payoffs using the actual frequencies of choices made by our subjects in each game, as presented in Table A.17 in the Appendix. These frequencies were then treated as probabilities to compute the expected payoffs for both actions in both games, which we report in Tables 6 and 7.

Table 6 indicates that when facing a low-cognitive-skills opponent, the strategy that maximizes the expected payoff is to choose the PD, and then Defect in the resulting game. Conversely, Table 7 shows that when playing against a high-cognitive-skills opponent, the optimal strategy is to choose the HG and then Cooperate. From these calculations, we infer that subjects with higher cognitive abilities tend to play more profitably than those with lower abilities.

Table 6: **Empirical Expected Payoffs with Low Raven Score Opponent**

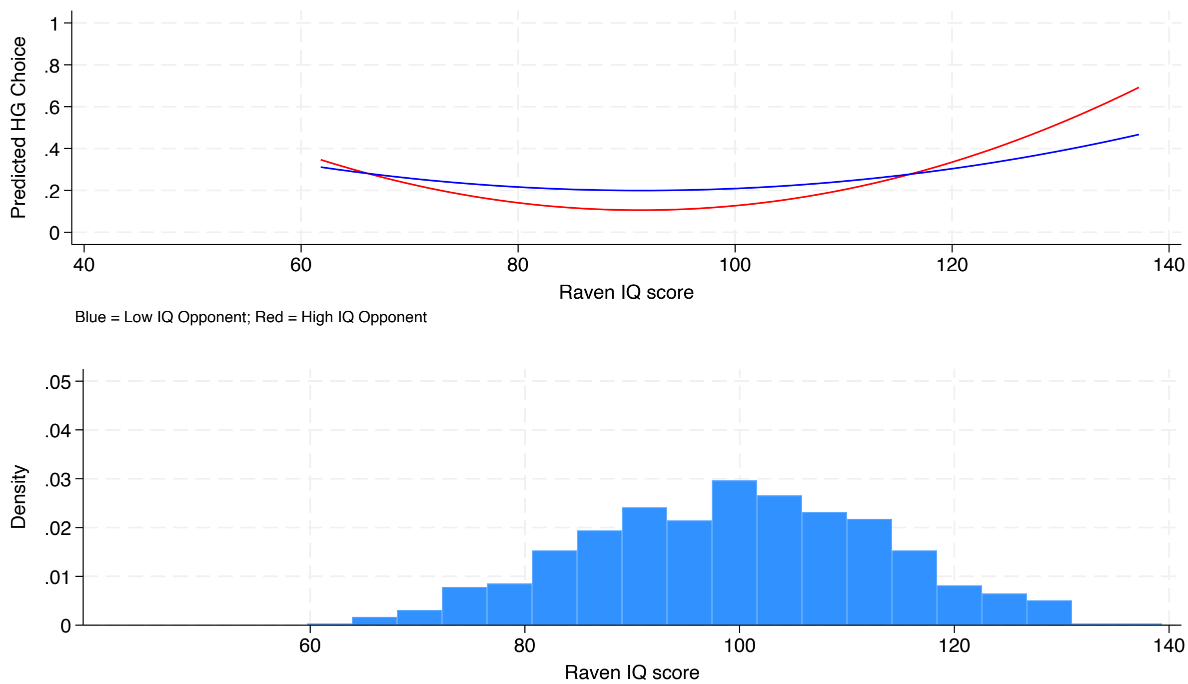
	Game	
Action in Game	**HG**	**PD**
C	7.22	6.39
D	6.22	8.39

Table 7: **Empirical Expected Payoffs with High Raven Score Opponent**

	Game	
Action in Game	**HG**	**PD**
C	8.00	4.58
D	7.00	6.58

In the top panel of Figure 6, we present the predicted probability of choosing the HG based on the IQ (calculated from Raven scores as described above) for each of the two treatments, using the quadratic models estimated in Table A.16 in the Appendix. The intersection of the two lines occurs at an IQ of approximately 115, which can be considered the threshold above which subjects are more likely to exhibit behavior that maximizes their payoffs. The bottom panel of Figure 6 shows the distribution of Raven-Based IQ among our subjects.

Figure 6: **Predicted Preference for Harmony Game** The two blue and red lines in the top panel represent the quadratic fit of choosing the Harmony Game against low and high IQ opponent respectively



5 Measure of Appreciation of Indirect Effects

To understand better the process leading to an appreciation of the indirect effects of policies, in Part 6 of our study, we presented participants with two open-ended questions aimed at eliciting their views on the potential direct and indirect consequences of policy reforms in the UK. We asked our subjects the following two questions:

- Minimum Wage: *Recently, the UK has increased the hourly gross minimum wage from 10.42 GBP to 11.44 GBP for 21-year-olds and over. What do you think will be the consequences in the UK? Please respond in full sentences.*

- Immigration: *In the UK, on 1 January 2021, there were 9.5 million foreigners, representing 14.4% of the UK population. What do you think would be the consequences of reducing the inflow of foreigners to the UK with the goal of reaching the European average? Please respond in full sentences.*

With help from ChatGPT and a Research Assistant, we compiled a list of all the indirect consequences of each reform mentioned by our respondents in their answers. For this purpose, we define a consequence as indirect if it involves economic agents changing their behavior in response to the new policy. The list, separately for the minimum wage and the immigration question, is available in Section C.2 of the Appendix. Endowed with these lists, we then manually coded the number of indirect consequences mentioned by each participant in each answer.

To give some examples, considere the minimum wage question. The answer "*Inflation is likely to stay high as prices are increased to cover the increase in wages. Some firms may be unable to stay in business and may close*" was attributed 2 indirect effects: "inflation" and "businesses closing" (respectively 2 and 9 in the list in Section C.2). On the other hand, we attributed zero indirect effects to the following answer: "*Those earning minimum wage will be slightly less poverty stricken, though I don't feel that this increase is enough. Wages across the board have fallen dramatically in real terms with only the rich actually making more than before.*" Considering the Immigration question, the answer "*Reduced diversity over time in country. Reduced dynamic labour force supply. Opportunities for some UK born people to re-enter the labour market where previously impossible. Increased inflation due to higher wage.*" was attributed 2 indirect effects: "UK nationals entering labor force, reduced reliance on government benefits, increased quality of public services" and "greater wages for workers, greater costs for entrepreneurs" (1 and 7 respectively).

In Table A.18 of the Appendix, we present descriptive statistics for the main variables of this textual analysis. On average, for the immigration question, respondents listed approximately 0.4 indirect consequences, whereas for the minimum wage question, they listed about 0.9 indirect consequences. The variable Right Leaning represents a self-assessment of political ideology on a scale from 1 to 10, with 1 indicating extreme left. The average score is 4.7, indicating a slight skew toward the left. In Table A.19 of the appendix, we present the cross correlations between the main variables involved in this analysis. We note in particular that the CRT score is significantly and positively correlated with the listed consequences, while the Raven score is not significantly correlated with Minimum Wage Consequences. Furthermore, we find that Right Leaning is significantly and negatively correlated with the Raven score. The correlation with CRT is negative but insignificant.[27]

We assumed that the number of untried policies' indirect effects identified by participants to be a measure of the respondents' understanding of indirect effects. We then used this measure as the dependent variable in the regression analyses shown in Table 8. The results indicate that while the CRT score is a significant positive predictor of the number of identified indirect effects, the Raven score is not. As in the experimental portion of the study, cognitive abilities are associated with understanding of indirect effects. However, in this second exercise, we find that it is the CRT score, rather than the Raven score, that emerges as a significant predictor.

This suggests that cognitive abilities contribute to understanding indirect effects in different ways across the two environments. In the first exercise, which focused on the abstract understanding of indirect effects, fluid intelligence—measured by the Raven score—was more relevant. Fluid intelligence involves skills such as pattern recognition and problem-solving in novel contexts, making it valuable for tasks that require abstract reasoning without specific contextual knowledge.

In contrast, in this second exercise, the CRT score is the key predictor. This finding

[27] Kahan (2013) finds a similar result.

Table 8: **Cognitive Abilities and Listed Indirect Effects:** The dependent variable is the number of indirect consequences listed in the open-ended questions in Part 6. Raven and CRT scores are standardized. The omitted education category is secondary or primary education. Robust Standard Errors are in brackets; * $p - value < 0.1$, ** $p - value < 0.05$, *** $p - value < 0.01$

	Baseline Immigration b/se	Baseline Minimum Wage b/se	Controls Immigration b/se	Controls Minimum Wage b/se
Raven Score (std)	–0.010	–0.043	–0.034	–0.023
	(0.0259)	(0.0347)	(0.0267)	(0.0363)
CRT Score (std)	0.128***	0.115***	0.127***	0.083**
	(0.0258)	(0.0360)	(0.0268)	(0.0380)
Age			–0.003*	0.006***
			(0.0015)	(0.0021)
Female			0.006	–0.082
			(0.0432)	(0.0597)
Income			0.001	–0.000
			(0.0007)	(0.0009)
Post-Graduate			0.117	–0.011
			(0.0716)	(0.1062)
Graduate			0.116*	0.122
			(0.0597)	(0.0875)
Further Secondary			0.002	–0.100
			(0.0669)	(0.0967)
Constant	0.441***	0.906***	0.463***	0.631***
	(0.0212)	(0.0301)	(0.1034)	(0.1480)
N	701	701	701	701

suggests that the task requires not only abstract reasoning but also the ability to override intuitive judgments, enabling individuals to consider and identify less obvious consequences. Here, the exercise focuses on the number of indirect effects listed, capturing both an individual's awareness of potential indirect consequences and their belief in these consequences' relevance within the current socio-economic context. Thus, the CRT score's significance likely reflects the importance of reflective thinking in recognizing and articulating indirect consequences that go beyond initial impressions.[28]

As shown in columns 1 and 2 of Table 9, introducing a measure of participants' political ideology (Right Leaning) does not alter the previous results, suggesting that the number

[28] In Table A.20, we show the regression with the Raven-CRT joint distribution dummies, which confirms the finding that the CRT score is a better predictor than the Raven score.

of indirect effects listed is not solely a reflection of one's personal views on the proposed reform. The coefficient for political ideology is statistically significant suggesting that an important factor influencing the ability to identify indirect consequences is motivated reasoning.[29] Right-leaning individuals list fewer indirect consequences for immigration reform, likely because they place greater emphasis on direct effects (e.g., changes in the number of foreigners in the country). On the other hand, they tend to list more indirect consequences of a minimum wage increase, suggesting that left-leaning individuals may focus more on the direct effects (e.g., higher wages for lower-income workers) while paying less attention to potential indirect effects (e.g., reduced labor demand or layoffs by employers). The coefficients for ideology are similar in magnitude to those for CRT, indicating that both effects are comparable in strength. In columns 3 and 4 of Table 9, we introduce vote intentions, and the effects observed can be interpreted similarly to those for ideology.

Result 3. *When prompted to describe the consequences of real-world (immigration and labor) policy reforms, participants with greater cognitive abilities are more likely to think of indirect effects.*

6 Additional Evidence on Cognitive Abilities and Political Attitudes from Understanding Society

In order to enhance the external validity of our result, in this Section, we analyze data from the UK Household Longitudinal Study (also known as Understanding Society) to present some descriptive statistics that are broadly consistent with the assumptions and predictions of the model outlined in Section 2. Cognitive ability is measured using an index of abstract problem solving skills based on completing numerical series. This measure is available in Wave 3 (January 2011–July 2013) of the Understanding Society dataset. We collected this

[29]This refers to the tendency of individuals to shape their assessments of information based on goals or ends that are extrinsic to accuracy (e.g. Kahan, 2013)

Table 9: **Cognitive Abilities, Political Preferences and Listed Indirect Effects:** The dependent variable is the number of indirect consequences listed in the open-ended questions in Part 6. Raven and CRT scores are standardized. Conservative is the omitted party. The omitted education category is secondary or primary education. Socio-Economics Controls include Age, Gender, Income and Education. Robust Standard Errors are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Ideology		Party Preferences	
	Immigration	Min.Wage	Immigration	Min.Wage
	b/se	b/se	b/se	b/se
Raven Score (std)	–0.038	–0.016	–0.038	–0.013
	(0.0267)	(0.0355)	(0.0269)	(0.0363)
CRT Score (std)	0.124***	0.087**	0.126***	0.084**
	(0.0265)	(0.0373)	(0.0265)	(0.0376)
Right Leaning (std)	–0.073***	0.106***		
	(0.0240)	(0.0303)		
Labour			0.125	–0.313**
			(0.0930)	(0.1266)
Lib Dem			0.110*	–0.230***
			(0.0592)	(0.0855)
SNP			–0.002	–0.054
			(0.0687)	(0.1112)
Other			0.133	–0.201*
			(0.0867)	(0.1203)
Constant	0.437***	0.670***	0.364***	0.850***
	(0.1050)	(0.1466)	(0.1174)	(0.1693)
Socio-Economics Controls	Yes	Yes	Yes	Yes
N	701	701	701	701

measure also in our experiment, with the preregistered aim of connecting our results with evidence from Understanding Society.[30]

Using the weighted data from Understanding Society, we estimated the mean and the standard deviation of the UK population to build an IQ measure, hereafter referred to as the NS-based IQ. Tables A.8 and A.13 demonstrate that, in our sample, the NS score is correlated with the probability of selecting the dominant action and the likelihood of preferring the Harmony Game, in a similar way as the Raven score.

Our model and test of political preference formation implies that political preferences and choices are the outcomes of a mental calculation. In Figure 7, we observe that the

[30] As mentioned earlier, this measure is conceptually similar to the Raven's Progressive Matrices test, although it also depends on numeracy levels. In Table A.1 of the Appendix, we observe that it is highly correlated with Raven scores and CRT scores.

average IQ of individuals who report voting based on rationally driven political evaluation is significantly higher than that of individuals who report voting without deliberate thought, such as always voting in a particular way, not knowing why they voted, or following others' advice. Notably, respondents who reported voting strategically exhibit the highest average IQ, which suggests how more intelligent voters are motivated by evaluations about other voters' behavior. This is perfectly in line with what it is stated in the hypothesis 2.

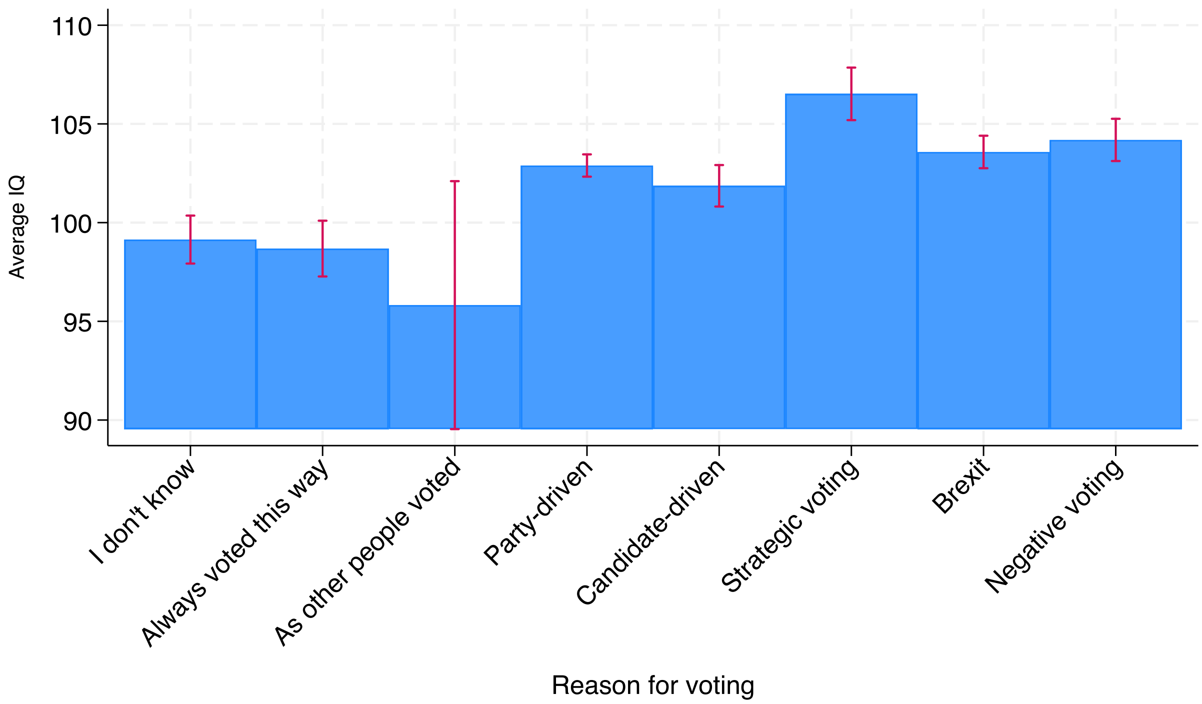


Figure 7: **Reason for Voting and Cognitive Abilities:** Responses to the question on voting reasoning in Understanding Society, Wave 11 (2019–2020). IQ is calculated using the numerical series score from Wave 3 (2011–2013). Averages are weighted using the survey's cross-sectional weights.

In the model, political preferences are influenced by people's ability to understand indirect effects, a capacity that can also be enhanced through external information. Figure 8 suggests that – assuming personal levels of information on politics are closely correlated with IQ – individuals generally have an accurate perception of their levels of political knowledge relative to others. Individuals with an average IQ (around 100) tend to believe they are as informed as the average person (neither agreeing nor disagreeing with the question asked), while those with above-average IQ perceive themselves as better informed, and those with below-average IQ believe they are less informed.

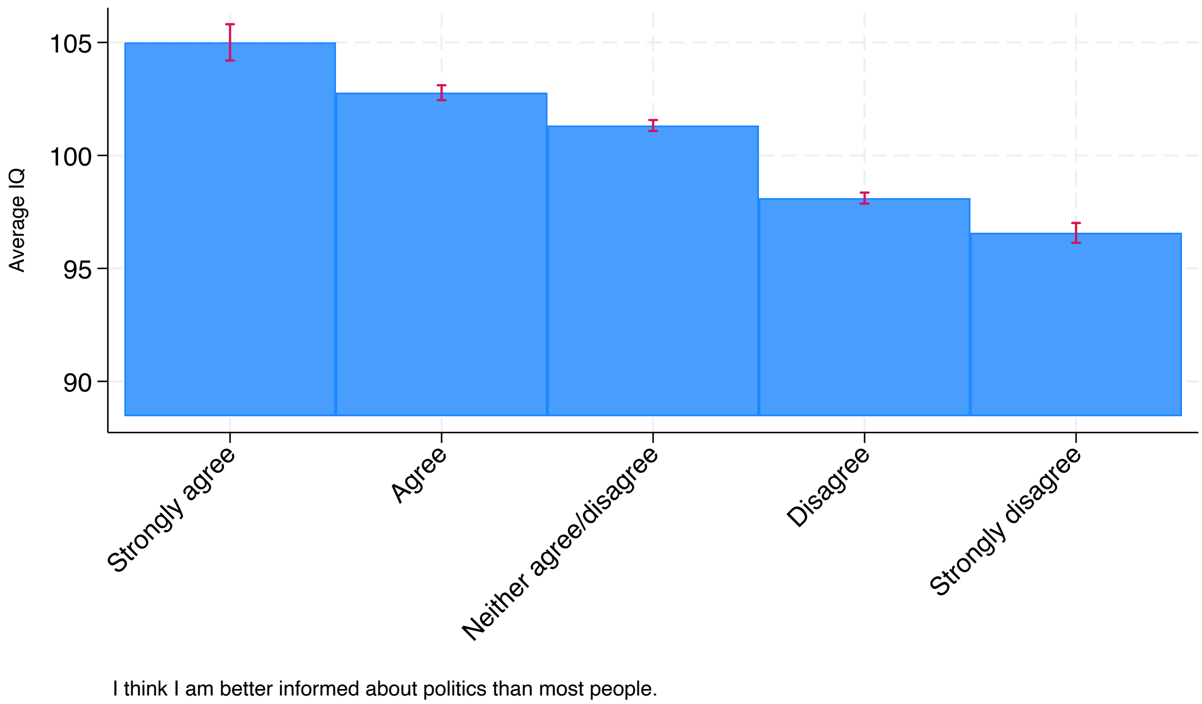


Figure 8: **Political Information and Cognitive Abilities:** Responses to the question: *"I am better informed about Politics than Most People"* in Understanding Society IQ is calculated using the numerical series score, both variables are from Wave 3 (2011–2013). Averages are weighted using the survey's cross-sectional weights.

Understanding indirect effects implies that individuals have nuanced opinions about policies, often resulting from implicit cost-benefit calculations. Consequently, we expect individuals with higher cognitive ability to provide moderate answers to key policy preferences questions. In the Understanding Society dataset, two key political preference questions are available: preferences on taxes versus public expenditures and preferences on UK integration with the EU.

In Figure 9, we show that respondents with higher IQ scores tend to position themselves between the extremes when asked about tax preferences. Similarly, Figure 10 illustrates a similar pattern for preferences on UK independence versus integration with the EU.

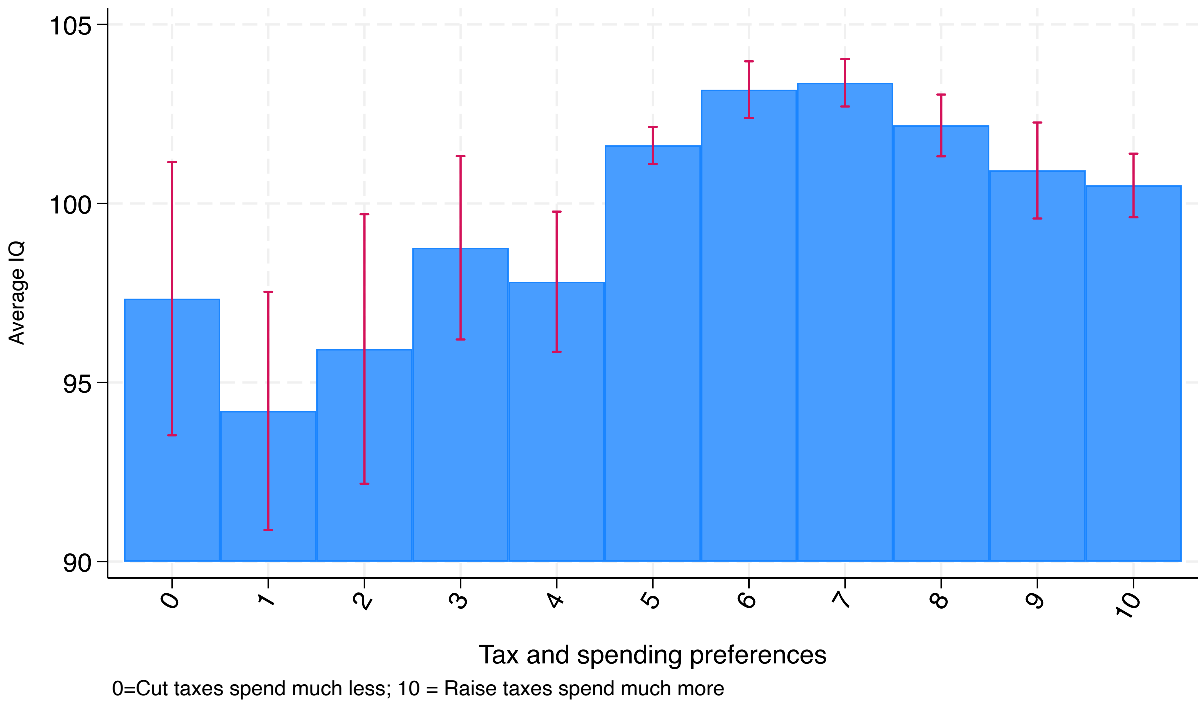


Figure 9: **Government Spending Preferences and Cognitive Abilities:** Responses to the question: *"On a 0 to 10 scale, where 0 means the government should cut taxes a lot and spend much less on health and social services, and 10 means the government should raise taxes a lot and spend much more on health and social services, where would you place yourself on this scale?"* in Understanding Society, Wave 11 (2019–2020). IQ is calculated using the numerical series score from Wave 3 (2011–2013). Averages are weighted using the survey's cross-sectional weights.

7 Conclusion

The research presented here explored the complex relationship between cognitive abilities and political preferences. Specifically, we examined individuals' capacity to understand the indirect effects of untested policies. Our model demonstrates that while personal cognitive abilities and an understanding of a policy's ultimate outcomes are necessary conditions for supporting welfare-enhancing reforms, they are not sufficient. We also hypothesized that the belief in the cognitive competence of one's peers is equally crucial, and it plays a pivotal role in fostering support for sound policies.

Our methodological approach was designed to capture the nuanced influence of beliefs on

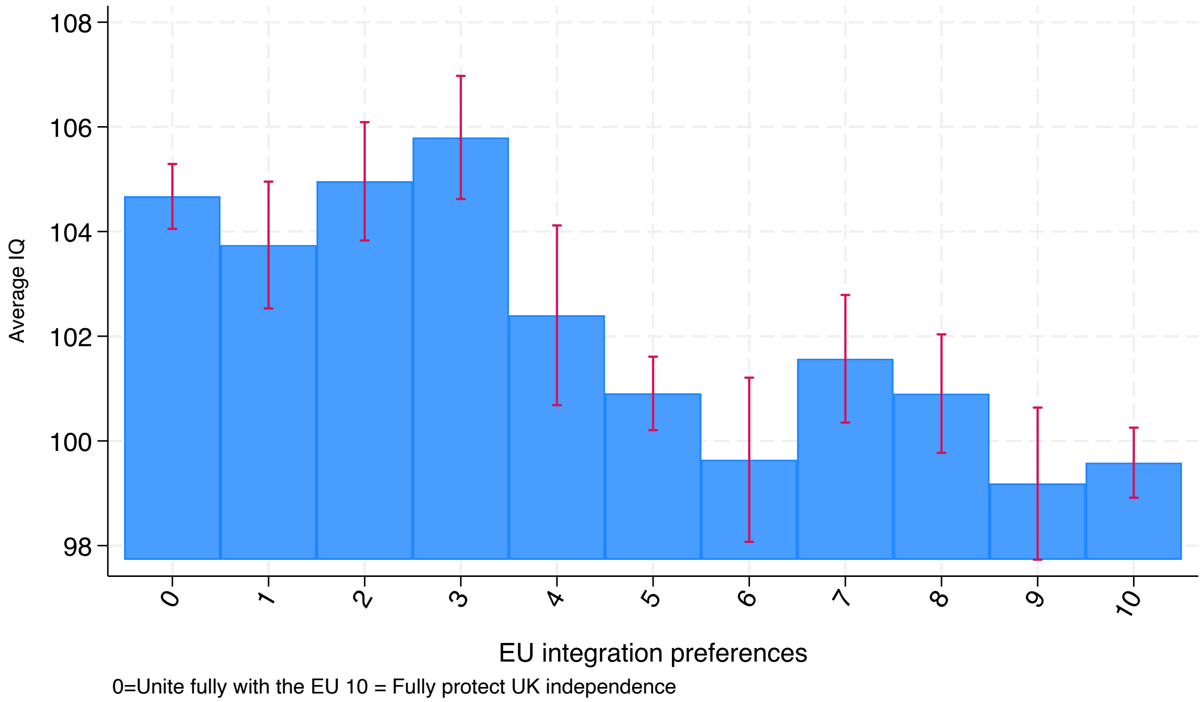


Figure 10: **UK Independence vs. Integration Preferences and Cognitive Abilities:** Responses to the question: *"On a 0 to 10 scale, where 0 means the UK should do all it can to unite fully with the European Union, and 10 means the UK should do all it can to protect its independence from the European Union, where would you place yourself on this scale?"* in Understanding Society, Wave 11 (2019–2020). IQ is calculated using the numerical series score from Wave 3 (2011–2013). Averages are weighted using using the survey's cross-sectional weights.

behavior. We employed pre-registered experimental treatments to manipulate participants' beliefs about the cognitive abilities of others. We then complemented this with textual analysis to deepen our understanding. The findings broadly confirm our hypotheses: individuals with higher cognitive abilities are more likely to support policies with beneficial indirect effects, even when these policies entail seemingly costly direct and immediate consequences. However, this preference is conditional on their expectations regarding other voters' capacity to comprehend the broader implications of such policies. Moreover, our results reveal an unexpected independent role of education in shaping the ability to grasp indirect effects.

Our findings carry significant implications for public policy. While the long-term impacts of early childhood interventions on cognitive development continue to be evaluated (e.g., Kautz et al., 2014), there is mounting evidence that such programs can successfully foster cognitive skills (e.g., Berger et al., 2025; Brown et al., 2024; García and Heckman, 2023; List and Uchida, 2024; Zhang et al., 2024).[31] Our results point to an additional benefit of enhancing cognitive abilities: not only could individual economic prospects improve, but the overall quality of societal decision-making could also advance.

The observation that individual education exerts an effect similar to—but independent of—the cognitive abilities measured in our tests further emphasizes the critical role of education in shaping demand for optimal policies. This finding opens avenues for future research, particularly on how education interacts with cognitive abilities to influence policy preferences. Moreover, the link between economic conditions and the demand for suboptimal policies is underscored by evidence that adverse economic shocks or poverty can impair cognitive functioning (Mani et al., 2013), while positive affect can enhance it (e.g., Oswald et al., 2015). This dynamic suggests the potential for self-reinforcing economic cycles that may perpetuate disparities.

Another dimension warranting attention is the role of media consumption in shaping cognitive abilities. Research indicates a correlation between childhood exposure to entertainment television and social media and reduced cognitive sophistication (Durante et al., 2019). Our study contributes to this dialogue by examining the broader societal repercussions of diminished cognitive abilities.

Finally, our findings highlight the critical importance of trust in the democratic process, particularly in shaping the demand for good policies. Here, trust transcends its traditional definition, encompassing not only confidence in the integrity and pro-social behavior of others but also faith in their rationality and ability to evaluate the long-term consequences of policies. This trust is essential for collective support of policies that may entail short-term costs but yield substantial long-term societal benefits.

[31] For instance, two iconic programs highlight these effects: the Perry Preschool Project increased cognitive abilities by half a standard deviation at age 54, while the Abecedarian Project achieved an increase of one-third of a standard deviation at age 45 (García and Heckman, 2023).

In summary, our work highlights the complex nexus between citizens' cognitive abilities, their trust in collective rationality, and the dynamics of support for public policies. It underscores the importance of promoting cognitive abilities and trust in others' rationality as collective objectives. These efforts have far-reaching implications, not only for individual well-being but also for the strength and effectiveness of democratic institutions.

References

Achen, Christopher and Larry Bartels, *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, Princeton University Press, 2017.

Agranov, Marina, Elizabeth Potamites, Andrew Schotter, and Chloe Tergiman, "Beliefs and Endogenous Cognitive Levels: An Experimental Study," *Games and Economic Behavior*, 2012, *75* (2), 449–463.

Alaoui, Larbi and Antonio Penta, "Endogenous Depth of Reasoning," *Review of Economic Studies*, 2016, *83* (4), 1297–1333.

_ , Katharina A Janezic, and Antonio Penta, "Reasoning about Others' Reasoning," *Journal of Economic Theory*, 2020, *189*, 105091.

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk, "Misperceived Social Norms and Willingness To Act Against Climate Change," *Review of Economics and Statistics*, 2024, pp. 1–46.

Beilharz, Hans-Jörg and Hans Gersbach, "Voting Oneself into a Crisis," *Macroeconomic Dynamics*, 2016, *20* (4), 954–984.

Berger, Eva M, Ernst Fehr, Henning Hermes, Daniel Schunk, and Kirsten Winkel, "The Impact of Working-Memory Training On Children's Cognitive and Noncognitive Skills," *Journal of Political Economy*, 2025, *133* (2), 000–000.

Besley, Timothy and Stephen Coate, "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 1997, *112* (1), 85–114.

Black, Duncan, "On the Rationale of Group Decision-Making," *Journal of Political Economy*, 1948, *56* (1), 23–34.

Brennan, Geoffrey and Alan Hamlin, "Expressive Voting and Electoral Equilibrium," *Public Choice*, 1998, *95* (1), 149–175.

_ **and Loren Lomasky**, *Democracy and Decision: The Pure Theory of Electoral Preference*, Cambridge University Press, 1997.

Brown, Christina, Supreet Kaur, Geeta Kingdon, and Heather Schofield, "Cognitive Endurance as Human Capital," *Quarterly Journal of Economics*, 12 2024.

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment," *Proceedings of The National Academy of Sciences*, 2009, *106* (19), 7745–7750.

Börgers, Tilman, "Costly Voting," *American Economic Review*, 2004, *94* (1), 57–66.

Caplan, Bryan, *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*, Princeton University Press, 2011.

Carter, John R and Stephen D Guerette, "An Experimental Study of Expressive Voting," *Public Choice*, 1992, *73* (3), 251–260.

Cattell, Raymond B, "Theory of Fluid and Crystallized Intelligence: A Critical Experiment.," *Journal of Educational Psychology*, 1963, *54* (1), 1.

Dal Bó, Ernesto, Frederico Finan, Olle Folke, Torsten Persson, and Johanna Rickne, "Who Becomes a Politician?," *Quarterly Journal of Economics*, 2017, *132* (4), 1877–1914.

— , **Pedro Dal Bó, and Erik Eyster**, "The Demand for Bad Policy when Voters Underappreciate Equilibrium Effects," *Review of Economic Studies*, 2018, *85* (2), 964–998.

Deary, Ian J, G David Batty, and Catharine R Gale, "Bright Children Become Enlightened Adults," *Psychological Science*, 2008, *19* (1), 1–6.

— , — , **and** — , "Childhood Intelligence Predicts Voter Turnout, Voting Preferences, and Political Involvement in Adulthood: The 1970 British Cohort Study," *Intelligence*, 2008, *36* (6), 548–555.

Dechezleprêtre, Antoine, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva, "Fighting Climate Change: International Attitudes Toward Climate Policies," 2022. Unpublished Manuscript.

Denny, Kevin and Orla Doyle, "Political Interest, Cognitive Ability and Personality: Determinants of Voter Turnout In Britain," *British Journal of Political Science*, 2008, *38* (2), 291–310.

Downs, Anthony, "An Economic Theory of Political Action in a Democracy," *Journal of Political Economy*, 1957, *65* (2), 135–150.

Durante, Ruben, Paolo Pinotti, and Andrea Tesei, "The Political Legacy of Entertainment TV," *American Economic Review*, 2019, *109* (7), 2497–2530.

Elinder, Mikael and Oscar Erixson, "An Inquiry into the Relationship Between Intelligence and Prosocial Behavior: Evidence from Swedish Population Registers," 2022. Unpublished Manuscript.

Fe, Eduardo, David Gill, and Victoria Prowse, "Cognitive Skills, Strategic Sophistication, and Life Outcomes," *Journal of Political Economy*, 2022, *130* (10), 2643–2704.

Fearon, James D, "Self-Enforcing Democracy," *Quarterly Journal of Economics*, 2011, *126* (4), 1661–1708.

Feddersen, Timothy and Wolfgang Pesendorfer, "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, 1997, pp. 1029–1058.

Fisher, Gwenith G, John J McArdle, Ryan J McCammon, Amanda Sonnega, and David Weir, "New Measures of Fluid Intelligence in the HRS," 2013. Unpublished Manuscript.

Frederick, Shane, "Cognitive Reflection and Decision Making," *Journal of Economic Perspectives*, 2005, *19* (4), 25–42.

García, Jorge Luis and James J Heckman, "Parenting Promotes Social Mobility within and across Generations," *Annual Review of Economics*, 2023, *15*, 349–388.

Gill, David and Victoria Prowse, "Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis," *Journal of Political Economy*, 2016, *124* (6), 1619–1676.

Halevy, Yoram, Johannes Hoelzemann, and Terri Kneeland, "Magic Mirror on the Wall, Who Is the Smartest One of All?," 2023. Unpublished Manuscript.

Hillygus, Sunshine D, "The Missing Link: Exploring the Relationship Between Higher Education and Political Engagement," *Political Behavior*, 2005, *27*, 25–47.

Kahan, Dan M, "Ideology, Motivated Reasoning, and Cognitive Reflection," *Judgment and Decision Making*, 2013, *8* (4), 407–424.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans, "Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success," 2014. Unpublished Manuscript.

Lambrecht, Marco, Eugenio Proto, Aldo Rustichini, and Andis Sofianos, "Intelligence Disclosure and Cooperation In Repeated Interactions," *American Economic Journal: Microeconomics*, 2024, *16* (3), 199–231.

List, John A and Haruka Uchida, "Here Today, Gone Tomorrow? Toward An Understanding of Fade-Out in Early Childhood Education Programs," 2024. Unpublished Manuscript.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir, and Jiaying Zhao, "Poverty Impedes Cognitive Function," *Science*, 2013, *341* (6149), 976–980.

McMurray, Joseph C, "Aggregating Information by Voting: The Wisdom of the Experts Versus the Wisdom of the Masses," *Review of Economic Studies*, 2013, *80* (1), 277–312.

Morton, Rebecca, Jean-Robert Tyran, and Erik Wengström, "Income and Ideology: How Personality Traits, Cognitive Abilities, and Education Shape Political Attitudes," 2011. Unpublished Manuscript.

Myatt, David P, "On the Theory of Strategic Voting," *Review of Economic Studies*, 2007, *74* (1), 255–281.

North, Douglass C, *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, 1990.

Ofcom, "Adults' Media Use and Attitudes Report 2020/21," 2021.

Oskarsson, Sven, David Cesarini, Christopher T Dawes, James H Fowler, Magnus Johannesson, Patrik KE Magnusson, and Jan Teorell, "Linking Genes and Political Orientations: Testing the Cognitive Ability As Mediator Hypothesis," *Political Psychology*, 2015, *36* (6), 649–665.

Oswald, Andrew J, Eugenio Proto, and Daniel Sgroi, "Happiness and Productivity," *Journal of Labor Economics*, 2015, *33* (4), 789–822.

Ottinger, Sebastian and Nico Voigtländer, "History's Masters: The Effect of European Monarchs on State Performance," *Econometrica*, Forthcoming.

Persson, Torsten and Guido Tabellini, *Political Economics: Explaining Economic Policy*, MIT Press, 2002.

Pons, Vincent and Clémence Tricaud, "Expressive Voting and Its Cost: Evidence from Runoffs with Two or Three Candidates," *Econometrica*, 2018, *86* (5), 1621–1649.

Proto, Eugenio, Aldo Rustichini, and Andis Sofianos, "Intelligence, Personality, and Gains from Cooperation in Repeated Interactions," *Journal of Political Economy*, 2019, *127* (3), 1351–1390.

_ , _ , and _ , "Intelligence, Errors, and Cooperation in Repeated Interactions," *Review of Economic Studies*, 2022, *89* (5), 2723–2767.

Raven, John, "Raven's Advanced Progressive Matrices and Mill Hill Vocabulary Scale: Occupational user's guide," 2003.

Romer, David, "Misconceptions and Political Outcomes," *The Economic Journal*, 2003, *113* (484), 1–20.

Smith, Adam, "An Enquiry into the Nature and Causes of the Wealth of Nations," 1776.

Tyran, Jean-Robert, "Voting When Money and Morals Conflict: An Experimental Test of Expressive Voting," *Journal of Public Economics*, 2004, *88* (7-8), 1645–1664.

Vora, Ravi N., "Carbon Pricing and Bounded Reasoning," 2023. Unpublished Manuscript.

Willadsen, Helene, Sarah Zaccagni, Marco Piovesan, and Erik Wengström, "Measures of Cognitive Ability and Choice Inconsistency," *Journal of Economic Behavior & Organization*, 2024, *220*, 495–506.

Zhang, Yuan S, Elizabeth Frankenberg, and Duncan Thomas, "Education and Adult Cognition in a Low-income Setting: Differences among Adult Siblings," 2024. Unpublished Manuscript.

Zmigrod, Leor, Peter Jason Rentfrow, and Trevor W Robbins, "The Partisan Mind: Is Extreme Political Partisanship Related to Cognitive Inflexibility?," *Journal of Experimental Psychology: General*, 2020, *149* (3), 407.

A Appendix

A.1 Experimental Design: Timing Overview

Day 1: Measuring Cognitive Skills

- Part 1: 18 Raven's Advanced Progressive Matrices
- Part 2: Number Sequences Task (from Understanding Society)
- Part 3: 6-question Cognitive Reflection Test

Day 2: Measuring Understanding of Policies' Effects

- Part 4: behavior in exogenous games
  - 5 rounds of PD with 5 different opponents and feedback
  - 1 round of HG without feedback
- Part 5: 2 rounds with choice of PD or HG
  - Within-subject manipulation of opponent's cognitive abilities
  - Subject learns opponent's score in Part 1 before choice
  - Treatments: high cognitive ability opponent (15/18 matrices solved correctly) versus low cognitive ability opponent (3/18)
- Part 6: open-ended questions about perceived consequences of raising minimum wage and reducing inflow of migrants

A.2 Experimental Design: Screenshots

Figure A.1: 5 Rounds of (Static) Prisoner's Dilemma with Feedback

Part 4, Interaction 1

The EPs you will receive in this interaction will depend on your choice and on the choice of Participant A.

The EPs Participant A received in this interaction depended on his/her choice and on the choice of Participant B.

When choosing, Participant A did not know Participant B's choice but saw the table on the right-hand side of your screen.

Your Choice	Participant A's Choice	Your Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Participant A's Choice	Participant B's Choice	Participant A's Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Please choose your action:

- ◯ Action 1
- ◯ Action 2

Figure A.2: 1 Round of Harmony Game without Feedback

Part 4, Interaction 6

The EPs you will receive in this interaction will depend on your choice and on the choice of Participant K.

The EPs Participant K received in this interaction depended on his/her choice and on the choice of Participant L.

When choosing, Participant K did not know Participant L's choice but saw the table on the right-hand side of your screen.

Your Choice	Participant K's Choice	Your Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

Participant K's Choice	Participant L's Choice	Participant K's Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

Please choose your action:

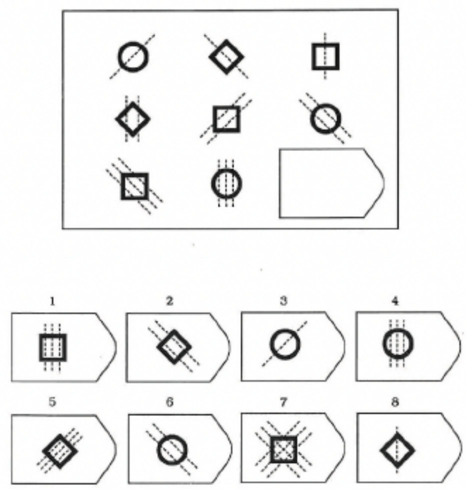
- Action 1
- Action 2

Figure A.3: Experimental Manipulation of Opponent's Cognitive Abilities

There are two important differences with respect to Part 4.

First, before each interaction, you will learn the other participants' performance in the 18 puzzles you solved yesterday.

Just to remind you, this is one of those puzzles:



These puzzles are Raven's Matrices and, according to Wikipedia, they are *"a non-verbal test typically used to measure general human intelligence and abstract reasoning and is regarded as a non-verbal estimate of fluid intelligence"*.

Figure A.4: Game Choice with High Cognitive Ability Opponent

Part 5, Interaction 1: Choose Table of Earnings

In Interaction 1, your earnings depend on your choice and on the choice of another participant, Participant K.

Participant K solved **15 out of 18** Raven's Matrices correctly.

Participants' performance ranged from 0 to 18 matrices solved correctly with an average of 9.

Participant K participated in two interactions, one using the Original Table of Earnings and one using the Alternative Table of Earnings. You can choose what tables of earnings to use for this interaction. Your choice determines (a) what table of earnings we use to determine your EPs depending on your choice and on Participant K's choice and (b) whether we use Participant K's choice in their interaction using the Original Table of Earnings or in their interaction using the Alternative Table of Earnings.

Original Tables of Earnings

Your Choice	Participant K's Choice	Your Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Participant K's Choice	Participant L's Choice	Participant K's Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Alternative Tables of Earnings

Your Choice	Participant K's Choice	Your Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

Participant K's Choice	Participant L's Choice	Participant K's Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

What tables of earnings do you choose for this interaction?

- ○ Original Tables of Earnings
- ○ Alternative Tables of Earnings

Figure A.5: Game Choice with Low Cognitive Ability Opponent

Part 5, Interaction 2: Choose Table of Earnings

In Interaction 2, your earnings depend on your choice and on the choice of another participant, Participant M.

Participant M solved **3 out of 18** Raven's Matrices correctly.

Participants' performance ranged from 0 to 18 matrices solved correctly with an average of 9.

Participant M participated in two interactions, one using the Original Table of Earnings and one using the Alternative Table of Earnings. You can choose what tables of earnings to use for this interaction. Your choice determines (a) what table of earnings we use to determine your EPs depending on your choice and on Participant M's choice and (b) whether we use Participant M's choice in their interaction using the Original Table of Earnings or in their interaction using the Alternative Table of Earnings.

Original Tables of Earnings

Your Choice	Participant M's Choice	Your Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Participant M's Choice	Participant O's Choice	Participant M's Earnings
Action 1	Action 1	9
Action 2	Action 1	11
Action 1	Action 2	3
Action 2	Action 2	5

Alternative Tables of Earnings

Your Choice	Participant M's Choice	Your Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

Participant M's Choice	Participant O's Choice	Participant M's Earnings
Action 1	Action 1	8
Action 2	Action 1	7
Action 1	Action 2	2
Action 2	Action 2	1

What tables of earnings do you choose for this interaction?

- ○ Original Tables of Earnings
- ○ Alternative Tables of Earnings

B Supplementary Analysis and Descriptive Statistics

Table A.1: **Correlation Between Different Measures of Cognitive Abilities and Education:** Raven Score 6 is calculated using 6 matrices; p-values in parentheses.

Variables	Raven	CRT	NS	Raven 6	Post-Grad	Graduate	Further Sec	Secondary	Primary
Raven	1.000								
CRT	0.472	1.000							
	(0.000)								
NS	0.407	0.418	1.000						
	(0.000)	(0.000)							
Raven 6	0.841	0.402	0.330	1.000					
	(0.000)	(0.000)	(0.000)						
Post-Grad	0.205	0.181	0.049	0.158	1.000				
	(0.000)	(0.000)	(0.192)	(0.000)					
Graduate	0.058	0.076	0.098	0.051	-0.481	1.000			
	(0.123)	(0.045)	(0.010)	(0.181)	(0.000)				
Further Sec	-0.120	-0.142	-0.036	-0.120	-0.245	-0.401	1.000		
	(0.001)	(0.000)	(0.343)	(0.001)	(0.000)	(0.000)			
Secondary	-0.180	-0.156	-0.132	-0.123	-0.229	-0.375	-0.191	1.000	
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)		
Primary	-0.049	-0.039	-0.080	-0.022	-0.058	-0.095	-0.049	-0.045	1.000
	(0.193)	(0.297)	(0.034)	(0.555)	(0.124)	(0.012)	(0.199)	(0.230)	

Table A.2: **Effect of Cognitive Abilities on Behavior in Games with Visible Controls' Coefficients:** The binary dependent variable is the action Defect. CRT and Raven scores are standardized. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Raven Score (std)	0.03712*	0.06014***	0.07066***	0.09487***	0.06973***	–0.02328**
	(0.0605)	(0.0018)	(0.0004)	(0.0000)	(0.0004)	(0.0499)
CRT Score (std)	–0.01364	–0.03543*	0.01612	–0.01651	–0.02316	–0.03677***
	(0.5026)	(0.0733)	(0.4146)	(0.3987)	(0.2568)	(0.0037)
Income	0.00017	–0.00016	–0.00026	–0.00072	0.00069	0.00007
	(0.7651)	(0.7596)	(0.6090)	(0.1679)	(0.1925)	(0.8214)
Female	0.00985	–0.00860	0.02282	–0.05230	0.02070	–0.06167***
	(0.7783)	(0.8045)	(0.5096)	(0.1265)	(0.5599)	(0.0050)
Age	0.00223*	0.00060	0.00213*	–0.00121	0.00105	–0.00050
	(0.0531)	(0.5919)	(0.0614)	(0.2925)	(0.3730)	(0.4913)
Post-Graduate	–0.01153	0.09667*	–0.02153	–0.03751	–0.02687	–0.05326
	(0.8535)	(0.0872)	(0.7211)	(0.5336)	(0.6594)	(0.1212)
Graduate	–0.05213	0.09505**	–0.03696	–0.05114	–0.02375	–0.04914*
	(0.3212)	(0.0439)	(0.4752)	(0.3139)	(0.6439)	(0.0865)
Further Secondary	–0.02823	0.09324*	–0.00432	–0.08759	–0.05820	–0.03436
	(0.6408)	(0.0919)	(0.9416)	(0.1268)	(0.3152)	(0.2804)
Sum of Previous Payoffs		0.01564***				
		(0.0026)				
Sum of Previous Payoffs			0.01303***			
			(0.0006)			
Sum of Previous Payoffs				0.00707**		
				(0.0312)		
Sum of Previous Payoffs					0.00957***	
					(0.0011)	
Sum of Previous Payoffs						–0.00637***
						(0.0008)
N	701	701	701	701	701	701

Table A.3: **Effect of Cognitive Abilities on Behavior in Games without Controls:** The binary dependent variable is the action Defect. CRT and Raven scores are standardized. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Raven Score (std)	0.02960	0.06264***	0.05913***	0.09370***	0.07031***	–0.02251**
	(0.1251)	(0.0008)	(0.0019)	(0.0000)	(0.0002)	(0.0484)
CRT Score (std)	–0.00876	–0.02993	0.01794	–0.01553	–0.02006	–0.03580***
	(0.6557)	(0.1129)	(0.3571)	(0.4042)	(0.2975)	(0.0034)
Sum of Previous Payoffs		0.01631***				
		(0.0017)				
Sum of Previous Payoffs			0.01338***			
			(0.0004)			
Sum of Previous Payoffs				0.00633*		
				(0.0522)		
Sum of Previous Payoffs					0.00984***	
					(0.0008)	
Sum of Previous Payoffs						–0.00628***
						(0.0012)
N	701	701	701	701	701	701

Table A.4: **Effect of Cognitive Skills on Behavior in Games with 6-Items Raven Score:** The binary dependent variable is the action Defect. CRT and Raven scores are standardized. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Raven Score 6 items (std)	0.02595	0.05273***	0.05797***	0.05721***	0.03979**	–0.01757
	(0.1644)	(0.0027)	(0.0021)	(0.0019)	(0.0368)	(0.1577)
CRT Score (std)	–0.00738	–0.02963	0.02482	0.00303	–0.00788	–0.04022***
	(0.7082)	(0.1261)	(0.1969)	(0.8759)	(0.6943)	(0.0020)
Income	0.00022	–0.00011	–0.00019	–0.00055	0.00082	0.00003
	(0.7024)	(0.8320)	(0.7108)	(0.2964)	(0.1253)	(0.9166)
Female	0.00827	–0.01216	0.01905	–0.05664*	0.01761	–0.06011***
	(0.8134)	(0.7258)	(0.5820)	(0.0998)	(0.6220)	(0.0061)
Age	0.00205*	0.00037	0.00182	–0.00175	0.00063	–0.00039
	(0.0749)	(0.7393)	(0.1103)	(0.1277)	(0.5898)	(0.5935)
Post-Graduate	–0.00422	0.10685*	–0.00919	–0.01811	–0.01187	–0.05735*
	(0.9460)	(0.0574)	(0.8784)	(0.7651)	(0.8452)	(0.0962)
Graduate	–0.04719	0.10166**	–0.02881	–0.03758	–0.01343	–0.05096*
	(0.3680)	(0.0306)	(0.5771)	(0.4609)	(0.7940)	(0.0749)
Further Secondary	–0.02499	0.09940*	0.00248	–0.08074	–0.05335	–0.03513
	(0.6790)	(0.0730)	(0.9664)	(0.1630)	(0.3599)	(0.2743)
Sum of Previous Payoffs		0.01561***				
		(0.0026)				
Sum of Previous Payoffs			0.01323***			
			(0.0005)			
Sum of Previous Payoffs				0.00756**		
				(0.0223)		
Sum of Previous Payoffs					0.00988***	
					(0.0008)	
Sum of Previous Payoffs						–0.00646***
						(0.0006)
N	701	701	701	701	701	701

Table A.5: **Joint Effects of Raven and CRT Scores on Behavior in Games:** The binary dependent variable is the action Defect. The baseline (omitted) is Low Raven and Low CRT Scores, defined as scores below the respective medians. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; $p - values$ are in brackets; * $p - value < 0.1$, ** $p - value < 0.05$, *** $p - value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Low Raven, High CRT	0.0328	–0.0407	0.1274**	0.0209	–0.0543	–0.1052***
	(0.5498)	(0.4750)	(0.0235)	(0.7159)	(0.3415)	(0.0054)
High Raven, Low CRT	0.0364	0.1823***	0.1482**	0.1764***	0.0587	–0.0891**
	(0.5612)	(0.0008)	(0.0206)	(0.0025)	(0.3481)	(0.0490)
High Raven, High CRT	0.0455	0.0734	0.1682***	0.1318***	0.0560	–0.1227***
	(0.3136)	(0.1027)	(0.0003)	(0.0040)	(0.2229)	(0.0003)
Socio-Demographics	Yes	Yes	Yes	Yes	Yes	Yes
N	701	701	701	701	701	701

Table A.6: **Effect of CRT Score on Behavior in Games:** The binary dependent variable is the action Defect. CRT score is standardized. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
CRT Score (std)	0.00237	–0.01024	0.04669**	0.02504	0.00707	–0.04649***
	(0.8972)	(0.5791)	(0.0105)	(0.1563)	(0.7013)	(0.0001)
Income	0.00029	0.00005	–0.00001	–0.00039	0.00094*	–0.00001
	(0.6045)	(0.9252)	(0.9775)	(0.4566)	(0.0763)	(0.9609)
Female	0.00915	–0.01122	0.02036	–0.05491	0.01904	–0.05934***
	(0.7939)	(0.7480)	(0.5586)	(0.1127)	(0.5945)	(0.0067)
Age	0.00190*	0.00006	0.00146	–0.00211*	0.00038	–0.00027
	(0.0986)	(0.9569)	(0.2034)	(0.0680)	(0.7460)	(0.7123)
Post-Graduate	0.00019	0.11644**	0.00099	–0.00804	–0.00585	–0.06012*
	(0.9976)	(0.0400)	(0.9868)	(0.8943)	(0.9228)	(0.0852)
Graduate	–0.04481	0.10816**	–0.02376	–0.03131	–0.00996	–0.05153*
	(0.3936)	(0.0225)	(0.6466)	(0.5372)	(0.8461)	(0.0738)
Further Secondary	–0.02725	0.09618*	–0.00312	–0.08473	–0.05661	–0.03238
	(0.6533)	(0.0870)	(0.9583)	(0.1429)	(0.3335)	(0.3057)
Sum of Previous Payoffs		0.01543***				
		(0.0030)				
Sum of Previous Payoffs			0.01330***			
			(0.0005)			
Sum of Previous Payoffs				0.00779**		
				(0.0195)		
Sum of Previous Payoffs					0.01005***	
					(0.0006)	
Sum of Previous Payoffs						–0.00653***
						(0.0005)
N	701	701	701	701	701	701

Table A.7: **Effect of Raven Score on Behavior in Games:** The binary dependent variable is the action Defect. Raven score is standardized. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
Raven Score (std)	0.03129*	0.04537**	0.07760***	0.08777***	0.05981***	–0.03847***
	(0.0796)	(0.0112)	(0.0000)	(0.0000)	(0.0006)	(0.0004)
Income	0.00015	–0.00022	–0.00024	–0.00074	0.00066	–0.00001
	(0.7935)	(0.6896)	(0.6431)	(0.1546)	(0.2148)	(0.9727)
Female	0.01375	0.00169	0.01831	–0.04740	0.02742	–0.05172**
	(0.6900)	(0.9608)	(0.5937)	(0.1613)	(0.4291)	(0.0170)
Age	0.00209*	0.00024	0.00230**	–0.00138	0.00080	–0.00089
	(0.0654)	(0.8284)	(0.0409)	(0.2209)	(0.4874)	(0.1990)
Post-Graduate	–0.01700	0.08175	–0.01534	–0.04400	–0.03608	–0.06449*
	(0.7828)	(0.1422)	(0.7975)	(0.4600)	(0.5470)	(0.0586)
Graduate	–0.05568	0.08598*	–0.03316	–0.05494	–0.02916	–0.05456*
	(0.2869)	(0.0667)	(0.5200)	(0.2766)	(0.5679)	(0.0563)
Further Secondary	–0.02887	0.09264*	–0.00424	–0.08765	–0.05851	–0.03287
	(0.6326)	(0.0943)	(0.9429)	(0.1255)	(0.3136)	(0.3103)
Sum of Previous Payoffs		0.01551***				
		(0.0029)				
Sum of Previous Payoffs			0.01296***			
			(0.0006)			
Sum of Previous Payoffs				0.00705**		
				(0.0317)		
Sum of Previous Payoffs					0.00955***	
					(0.0012)	
Sum of Previous Payoffs						–0.00633***
						(0.0008)
N	701	701	701	701	701	701

Table A.8: **Effect of Numerical Series Score on Behavior in Games:** The binary dependent variable is the action Defect. Numerical Series (NS) score is standardized. The omitted category of education is Secondary and Primary. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Prisoners' Dilemma					Harmony Game
	Round 1	Round 2	Round 3	Round 4	Round 5	Round 1
	b/p	b/p	b/p	b/p	b/p	b/p
NS Score (std)	0.01945	0.01331	0.04367**	0.03842**	0.04445**	–0.02658**
	(0.2722)	(0.4410)	(0.0113)	(0.0268)	(0.0105)	(0.0140)
Income	0.00026	–0.00002	0.00005	–0.00038	0.00086	–0.00013
	(0.6423)	(0.9732)	(0.9175)	(0.4602)	(0.1049)	(0.6543)
Female	0.01706	–0.00134	0.02337	–0.04748	0.03602	–0.05348**
	(0.6291)	(0.9696)	(0.5068)	(0.1730)	(0.3097)	(0.0137)
Age	0.00192*	–0.00002	0.00186	–0.00190*	0.00045	–0.00070
	(0.0907)	(0.9884)	(0.1018)	(0.0939)	(0.6986)	(0.3176)
Post-Graduate	–0.00499	0.10476*	0.01697	–0.00502	–0.01661	–0.07504**
	(0.9352)	(0.0590)	(0.7748)	(0.9327)	(0.7787)	(0.0310)
Graduate	–0.05174	0.09828**	–0.02131	–0.03668	–0.02528	–0.05457*
	(0.3236)	(0.0377)	(0.6792)	(0.4673)	(0.6197)	(0.0574)
Further Secondary	–0.03144	0.09270*	–0.00956	–0.09243	–0.06589	–0.02527
	(0.6051)	(0.0987)	(0.8730)	(0.1114)	(0.2627)	(0.4321)
Sum of Previous Payoffs		0.01519***				
		(0.0035)				
Sum of Previous Payoffs			0.01294***			
			(0.0006)			
Sum of Previous Payoffs				0.00775**		
				(0.0200)		
Sum of Previous Payoffs					0.00997***	
					(0.0007)	
Sum of Previous Payoffs						–0.00624***
						(0.0010)
N	701	701	701	701	701	701

Table A.9: **Effect of Raven Score, CRT Score, Education and Opponent's Cognitive Abilities on Preferences over Games, without Controls:** The binary dependent variable is the choice of the HG. Raven and CRT Scores are standardized. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	High Rav.Opp.	Low Rav.Opp.	All	All
	Vote HG	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p	b/p
Raven Score (std)	0.07773***	–0.01181	–0.01130	–0.01041
	(0.0000)	(0.4796)	(0.4967)	(0.5418)
CRT Score (std)	0.02078	0.02806*	0.02823*	0.02724*
	(0.2327)	(0.0913)	(0.0861)	(0.0982)
High Raven Opp.			–0.01148	–0.11982**
			(0.5535)	(0.0190)
High Raven Opp.× Raven Score			0.09280***	0.08355***
			(0.0000)	(0.0003)
High Raven Opp.× CRT Score			–0.00634	–0.01323
			(0.7655)	(0.5346)
High Raven Opp.× Graduate				0.12920**
				(0.0287)
High Raven Opp.× Post-Graduate				0.15002**
				(0.0200)
High Raven Opp.× Further Secondary				0.09129
				(0.2419)
Post-Graduate				–0.03377
				(0.4679)
Graduate				–0.06474
				(0.1214)
Further Secondary				–0.10348**
				(0.0489)
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	701	701	1402	1402

Table A.10: **Effect of Raven Score, CRT Score, Education and Opponent's Cognitive Abilities on Preferences over Games without Interactions:** The binary dependent variable is the choice of the HG. Raven and CRT Scores are standardized. The omitted education category is secondary or primary education. Socio-Demographics controls include: Age, Gender, and Income. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	All	All
	Vote HG	Vote HG
	b/p	b/p
Raven Score (std)	0.02348*	0.02349*
	(0.0914)	(0.0915)
CRT Score (std)	0.02314*	0.02314*
	(0.0815)	(0.0815)
High Raven Opp.		0.00259
		(0.8920)
Post-Graduate	0.01018	0.01019
	(0.7946)	(0.7944)
Graduate	–0.02183	–0.02183
	(0.5399)	(0.5398)
Further Secondary	–0.07746*	–0.07745*
	(0.0777)	(0.0778)
Socio-Demographics	Yes	Yes
Previous PD Payoffs	Yes	Yes
Treatment Order	Yes	Yes
N	1402	1402

Table A.11: **Effect of Raven Score, CRT Score, Education and Opponent's Cognitive Abilities on Preferences over Games using 6-items Raven Score:** The binary dependent variable is the the choice of the HG the HG. CRT score is standardized. The omitted education category is secondary or primary education. The interacted term Graduate+ indicates subjects with an Undergraduate and Postgraduate degree. Socio-Demographics controls include: Age, Gender and Income. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value<0.1$, ** $p-value<0.05$, *** $p-value<0.01$

	High Raven Opp.	Low Raven Opp.	All	All
	Vote HG	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p	b/p
Raven Score 6 items (std)	0.04406***	–0.02672*	–0.03053**	–0.02739*
	(0.0076)	(0.0894)	(0.0498)	(0.0787)
CRT Score (std)	0.02799*	0.03060*	0.02733*	0.03206**
	(0.0928)	(0.0652)	(0.0891)	(0.0486)
High Raven Opp.			–0.00709	–0.12742**
			(0.7135)	(0.0117)
High Raven Opp.× Raven Score 6 items			0.08043***	0.07416***
			(0.0001)	(0.0003)
High Raven Opp.× CRT Score			0.00506	–0.00503
			(0.8010)	(0.8061)
High Raven Opp.× Graduate				0.14113**
				(0.0156)
High Raven Opp.× Post-Graduate				0.16757***
				(0.0082)
High Raven Opp.× Further Secondary				0.10342
				(0.1797)
Post-Graduate	0.10369*	–0.06520	0.01568	–0.05953
	(0.0598)	(0.1709)	(0.6869)	(0.2026)
Graduate	0.06273	–0.07830*	–0.01753	–0.07843*
	(0.2158)	(0.0681)	(0.6210)	(0.0644)
Further Secondary	–0.02049	–0.12086**	–0.07688*	–0.11869**
	(0.7491)	(0.0245)	(0.0787)	(0.0259)
Socio-Demographics	Yes	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	701	701	1402	1402

Table A.12: **Effect of Raven Score, CRT Score, Education and Opponent's Cognitive Abilities on Preferences over Games, Excluding Subjects who Studied Game Theory:** The binary dependent variable is the choice of the HG. The 45 subjects who declared taking course in game theory have been excluded. Raven and CRT Scores are standardized. The omitted education category is secondary or primary education. Socio-Demographics controls include: Age, Gender, and Income. The coefficients represent the marginal effect and are calculated using a Logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p - values$ are in brackets; * $p - value < 0.1$, ** $p - value < 0.05$, *** $p - value < 0.01$

	High Rav.Opp.	Low Rav.Opp.	All	All
	Vote HG	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p	b/p
Raven Score (std)	0.05751***	–0.01503	–0.01860	–0.01387
	(0.0017)	(0.4037)	(0.2808)	(0.4264)
CRT Score (std)	0.01645	0.02906*	0.02546	0.02885*
	(0.3531)	(0.0959)	(0.1300)	(0.0895)
High Raven Opp.			–0.01268	–0.13032**
			(0.5196)	(0.0102)
High Raven Opp.× Raven Score			0.08777***	0.07800***
			(0.0001)	(0.0007)
High Raven Opp.× CRT Score			–0.00668	–0.01436
			(0.7535)	(0.5061)
High Raven Opp.× Graduate				0.14583**
				(0.0132)
High Raven Opp.× Post-Graduate				0.15649**
				(0.0153)
High Raven Opp.× Further Secondary				0.10014
				(0.2120)
Post-Graduate	0.09900*	–0.05944	0.01365	–0.05388
	(0.0768)	(0.2213)	(0.7327)	(0.2550)
Graduate	0.06588	–0.08151*	–0.02100	–0.08262*
	(0.1967)	(0.0596)	(0.5565)	(0.0516)
Further Secondary	–0.02848	–0.13090**	–0.08703**	–0.12568**
	(0.6630)	(0.0159)	(0.0466)	(0.0189)
Socio-Demographics	Yes	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	656	656	1312	1312

Table A.13: **Effect of Numerical Series Score and Opponents' Cognitive Abilities on Preferences over Games:** The binary dependent variable is the choice of the HG. Numerical Series (NS) score is standardized. The omitted education category is secondary or primary education. The interacted term Graduate+ indicates subjects with an Undergraduate and Postgraduate degree. Socio-Demographics controls include: Age, Gender and Income. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	High Raven Opp.	Low Raven Opp.	All
	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p
NS Score (std)	0.02543	–0.01144	–0.01391
	(0.1309)	(0.4961)	(0.3962)
High Raven Opp.			0.00033
			(0.9860)
High Raven Opp.× NS Score			0.04539**
			(0.0254)
Post-Graduate	0.13303**	–0.05180	0.03712
	(0.0133)	(0.2649)	(0.3325)
Graduate	0.07650	–0.06918	–0.00636
	(0.1350)	(0.1050)	(0.8600)
Further Secondary	–0.02320	–0.11593**	–0.07526*
	(0.7184)	(0.0311)	(0.0874)
Socio-Demographics	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes
N	701	701	1402

Table A.14: **Effect of Raven Score Only and Opponents' Cognitive Abilities on Preferences over Games:** The binary dependent variable is the choice of the HG. Raven score is standardized. The omitted education category is secondary or primary education. The interacted term Graduate+ indicates subjects with an Undergraduate and Postgraduate degree. Socio-Demographics controls include: Age, Gender and Income. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	High Raven Opp.	Low Raven Opp.	All	All
	Vote HG	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p	b/p
Raven Score (std)	0.06925***	–0.00572	–0.01020	–0.00442
	(0.0000)	(0.7327)	(0.5239)	(0.7866)
High Raven Opp.			–0.01185	–0.11820**
			(0.5383)	(0.0195)
High Raven Opp.× Raven Score			0.09028***	0.07811***
			(0.0000)	(0.0003)
High Raven Opp.× Graduate				0.12702**
				(0.0291)
High Raven Opp.× Post-Graduate				0.14496**
				(0.0227)
High Raven Opp.× Further Secondary				0.09012
				(0.2465)
Post-Graduate	0.09740*	–0.05236	0.01865	–0.04514
	(0.0726)	(0.2704)	(0.6296)	(0.3319)
Graduate	0.05856	–0.07164*	–0.01546	–0.06947*
	(0.2424)	(0.0936)	(0.6607)	(0.0981)
Further Secondary	–0.02686	–0.11777**	–0.07766*	–0.11310**
	(0.6758)	(0.0279)	(0.0757)	(0.0323)
Socio-Demographics	Yes	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	701	701	1402	1402

Table A.15: **Effect of CRT Score Only and Opponents' Cognitive Abilities on Preferences over Games:** The binary dependent variable is the choice of the HG. Raven score is standardized. The omitted education category is secondary or primary education. The interacted term Graduate+ indicates subjects with an Undergraduate and Postgraduate degree. Socio-Demographics controls include: Age, Gender and Income. The coefficients represent the marginal effect and are calculated using a logit estimator. Robust Standard Errors; in columns 3 and 4 the errors are clustered at the individual level; $p - values$ are in brackets; * $p - value < 0.1$, ** $p - value < 0.05$, *** $p - value < 0.01$

	High Raven Opp.	Low Raven Opp.	All	All
	Vote HG	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p	b/p
CRT Score (std)	0.04419***	0.02077	0.01513	0.02218
	(0.0058)	(0.1946)	(0.3308)	(0.1607)
High Raven Opp.			–0.00279	–0.13689***
			(0.8848)	(0.0064)
High Raven Opp.× CRT Score			0.03648*	0.02181
			(0.0640)	(0.2816)
High Raven Opp.× Graduate				0.15519***
				(0.0076)
High Raven Opp.× Post-Graduate				0.19238***
				(0.0022)
High Raven Opp.× Further Secondary				0.10835
				(0.1584)
Post-Graduate	0.11211**	–0.06888	0.01775	–0.07005
	(0.0425)	(0.1431)	(0.6479)	(0.1308)
Graduate	0.06761	–0.08148*	–0.01723	–0.08512**
	(0.1850)	(0.0551)	(0.6276)	(0.0442)
Further Secondary	–0.02170	–0.11865**	–0.07594*	–0.12046**
	(0.7360)	(0.0269)	(0.0831)	(0.0249)
Socio-Demographics	Yes	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes	Yes
N	701	701	1402	1402

Table A.16: **Effect of Raven Score on Preferences over Games in a Quadratic Model:** The binary dependent variable is the choice of the HG. Raven score is standardized. Socio-Demographics controls include: Age, Gender, Income and Education. The coefficients represent the marginal effect and are calculated using a logit estimator. The coefficients represent the marginal effects and are calculated using a logit estimator;. Robust Standard Errors; in columns 3 the errors are clustered at the individual level; $p-values$ are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	High Raven Opp.	Low Raven Opp.	All
	Vote HG	Vote HG	Vote HG
	b/p	b/p	b/p
Raven Score	–0.04989***	–0.03362*	–0.03585**
	(0.0037)	(0.0601)	(0.0414)
Raven Score^2	0.00371***	0.00178*	0.00183*
	(0.0000)	(0.0652)	(0.0542)
High Raven Opp.			–0.06727
			(0.4718)
High Raven Opp.$\times$ Raven Score			–0.01553
			(0.4692)
High Raven Opp.$\times$ Raven Score^2			0.00214*
			(0.0700)
Socio-Demographics	Yes	Yes	Yes
Previous PD Payoffs	Yes	Yes	Yes
Treatment Order	Yes	Yes	Yes
N	701	701	1402

Figure A.6: **Raven Scores and Preferences over Games by Treatment and Education:** Each rectangle of the histogram represents the share of the choice of the HG by Education in each of the two treatments (in blue and red). The bars represent the 95% confidence intervals

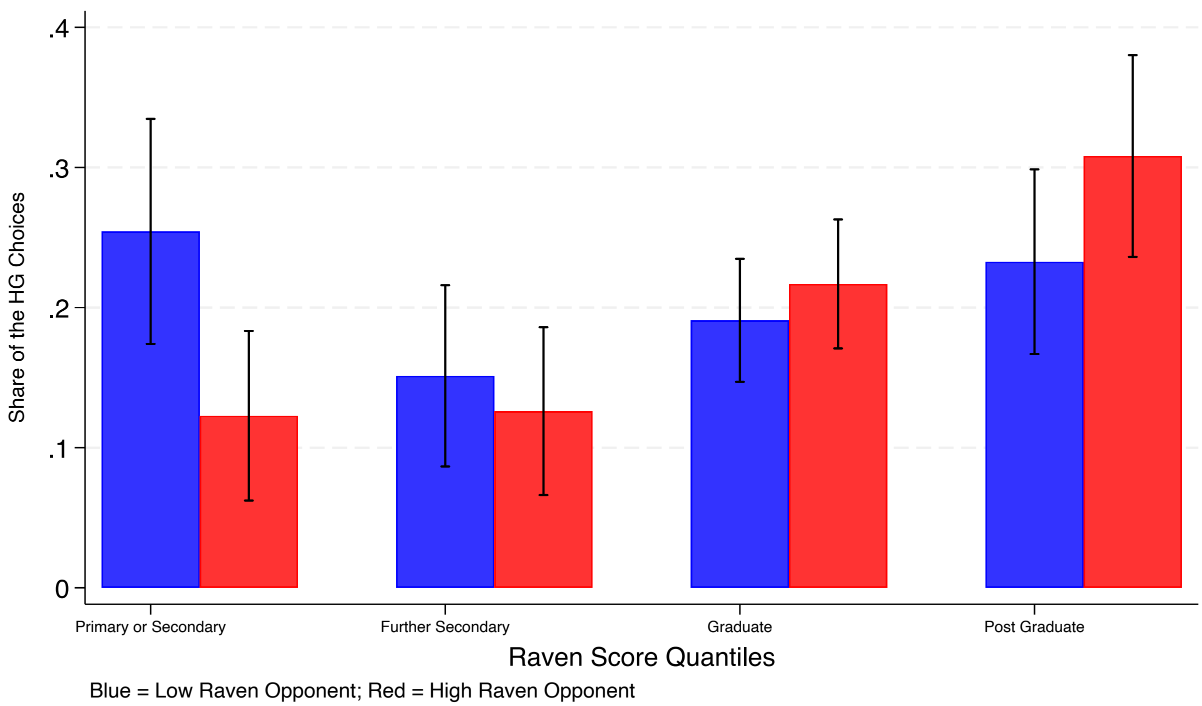


Table A.17: **Frequency of Cooperation for Different Opponents.** The cooperation rates are computed for the 5^{th} round of PD (fifth interaction number 5 among the participants) and to the 1^{st} round of HG (or interaction 6)

Subjects	**Low Raven (=3)**	**High Raven (= 15)**
% Cooperation in PD	0.565	0.263
% Cooperation in HG	0.869	1

C Textual Analysis

C.1 Descriptive Statistics of Textual Analysis

Table A.18: **Textual Analysis**: Main Variables

Variable	Mean	Std. Dev.	Min.	Max.	N
Imm. Conseq.	0.441	0.574	0	3	701
Min. Wage Conseq.	0.906	0.802	0	5	701
Right Leaning	4.738	1.903	1	10	701

Table A.19: **Correlation Between Main Variables of the Textual Analysis:** p-values in brackets

Variables	Imm. Conseq.	Min. Wage Conseq.	Right Leaning	Raven	CRT	Num. Series
Imm. Conseq.	1.000					
Min. Wage Conseq.	0.221	1.000				
	(0.000)					
Right Leaning	-0.150	0.148	1.000			
	(0.000)	(0.000)				
Raven	0.088	0.015	-0.117	1.000		
	(0.020)	(0.701)	(0.002)			
CRT	0.215	0.119	-0.059	0.472	1.000	
	(0.000)	(0.002)	(0.120)	(0.000)		
Num. Series	0.078	0.042	-0.038	0.407	0.418	1.000
	(0.039)	(0.268)	(0.315)	(0.000)	(0.000)	

C.2 Indirect Consequences Identified by Respondents

Indirect Consequences of Minimum Wage Reform

1. Increase in workers' consumption
2. Inflation
3. Greater inward migration to UK
4. Greater unemployment (due to, e.g., outsourcing)
5. Change in employers' recruiting strategy
6. Change in citizens' political behavior
7. Reduced reliance on government benefits
8. Increase in illegal labor agreements
9. Businesses closing
10. Greater workers' productivity

Indirect Consequences of Immigration Reform

1. UK nationals entering labor force, reduced reliance on government benefits, increased quality of public services
2. Reduction in quantity of goods and services produced
3. Reduction in quality of goods and services produced
4. Greater outward migration from the UK
5. Greater public spending on border and immigration control
6. Greater investment in human capital of UK nationals

7. Greater wages for workers, greater costs for entrepreneurs
8. Businesses closing
9. Reduced CO_2 emissions due to fewer houses being built

C.3 Supplementary Analysis

Table A.20: **Joint Effects of Raven and CRT Scores and Listed Indirect Effects:** Dependent Variable: Number of Indirect Consequences Open Questions on Policies' Indirect Effects. The baseline (omitted) is Low Raven and Low CRT Scores, defined as scores below the respective medians. The omitted education category is secondary or primary education. Robust Standard errors are in brackets; * $p-value < 0.1$, ** $p-value < 0.05$, *** $p-value < 0.01$

	Baseline	Baseline	Controls	Controls
	Immigration	Minimum Wage	Immigration	Minimum Wage
	b/se	b/se	b/se	b/se
Low Raven, High CRT	–0.072	0.053	–0.097	0.090
	(0.0662)	(0.1025)	(0.0692)	(0.1030)
High Raven, Low CRT	0.169**	0.234**	0.159**	0.191*
	(0.0667)	(0.0982)	(0.0670)	(0.0988)
High Raven, High CRT	0.180***	0.235***	0.129**	0.212***
	(0.0508)	(0.0738)	(0.0531)	(0.0801)
Age			–0.008	–0.013
			(0.0093)	(0.0129)
Age Squared			0.000	0.000
			(0.0001)	(0.0001)
Female			–0.015	–0.085
			(0.0431)	(0.0589)
Income			0.001	–0.000
			(0.0007)	(0.0009)
Post-Graduate			0.139*	–0.024
			(0.0716)	(0.1059)
Graduate			0.127**	0.117
			(0.0592)	(0.0877)
Further Secondary			–0.004	–0.112
			(0.0670)	(0.0965)
Constant	0.335***	0.750***	0.468**	0.885***
	(0.0383)	(0.0595)	(0.2264)	(0.2944)
N	701	701	701	701