Fostering Strategic Reasoning through Teamwork and Information about Opponent*

Martin J. Chegere† Marcel Fafchamps‡ Paolo Falco§

May 27, 2026

Abstract

We conduct lab experiments in Ghana and Tanzania to foster strategic thinking in two-person level-k games. Participants perform a series of tasks in which they guess a number as close as possible to an anonymous opponent's guess, times a multiplier. In a sample of small Ghanaian entrepreneurs, we investigate whether working in a team encourages strategic thinking. We find that team leaders display more strategic sophistication than individual players or team members. In Tanzania, we test whether participants' expectation of their opponent's level of strategic reasoning and their own revealed level of strategic reasoning are affected by (a) receiving written factual information about their opponent and (b) watching a video of their opponent introducing themselves. Receiving factual information increases participants' level of strategic reasoning as well as the level of strategic reasoning that they expect from their opponent. The video per se does not change strategic reasoning, but it enhances the impact of factual information when combined with it. These results suggest that human contact may either enhance perspective taking, e.g., via emotive and affective empathy, or facilitate the processing of information that is relevant for strategic reasoning, e.g., due to increased attention.

JEL codes: C91, D84, D91.

Keywords: strategic reasoning, perspective taking, empathy, expectations, level-k thinking, two-person guessing game.

*We are grateful to Vojtěch Bartoš, Stefano Caria, Vicky Fouka, Arman Rezaee, Pieter Serneels, and to participants in the 2024 Pacific Conference for Development Economics, the CSAE Conference 2024, and SEEDEC 2024 for their useful comments, and to all our research assistants for their outstanding work. We thank DANIDA for financial support through the Embassy of Denmark in Dar es Salaam (GDRP Phase II Project). The trial was pre-registered in the AEA RCT Registry (ID: AEARCTR-0010717) and received IRB approval from the ethics committee of the Department of Economics at the University of Copenhagen. All errors are our own.

†University of Dar es Salaam: chegere.martin@udsm.ac.tz

‡Stanford University: fafchamp@stanford.edu

§University of Copenhagen: paolo.falco@econ.ku.dk

1 Introduction

Strategic reasoning is central to economic environments in which optimal choices depend on the anticipated actions of others. In such settings, agents must form beliefs about how others will behave and choose best responses accordingly (e.g., Gans, 2026). While this logic is foundational in economics, there is still limited evidence on what determines the depth of strategic reasoning in practice, and on whether it can be improved through relatively simple interventions.

A large experimental literature shows that strategic sophistication is often limited. In beauty-contest and related guessing games, many participants either choose as if their opponents were guessing randomly or display behavior that cannot easily be rationalized by higher-order reasoning (Stahl II and Wilson, 1994; Stahl and Wilson, 1995; Costa-Gomes and Crawford, 2006; Grosskopf and Nagel, 2008; Sbriglia, 2008; Crawford et al., 2013; Nagel, 1995; Fragiadakis et al., 2016). This literature points to substantial heterogeneity in strategic thinking, but it leaves open a central question: what helps individuals reason more deeply about the behavior of others?

In this paper, we study two distinct mechanisms that may foster strategic sophistication. The first is *teamwork*. Deliberating with others may improve strategic reasoning by exposing individuals to alternative perspectives, pooling cognitive resources, or increasing accountability for the final choice. The second is *information about the opponent*. Providing information about an opponent may help a player form a more structured view of the opponent's likely behavior and thereby improve strategic reasoning. Information may work through a purely cognitive channel, by making relevant traits more salient. It may also operate through a more affective channel if it induces participants to view the situation from the perspective of their opponent. We test both channels.

We investigate these mechanisms through two separate experiments based on two-person level-k guessing games. In the first experiment, conducted with small entrepreneurs in Ghana, we test whether decisions made in teams display greater strategic sophistication than decisions made individually. In the second experiment, conducted with university students in Tanzania, we test whether strategic reasoning changes when participants receive factual information about their opponent, view a short introductory video of the opponent, or receive both. The two experiments are connected by a common objective: identifying conditions under which individuals engage in deeper reasoning about others' choices.

To study these questions, we rely on the generalized two-person guessing game of Costa-Gomes and Crawford (2006), which extends the original beauty-contest framework of Stahl II and Wilson (1994) and Stahl and Wilson (1995). This design is particularly useful because observed choices can be mapped into different depths of strategic reasoning. A level-0 player chooses randomly. A level-1 player best responds to random play. A level-2 player best responds to level-1, and so on. The game therefore provides a tractable way to measure how many steps of strategic reasoning participants appear to take when making a decision.

The Ghana experiment studies whether teamwork improves strategic sophistication. Participants are small entrepreneurs in Accra who play a sequence of guessing games either individually or in teams of four. Within teams, one participant is randomly assigned the role of captain and is responsible for submitting the team's final answer for that game, while other team members are encouraged to also submit their own individual guesses. This design allows us to distinguish between two possible channels. Team interaction may improve reasoning by exposing all members to additional viewpoints and discussion, in which cases all team guesses should indicate increased strategic sophistication. Alternatively, the effect may operate primarily through responsibility: being made accountable for a collective decision may induce the captain to think more carefully and exert more cognitive effort, in which case only decisions made by captains should be more strategic.

The Tanzania experiment studies whether information about the opponent affects strategic reasoning. Participants play a sequence of guessing games under different information conditions. In one treatment, they receive brief factual information about the opponent, namely gender, field of study, and grade performance. In another, they watch a short video in which the opponent introduces himself or herself and shares broadly relatable personal facts (Paluck et al., 2019; Andries et al., 2024). In a third treatment, factual information and video are combined. We also elicit participants' beliefs about their opponents' likely choices, which allows us to study not only own strategic reasoning but also the degree of strategic sophistication participants attribute to others.

The two experiments speak to distinct but related ideas. The Ghana design focuses on whether strategic reasoning can be improved through collective deliberation and responsibility within the decision-making process itself. The Tanzania design focuses on whether strategic reasoning can be improved by changing how players think about the person they are facing. Both treatments are related because they both rely on social interactions to activate perspective taking, a mechanism believed to play a key part in achieving strategic sophistication (e.g., Kalla and Broockman, 2023; Enke et al., 2022).

Casanelles and Gonçalves (2026) show that material incentives affect strategic behavior by reducing payoff-dependent mistakes and by increasing the effort devoted to reasoning. Our findings indicate that strategic reasoning is also malleable through mechanisms other than incentives. In Ghana, team captains display substantially greater strategic sophistication than individual players, while ordinary team members show smaller gains. This pattern suggests that the main effect of teamwork operates through responsibility for the team's outcome rather than through broad-based learning within the team. In Tanzania, factual information about the opponent increases both the level of strategic reasoning that participants attribute to their opponent and their own level of strategic reasoning. By contrast, the video treatment on its own has little effect. However, when combined with factual information, the video strengthens the impact of that information. This pattern suggests that cognitively relevant information is the main driver of improved strategic reasoning in that case, but that more personal or affective exposure can enhance its effectiveness. Furthermore, when participants are asked to guess their opponent's action before making their own choice, the effect of providing information is stronger on the participant's own strategic sophistication and on the level of strategic sophistication they attribute to their opponent. We also show that this stronger effect arise entirely because participants best respond to what they believe their opponent will do.

Our paper contributes to the economics literature on strategic thinking and iterative reasoning in games. Building on well-known evidence of heterogeneity in initial strategic responses and level-k behavior, we move beyond documenting cross-sectional variation and instead study interventions that shift strategic sophistication itself (Costa-Gomes and Crawford, 2006; Crawford et al., 2013; Fragiadakis et al., 2016). In doing so, we speak to foundational experimental work on guessing games and boundedly rational "models of other players", as well as to adjacent iterative-reasoning frameworks such as cognitive hierarchy (Nagel, 1995; Stahl II and Wilson, 1994; Stahl and Wilson, 1995; Camerer et al., 2004). We also connect to research that explicitly compares strategic performance under individual versus group decision-making, which is particularly relevant for interpreting our teamwork and responsibility treatments (Kocher and Sutter, 2005; Kocher et al., 2006).

Second, our paper contributes to the literature on beliefs and reasoning about others by showing that the *type* of signal received about an opponent affects both beliefs and strategic choices (Gill and Rosokha, 2024). This relates to economics work that measures stated beliefs and studies how well observed play can be interpreted as best responses to those beliefs (Costa-Gomes and Weizsäcker, 2008). It also connects to behavioral models in which agents systematically mis-handle the mapping from others' information or types into predicted actions (Eyster and Rabin, 2005). More broadly, because our opponent profiles provide (or deliberately withhold) information that can be interpreted as a signal of strategic skill, our results relate to the economics of signaling and to experimental evidence that salient identity cues can shape expectations and behavior in strategic interactions (Spence, 1973; Fershtman and Gneezy, 2001; Hoff and Pandey, 2006).

Third, our paper relates to the literature on perspective taking and empathy. In psychology and economics, perspective taking has been linked to universalism, empathy, and prosocial behavior (Krebs, 1975; Davis, 1996; Batson, 2009; Enke et al., 2022; Kalla and Broockman, 2023; Andries et al., 2024). A related literature shows that social contact can affect attitudes and behavior toward others (Pettigrew and Tropp, 2006; Paluck et al., 2019), and economics-based experimental work highlights that informational cues can sometimes *interact* with vivid experiential or contact-based exposure to amplify empathic responses (Andries et al., 2024; Scacco and Warren, 2018). Our findings suggest that these ideas also matter in strategic settings, but in a nuanced way: affective cues alone are not sufficient to improve strategic reasoning, yet they can amplify the effect of cognitively relevant information.

Finally, our results also connect to a broader literature on the relationship between emotions and reasoning. Prior work suggests that emotions can sometimes hinder deliberation and induce less normatively correct reasoning, but may in other contexts motivate attention and cognitive effort (Schwarz, 2000; Blanchette and Richards, 2004; Jung et al., 2014; Loewenstein and Lerner, 2003; Rick and Loewenstein, 2008; Blanchette and Richards, 2010; Blanchette et al., 2017; Damasio, 2006; Dolan, 2002). Our Tanzania findings are consistent with this more nuanced view: affective cues do not independently improve strategic reasoning, but they appear to strengthen the effect of factual information when the latter is present.

The remainder of the paper proceeds as follows. Section 2 presents the conceptual framework and estimation approach. Section 3 reports the results from the Ghana experiment on teamwork and responsibility. Section 4 presents the Tanzania experiment on factual and visual information about opponents. Section 5 concludes.

2 Conceptual framework

2.1 Experimental Design

Our experimental design is based on the beauty contest game. In the original two-person beauty contest of Stahl II and Wilson (1994), player i is asked to guess a value $y_i \in [a, b]$ equal to γ times the value $y_j \in [a, b]$ guessed by opponent j. The closer y_i is to γy_j, the higher is i's payoff. The uninformative prior of a player who assigns no strategic thinking to the opponent is distributed uniformly over the admissible interval $[a, b]$. The best response given this belief is $y_i = \gamma \frac{a+b}{2}$. Observing this choice indicates that player i chooses strategically, but assigns no strategic reasoning to opponent j. Let us denote this as level-1 thinking or L^1. The best response to L^1 by player j is $y_j = \gamma L^1$, which we denote as L^2, and similarly the best response to L^2 is $L^3 = \gamma L^2$. Choices L^2 and L^3 ascribe increasing levels of strategic thinking to the opponent. In contrast, a non-strategic player chooses at random any value in $[a, b]$. This strategy is denoted as L^0.[1]

Costa-Gomes and Crawford (2006) generalize this setting to multiple rounds indexed by t that are played by pairs of players i and j with multipliers γ_{it} and γ_{jt} and intervals $[a_{it}, b_{it}]$ and $[a_{jt}, b_{jt}]$ that vary across players and rounds. Both i and j know the intervals $[a_{it}, b_{it}]$ and $[a_{jt}, b_{jt}]$ and the multipliers γ_{it} and γ_{jt} but not the choice of the other player. The payoff to i is highest when $y_{it} \in [a_{it}, b_{it}]$ is closest to $\gamma_{it} \times y_{jt}$ with $y_{jt} \in [a_{jt}, b_{jt}]$ – and similarly for player j. Level-k strategies L^0, L^1, L^2, and L^3 and index S_i are defined as before, except that in this case they differ between players i and j and vary across t rounds:[2] A simple index of level-k strategic sophistication of a player can therefore be constructed as:[3]

$$S_{it} = I(k_{it} = 0) + 2 * I(k_{it} = 1) + 3 * I(k_{it} = 2) + 4 * I(k_{it} = 3) \tag{1}$$

where $I(.)$ is an indicator function equal to 1 if the equality holds and 0 otherwise, the four possible events are mutually exclusive and their probabilities sum to 1. The higher S_i, the higher the level of strategic sophistication of player i.

[1]In the original game, $[a, b] = [0, 100]$ and $\gamma = 0.5$ hence $L^1 = 0.25$, $L^2 = 12.5$ and $L^3 = 6.25$.

[2]In the original game, subsequent level-k iterations yield the Nash equilibrium $L^\infty = 0$. In the generalized version, a Nash equilibrium in pure strategies L^∞ does not exist for many games, in which case level-k iterations lead to a limit cycle. For this reason, we ignore Nash strategies from our index.

[3]Costa-Gomes and Crawford (2006) also consider strategies other than level-k. These strategies are ignored here for two reasons: they correspond to a different type of strategic thinking and hence do not fit into a level-k index of increasing strategic sophistication; and, more importantly, our evidence suggests that they do not arise in our empirical setting. More about this later.

2.2 Estimating Strategic Sophistication

The objective of the paper is to estimate the effect of various treatments on strategic sophistication by comparing the values of S_{it} across treatments. In order to do this, we would need estimates of the various identity functions entering the definition of S_{it}. Unfortunately, the identity functions $I(k_{it} = K)$ for $K = 0, 1, 2, 3$ can not be inferred from the data given the overlap between the range of theoretically predicted values between L^0_{it} and L^k_{it} for $k = 1, 2, 3$, and given the occasional overlap between L^k_{it} for $k = 1, 2, 3$. Since by construction L^1_{it}, L^2_{it}, and L^3_{it} always $\in [a_{it}, b_{it}]$, observing $y_{it} = L^k_{it}$ for $k = \{1, 2, 3\}$ does not imply that L^k_{it} was purposefully chosen by player i: it could also have arisen under $L^0_i t$. The same difficulty arises whenever $L^k_{it} = L^{k'}_{it}$ for k and $k' \in \{1, 2, 3\}$: observing $y_{it} = L^k_{it}$ does not tell us whether the level of strategic sophistication of player i at time t is k or k'.

While we cannot infer S_{it} from observed choices, we can nonetheless use the sample of observations for treatment $T = s$ to estimate $\widehat{Pr}(k_{it} = K|T = s)$ for $K = \{0, 1, 2, 3\}$, and use the results to calculate the expected value $E[S_{it}|T = s]$ of the strategic sophistication index as:

$$\hat{S}^s_{it} = \widehat{Pr}(k_{it} = 0|T = s) + 2 * \widehat{Pr}(k_{it} = 1|T = s) \tag{2}$$

$$+ 3 * \widehat{Pr}(k_{it} = 2|T = s) + 4 * \widehat{Pr}(k_{it} = 3|T = s) \tag{3}$$

in a given treatment sample $T = s$. By comparing $\hat{S}^s_{it}$ across different treatments, we can estimate treatment effects on strategic sophistication.

To obtain sample estimates of the four probabilities in index (3), we rely on a well-known property of conditional probabilities, namely that:

$$Pr(y_{it} = Y) = \sum_{K=0}^{3} Pr(y_{it} = Y|k_{it} = K)Pr(k_{it} = K) \tag{4}$$

where the unconditional probability $Pr(y_{it} = Y)$ can be estimated from its sample equivalent, the values of $Pr(y_{it} = Y|k_{it} = K)$ are known from the model for each Y and K, and the $Pr(k_{it} = K)$ for $K \in \{0, 1, 2, 3)$ are unknown values to be estimated. We estimate equation (4) by OLS as follows. We start by expanding each actual observation into $N_t = b_{it} - a_{it}$ counterfactual observations covering all possible values of $y_{itn} \in [a_{it}, b_{it}]$.[4] We then create a dummy variable $d_{itn} = 1$ if $y_{itn} = 1$ if $y_{it} = n$ and 0 otherwise. By construction, for each sample observation it, $d_{itn} = 1$ for only one value n, and 0 otherwise. We then use the model to compute, for each of the N_t observations, counterfactual values of $Pr(y_{it} = y_{itn}|k_{it} = K)$ for each valid value $y_{itn} \in [a_{it}, b_{it}]$ and each level $K = \{0, 1, 2, 3\}$. We then estimate the following regression by OLS:

$$\begin{aligned} d_{itn} = \beta_0 Pr(y_{it} = n|k_{it} = 0) + \beta_1 Pr(y_{it} = n|k_{it} = 1) \\ + \beta_2 Pr(y_{it} = n|k_{it} = 2) + \beta_3 Pr(y_{it} = n|k_{it} = 3) + u_{itn} \end{aligned} \tag{5}$$

on the sample of observations for treatment T_s. Since $\sum_{K=0}^{3} Pr(k_{it} = K) = 1$ by assumption and $d_{itn} = 1$ only once over the $[a_{it}, b_{it}]$ interval, it follows that, mechanically, $\sum_{K=0}^{3} \beta_K = 1$ when model (5) is estimated via OLS without intercept. Identification comes from the proportion of choices falling within a restricted L_{it}^k range for $k = \{1, 2, 3\}$ relative to what would occur by chance under L_{it}^0. The resulting values of $\hat{\beta}_K$ are our method-of-moments estimates of $\widehat{Pr}(k_{it} = K)$ for $K = \{0, 1, 2, 3\}$ for a given treatment in a given session. This process is repeated for all treatments. An estimate of the expected sophistication index in a specific treatment $T = s$ can then be constructed as:[5]

$$\hat{S}_{it}^s = \hat{\beta}_0^s + 2\hat{\beta}_1^s + 3\hat{\beta}_2^s + 4\hat{\beta}_3^s \tag{6}$$

We then stack $\hat{S}_{it}^s$ for all observations and treatment and regress the resulting variable on treatment dummies to estimate the effect of the different treatments on average strategic sophistication.

The estimated values of $\widehat{Pr}(k_{it} = K)$ for $K = \{0, 1, 2, 3\}$ can also be used to

[4]This approach is inspired of estimation models for discrete choice models with common regressors and coefficients (Train, 2002).

[5]Since probabilities cannot be negative, if one of the $\hat{\beta}_K$ happens to be < 0, we re-estimate the regression without the offending regressor and set the missing coefficient to 0 for the purpose of constructing the index. This is done to avoid biasing the index.

obtain more precise inference about each observed choice using Bayes Law. To recall:

$$Pr(k_{it} = K|y_{it} = Y) = \frac{Pr(y_{it} = Y|k_{it} = K)Pr(k_{it} = K)}{\sum_{L=0}^{3} Pr(y_{it} = Y|k_{it} = L)Pr(k_{it} = L)} \tag{7}$$
$$\text{with} \sum_{L=0}^{3} Pr(k_i = L) = 1$$

As before, each $Pr(y_{it} = Y|k_{it} = K)$ term is given by our model and each $Pr(k_{it} = K)$ has been estimated using regression (6). Applying this formula to each observation yields refined estimates of $\widehat{Pr}(k_{it} = K|y_{it} = Y)$ for $K = \{0, 1, 2, 3\}$ that vary across observations within treatment. These estimates can in turn be used to construct a choice-specific index of strategic sophistication as:

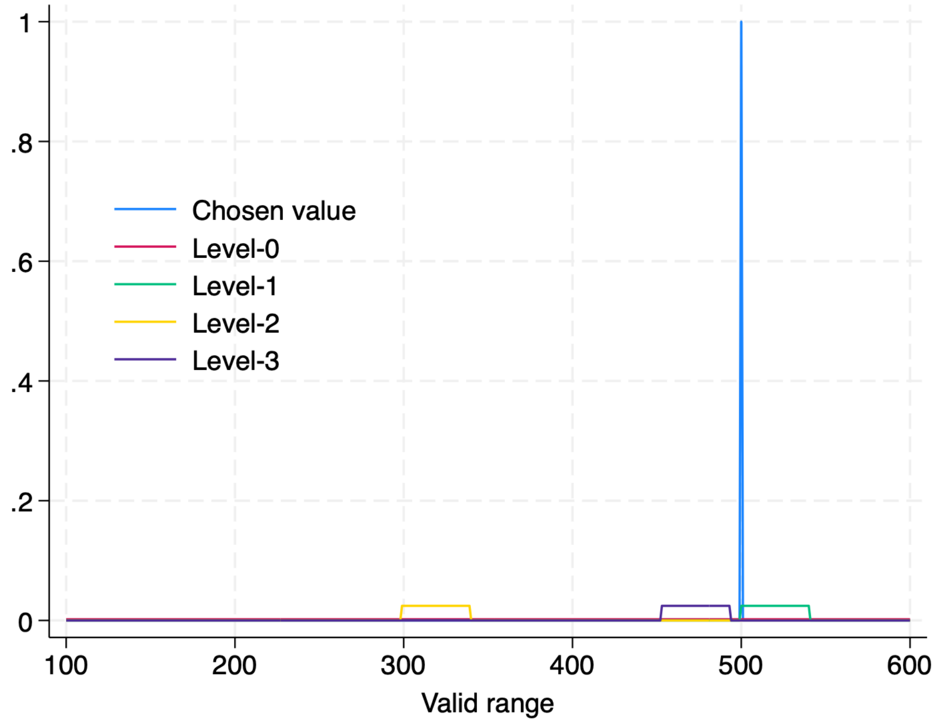
$$\begin{aligned}\hat{S}_{it}^{r} = \widehat{Pr}(k_{it} = 0|y_{it} = Y) + 2\widehat{Pr}(k_{it} = 1|y_{it} = Y) \\ + 3\widehat{Pr}(k_{it} = 2|y_{it} = Y) + 4\widehat{Pr}(k_{it} = 3|y_{it} = Y)\end{aligned} \tag{8}$$

This observation-varying index can then be used for more detailed analysis of the data, such as additional treatment effect analysis or heterogeneity analysis.

In practice, we refine the above methodology in two ways. First, instead of associated a single value $y_{it}(K)$ with a single K level, we broaden it to a small interval of size κ centered on the target value $y_{it}(K)$. This is done to allow for small mistakes made by experimental participants. Figure 1 provides an illustration for one specific observation. Participants must guess an integer value in the valid range. The chosen value is shown as a spike at d_{itn}. Level-0 can take any value over the valid range with equal probability. Intervals for level-k, $k \in \{1, 2, 3\}$ are shown as well. The height of all intervals is chosen such that all have an integral equal to 1. Intervals can overlap if the same choice can be predicted by different level-k reasoning. Second, we exclude from the analysis the choices that fall outside the allowed $[a_{it}, b_{it}]$ range, since they cannot be rationalized by any of the L_{it}^{k} strategies. These choices are analyzed separately.

In the remainder of the paper we present the results from two separate experiments, one conducted with Ghanaian entrepreneurs in West Africa, and one conducted

Figure 1: Example of intervals for $Pr(y_{it} = n|k_{it} = K)$



Notes: The value chosen by the participant is a spike at $d_{itn} = 1$. Values of $Pr(y_{it} = n|k_{it} = 0) = 1/(b_{it} - a_{it} + 1)$. Values $Pr(y_{it} = n|k_{it} = \{1, 2, 3\}) = 1/\kappa$ centered on L^k_{it} for $k \in \{1, 2, 3\}$.

with university students in Tanzania, East Africa. Both experiments seek to identify treatments that increase strategic sophistication among participants. The first experiment compares individually made choices with choices made by teams. The second experiment provides participants with different types of information about their opponent: a short CV, a short video, or both.

3 Ghanaian entrepreneurs

Strategic thinking is typically seen as a requirement for success in business. This implies that established entrepreneurs ought to be more proficient in level-k games. It is for this reason that, for our first experiment, we selected a population composed of men and women who run their own enterprise. Experimental subjects from such population are difficult to recruit in high-income countries for two reasons: the proportion of self-employed individuals in the working population is relatively small; and the opportunity cost of time of entrepreneurs is high and they seldom accept to participate. To find a large enough willing sample of self-employed participants, we turn to West Africa where entrepreneurship is by far the predominant form of employment. The relatively low level of educational attainment of our target population may hinder its ability to perform well in an abstract game such as a level-k experiment (Dean et al., 2025). While this may result in a low level of measured strategic thinking on average, it nonetheless raises the external validity of our findings relative to university student samples, given that the overwhelming majority of the world population has not gone to college.

3.1 Implementation

In the summer of 2011, we conducted an experiment with small-entrepreneurs recruited in Accra, the capital city of Ghana, West Africa. The participants were randomly selected among respondents to a survey of small entrepreneurs survey conducted in the capital city of Accra (Fafchamps et al., 2014).

Each group of participants was invited to come to a central city location for a half-day of instruction and training. They were introduced to the experimental design, explained the rules of the game, shown how payoffs were calculated, and given a chance to play several games to familiarize themselves with their difficulty. They were provided a per diem for their participation to the training session and asked if they wanted to participate to the experiment proper, at which time they were asked to provide informed consent. Each cohort of willing participants who came to the same training session is regarded as forming a distinct experimental session.

The experiment itself took place over a period of 16 days, with each day dedicated to a single game.[6] At the beginning of the day, participants were given all the relevant parameters for the game of the day, that is, multipliers γ_{it} and γ_{jt} and intervals $[a_{it}, b_{it}]$ and $[a_{jt}, b_{jt}]$. This information was provided to them by SMS over the phone. Participants never played a game with the same parameter set twice. Each day, they were re-paired with a new random opponent from among the other participants in the

[6]One cohort played over 17 days due to a phone network failure on one of the days.

experiment. Participants were instructed to reply to our SMS with their choice y_{it}. They were never told who their opponent was or what the outcome of the day's game was, except that each day they were assigned a different opponent.[7]

3.2 Treatments

In this first experiment, we investigate whether teamwork improves strategic thinking. To this effect, we randomized participants into two treatment conditions: individual play; and team play. We also aim to distinguish between two possible channels through which teamwork increases strategic sophistication: by improving the strategic thinking of each member of the team; or by making someone responsible for the welfare of the team.

These two channels operate through different types of perspective taking. By bouncing information and ideas between members, team play can raise awareness among all team members about the strategic perspective of others, which in turn may increase the level of strategic sophistication through *strategic* (i.e., cognitive) perspective taking: team members start thinking about the other team as also composed sentient individuals capable of rational thought. The second channel relies on picking a team leader who is given final say on the team's choice and therefore is made responsible for the welfare of the team. This can trigger *empathic* perspective taking: the leader invests more attention and effort into the decision because he or she cares about the members of the team, and increased effort in decision making results in more sophisticated choices. To distinguish between these two channels, we randomly rotate the role of team 'captain' to each team member over the course of the experiment; and we collect choices from both the captain and each team member individually.

This design is implemented as follows. In six of the cohorts (sessions), participants played each game individually. This is treatment T_0. In nine of them, participants played each game in a group of four, and one of the four, the 'captain', was randomly assigned the task of reporting by SMS the agreed-upon choice y_{it} decided by the group. Each group member received information by SMS on *one* of the four game parameter

[7] In order to make it difficult for participants to work out who their opponent was by comparing game parameters, participants were matched with participants from another day.

sets (γ_{it}, γ_{jt}, $[a_{it}, b_{it}]$, $[a_{jt}, b_{jt}]$). This means that they had to share parameter information with each other in order to play the game strategically. Only the choice made by the team captain was used to decide payoffs. But group members other than the captain were invited to submit their own individual choice of y_{it} as well, which may differ from that reported by the captain. The composition of each group remained unchanged for the duration of the experiment, but the role of captain rotated each day in random order across the four group members. Treatment T_2 denotes observations in which the participant was team captain while T_1 denotes observations in which the participant was a team member but not team captain. In total there were 221 participants, 69 of whom played individually and 152 in teams. The number of participants varied across cohorts, with a mean of 14.7, a minimum of 5 and a maximum of 28.

All game parameters are presented in Table B2 in Appendix. In order to maximize the information content of participants' choices, we carefully selected each of the 16 parameter sets to be as discriminating as possible between different strategies. This meant eliminating nearly all cases of overlap between choices corresponding to six different strategies: L^1, L^2, and L^3, but also the iterated dominated strategies D^1, D^2, and D^3 of Costa-Gomes and Crawford (2006). This had the side-effect of making the games strategically more difficult than those used by these authors, which contained a lot of overlapping strategic choices, especially concentrated at the upper and lower bounds of the allowed range.

Details on the Ghana experimental protocol are given in the Online Appendix E. At the exchange rate prevailing at the time, lab subjects received the equivalent of \$10 as a fixed show-up fee for participating in the experiment, plus gains of up to \$40 more depending on the outcome of the game.

3.3 Engagement

Participants were recruited among respondents to the self-employment survey conducted by Fafchamps et al. (2014) to study the effect of cash and in-kind grants to micro-entrepreneurs in and around Accra, Ghana (West Africa). They tend to be middle-aged men and women running a small urban business, primarily in retail and services. The only constraint imposed for recruitment was access to a phone capable of sending and receiving SMS messages.

Given that this experiment was implemented over the phone, we collected detailed information at endline regarding participants' engagement with the games and cooperation within teams. Our findings are summarized in Appendix E.4. Taken together, this evidence indicates a very high level of engagement with the experimental games, which is remarkable given that the experiment was conducted while participants were going on with their daily life. There is a moderate proportion of missed guesses, partly due to phone connection problems. But only a small minority of participants failed to report guesses for all 16 games. Importantly, we find that participants assigned to the team treatment report cooperating with each other, with a considerable density of phone and SMS exchanges across team members. From this evidence we conclude that participants did seriously engage with the experiment, enjoyed it, discussed strategy with their team, and perceived the difficulty of the games to be mostly average in both the individual and the team treatments.

3.4 Empirical results

We start by estimating equation (3) for each of the three treatment conditions: individual play T_0; team member T_1; and team captain T_2. Choices that fall outside the allowed range $[a_i t, b_i t]$ are omitted from the estimation since they cannot be accounted for by any of the level-k predictions and thus violate the requirement that $\sum_{K=0}^{3} Pr(k_{it} = K) = 1$. Out-of-range choices are examined separately below.

Regression estimates, presented in Table 1, represent, for each treatment, the probability that choices follow each of the level-k predictions.[8] Each coefficient $\hat{\beta}_K^s$ is an estimate of $Pr(k_{it} = K|T_m)$ for $K \in \{0, 1, 2)$ and treatment $T_m \in \{T_0, T_1, T_2\}$. When we include level-3 choices in the regression, they show up with coefficients that are small

[8]The Table was produced using value of $\kappa = 20$, where κ is the size of the interval around level-1, level-2 and level-3 choices. With a narrower interval $\kappa = 10$, the pattern of treatment effects remains the same but, unsurprisingly perhaps, the inferred probability of random play is higher.

in magnitude and occasionally negative, which would indicate that participants have a slight tendency to stay away from choices that correspond to a level-3 degree of strategic thinking. For this reason, level-3 is omitted from Table 1 – but they can be found in Appendix Table A1.

We see that choices mostly follow level-0 predictions, that is, random play over the allowed interval. This is true for all treatments, but varies systematically with treatment: the inferred probability of random play falls from 84% in individual play T_0 to 82% when the participant is a team member and 78% when the participant is team captain. This probably reflects the high difficulty level of the game parameters we chose. The inferred probability of level-1 choices is estimated to be around 12-14%, with a slight increase across treatments, and the inferred probability of level-2 behavior shows a large increase from 3.8% among individual players to 5.0% for team members and 8.1% for team captains. This indicates that strategic thinking increases in sophistication in the team treatment, but primarily when the participant is randomly assigned to the role of team captain.

Table 1: Treatment effects on $Pr(k_{it} = K)$

	T0	T1	T2
Level-0			
Coefficient	0.843***	0.823***	0.783***
Std. error	0.035	0.033	0.046
Level-1			
Coefficient	0.120***	0.127***	0.136***
Std. error	0.032	0.031	0.043
Level-2			
Coefficient	0.038***	0.050***	0.081***
Std. error	0.012	0.010	0.016

Note: The chosen interval step size κ is 20. Standard errors are clustered at the individual player level. *p < 0.10; **p < 0.05; ***p < 0.01.

Next, we turn to the estimation of equation (8), having constructed the choice-varying index of strategic thinking $\hat{S}^r_{it}$ from the coefficients reported in Table 1.[9] Column (1) presents OLS estimates with standard errors clustered at the participant level.

[9]Very similar results are obtained if we instead use that coefficient estimates with level-3 predictions that are reported in Appendix Table A1.

Columns (2) and (3) uses inverse probability weighting to correct for the possible bias introduced by missing guesses. The two columns differ by the choice of regressors when predicting missing guesses. Column (4) introduces as additional regressor the number of guesses missing from the three team members: if team participation is essential for team play to improve strategic sophistication, then a reduction in this participation should lower decision quality.

The results, presented in Table 2, are similar across columns 1, 2 and 3, suggesting that the results are unlikely to be an artifact of missing guesses. We also find that the number of missing team guesses has the anticipated sign but is not significant – providing no support to the idea that full member participation is essential to the success of team play. Results confirm that both treatments T_1 and T_2 produce a large and significant increase in participants' strategic sophistication. The effect of randomly being assigned the role of team captain is particularly large: given that index $\hat{S}^r_{it}$ is normalized to 1 for level-0 play (i.e., no strategic thinking) and that the value of the index is 1.14 for individual play, an increase of around 0.1 index points represents a 69-70% increase in strategic thinking, which is remarkable. The increase associated with being a team member (but not a captain) is marginally significant but much smaller in magnitude.

Since the role of team captain is randomized, in equal proportion, among all members over the course of the experiment, increased strategic sophistication cannot result from a leader selection effect, our findings suggest that teamwork increases the strategic sophistication of team decisions primarily through increased effort by the team leader. We call this a responsibility effect: by forcing participants to make decisions that affect others, we induce them to take better care which, in this context, means thinking more deeply about the strategic implications of one's choices. This implies that strategic sophistication can be raised even in low education populations by making people responsible for the effect of their choices on others.

We also investigate the effect of the treatments on out-of-range play. To this effect, we set index $\hat{S}^r_{it} = 0$ when a guess is out of range, that is, either $> b_{it}$ or $< a_{it}$, and we plot the cumulative distribution function of $\hat{S}^r_{it} = 0$ for each of the three treatments.

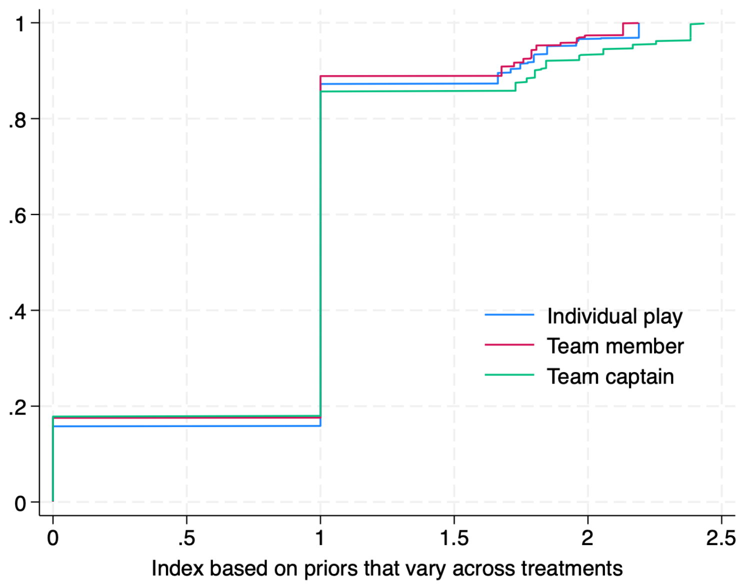
Table 2: Treatment effects on the strategic thinking index $\hat{S}^r_{it}$ (Ghana)

	1	2	3	4
Team member				
Coefficient	0.0236	0.0252*	0.0248*	0.0252*
Std. error	0.0150	0.0150	0.0150	0.0150
Team captain				
Coefficient	0.0976***	0.0994***	0.0987***	0.1113***
Std. error	0.0223	0.0224	0.0223	0.0278
N.missing guesses when team captain				
Coefficient				-0.0193
Std. error				0.0244
Intercept				
Coefficient	1.1432***	1.1413***	1.1411***	1.1413***
Std. error	0.0118	0.0117	0.0117	0.0117
N	2991	2991	2991	2991

Note: The dependent variable is the index of strategic thinking. It takes value 1 for random play (level-0), value 2 for level-1 thinking, etc. Guesses outside the allowed range are excluded from the regression. Column 1 includes no weighting. Column 2 includes Inverse Probability Weights (IPW) for missing guesses. Column 3 includes an IPW correction with more predictors. Column 4 adds a regressor for the number of missing guesses by team members when captain. Standard errors are clustered at the individual player level. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

The results, shown in Figure 2, visually confirm that treatment T_1 and T_2 produce large increase in strategic thinking of level-1 and 2. But they also *increase* the proportion of out-of-range play (index=0) relative to T_0. Why this is the case is unclear. One possible interpretation is that T_1 and T_2 induce some participants who are playing random in T_0 to pay more attention to the two multipliers and to the opponent's choice range, and this leads them to make the mistake of overlooking their own allowed range. In any case, the effect is too small in magnitude to be economically significant.

Figure 2: CDF of the index of strategic thinking, by treatment (Ghana)



Notes: The graph shows the cumulative distribution of the index of strategic thinking. Guesses outside the allowed range are included and take value 0.

4 Tanzanian university students

We conducted a second experiment with 200 university students. The focus of this experiment is on the effect that being provided with factual or visual information about an opponent has on strategic thinking. While we do not seek to measure perspective taking directly, these treatments are chosen because perspective taking is a potential causal mechanism by which strategic thinking improve – e.g., factual and visual input can affect strategic sophistication by raising participants' level of attention to the thinking process of others.

To gather additional evidence about perspective taking, we also elicit participants' beliefs about the choice made by their opponent in each game. We do not, however, ask participants to guess what they think their opponent believe their choice was – something that, following the literature (Agranov et al., 2012; Georganas et al., 2015), would be needed to measure perspective taking directly. The rationale for not asking this question is that it is artificial in our setting, given that participants are paired with an opponent from a previous session. We also worry that the fact of asking this question may itself affect perspective taking, thereby affecting measurement itself. To circumvent this concern, we choose instead to test whether strategic sophistication increases when a participant is asked to guess the choice made by their opponent. This is accomplished by randomly varying whether participants are asked to guess before or after they report their own choice. If measured strategic sophistication increases when the guess is asked before, this constitutes evidence that asking questions on beliefs about opponents affects perspective taking via a kind of Hawthorne effect.

We also cross this question Q with the four information conditions. If the effect of the timing treatment Q works purely through an attention channel, there is no reason to expect the effect of Q on strategic sophistication to vary with the amount of information provided on opponents: if anything, providing information may distract the participant by overloading his or her cognitive capacities when combined with a cognitively demanding question on beliefs. But if information treatments operate through a perspective taking channel, asking participants to guess the action of the opponent may prime them into using opponent information to make a more sophisticated guess, a process that would, in turn, increase the sophistication of their own choice.

4.1 Implementation

The experiment took place at the University of Dar es Salaam, Tanzania, where an active experimental lab supports a rich program of work in behavioral economics. This is an ideal context for this study given that it is a relatively young universitywith a student body characterized by a high level of ethnic and religious diversity – a context potentially less conducive to perspective taking. As such, it is representative of the bustling city of Dar es Salaam, the economic capital of Tanzania, which has grown from less than a million inhabitants in 1978 to more than seven million now by attracting migrants from all over a country twice the size of California. These features also make Dar es Salaam and its University representative of much of the Global South. The subject pool, described below, is otherwise akin to students in similar departments across the developed and developing world.

The basic experimental setup is similar to the one used in Ghana. Participants perform a series of guessing games, 22 in this case, in which they guess a number as close as possible to an anonymous opponent's guess times a multiplier. Participants play individually and, in each task, they are each paired with a new, randomly selected opponent. At the end of each game, participants are asked to guess the choice made by their opponent.

At the end of training, participants took a quiz to gauge their understanding of the game. Most participants correctly answered 5 or 6 of the 7 questions, which indicates a reasonable level of understanding of the mechanism of the game among our subjects.

4.2 Treatments

Participants play each game under one of four possible information conditions. Under the control condition, they are provided no information about their opponent other than it is someone recruited from the same student population as them. There are three information treatments, randomized across games and sessions: Factual Information (T1); Video (T2) and Combined Video and Factual Information (T3). The set of treatment conditions varies across sessions: in some of them, participants play games in the control condition plus T1 and T2, while in other sessions, they play in the control condition plus T1 and T3. Within sessions, each participant plays a series of games in the three different conditions used in that session. We now outline each treatment condition in detail.

The first treatment reveals *factual information* (T1) about the opponent: gender (male/female); field of study (STEM/non-STEM subject); and Grade Point Average (high grade/low grade).[10] Crossing these three traits produces, eight possible opponent profiles, which we match with the characteristics of eight actual players in the opponent sessions (described below).

Our second treatment shows *videos* (T2) of the opponents introducing themselves in less than one minute and sharing some fun facts about themselves, which are easily and broadly relatable but carry no specific signals about their strategic sophistication. Examples include their favorite dish.[11] The objective is to trigger a reaction based on emotive and affective empathy. As in T1, we have eight opponents, who were selected to have the same characteristics as the eight opponents whose information was shared in T1 (male/female, stem/non-stem, high-grade/low-grade), but they are *not* the same opponents.[12] As in T1, therefore, we generated eight different videos to match the eight possible factual profiles. Watching the videos was a mandatory component of the experiment, and participants could not skip this. As an attention check, after being shown each video, participants were asked to confirm they had watched it otherwise they could not move forward. In addition, the activity was conducted in a classroom and in small groups, under the supervision of trained enumerators – who were explicitly instructed to

[10] A high grade was a First class attained in the previous academic year. A low grade was a Lower Second class.

[11] Having several opponents of both genders and backgrounds ensured sufficient diversity in the fun facts that were provided. This ensures that the average impact of the video documented below cannot be driven by any specific type of information shared in the videos.

[12] Using the same opponents in T1 and T2 would have resulted in participants seeing the same opponent twice across games played under T1 and T2, respectively.

ensure that respondents were watching the videos. Based on the feedback we received, the videos were generally enjoyable and respondents were genuinely curious and motivated to watch them.

Our third treatment (T3) *combines* providing factual information about the opponent and showing the video. The profiles used for this treatment are the same as in T2, but we now show factual written information as in T1 (gender, field of study, GPA) in addition to playing the video.

Out of the 22 games participants played in each session, the first four (1-4) and the last two (21-22) were control tasks (C). The two treatments used in a session (i.e., wither T1 and T2, or T1 and T3) were randomly distributed over the remaining sixteen games (5-20). This ensures that the combination of eight opponent profiles and two treatment arms per subject results in exactly 16 unique *opponent* $\times$ *treatment* combinations per subject. The order in which treatments appear was randomized between participants.

The four information treatments were randomly crossed with the priming treatment Q varying the timing of the question on guessing the opponent's choice. Half of the time, the participant was asked to make their own choice first and then to guess their opponent's choice (Control condition), or to guess first and make their own choice second (Treatment condition). As already explained, this allows us to test whether priming participants to think about their opponent's choice increases the effect of information on strategic sophistication.

4.3 Experiment Implementation

The experiment was conducted over 10 sessions with 20 participants in each. Participants in sessions 1 to 7 played T1 and T3 games in addition to control games, while those in sessions 8 to 10 played T1 and T2 games plus control games. The games were played on digital tablets. Each session started with an explanation by an enumerator to the entire group of participants, after which each player completed the experiment individually. The initial explanation included standard examples to maximize understanding. After the 22 guesses, participants completed a short survey. The full experimental instructions and some sample screenshots from the survey are included in Appendix F.1 and F.2.

Participants in a session did not play against other participants in the same session. Since asking participants to produce videos could potentially affect their behavior in the experiment, we instead ran two auxiliary *opponent sessions* at the beginning of the experiment to generate opponent choices that could be paired to the choices of participants in the main sessions so as to generate winnings. At the end of these auxiliary sessions, we asked participants to produce the videos we used for the treatments.[13] In the two auxiliary sessions, participants were shown no information about their opponents (i.e., they played all the games under the control condition).[14] This ensures that opponents are not given an advantage over the main participants in terms of strategic reasoning.[15] We then matched participants in the main experimental sessions with opponents from these two auxiliary (opponent) sessions – which are not themselves used in the analysis presented here.

To summarize, players in the two auxiliary sessions are used as counterparts to all the players in the ten main sessions. This design offered the added benefit of avoiding deception: it allowed us to truthfully tell the main participants that they were playing against random anonymous students from their university who had no information about them.

[13] To avoid a potential confound due to the cold-hot empathy gap (the idea that empathy may depend on the emotional state or situation one is in, and on how distant one is from the object of one's empathy, see Loewenstein (1996)), participants in the main session were truthfully told that the opponents were other random students. We did not mention that opponents came from auxiliary sessions, which might have increased perceived social distance. Nonetheless, we acknowledge that we cannot fully control for this mechanism, as participants may have noticed that the participants shown in the videos were not present in their sessions.

[14] participants in the main sessions were informed about this to ensure full knowledge of the experimental conditions.

[15] For convenience, since the opponent sessions took place before the main sessions, we calculated the payoffs of participants in the opponent sessions by matching them with one another within those sessions. As those players were told that they were playing against random, anonymous peers and they had no information about their opponents, this was not deceitful. Importantly, this approach avoided having to wait for the main sessions to take place before paying participants in the opponent sessions. Such a delayed payment may have been problematic as it may have generated differences in behavior, undermining our design.

4.4 Sampling

We recruited participants from various undergraduate (second year and above, with prior university results) and graduate degrees at the University of Dar es Salaam. Recruitment was primarily conducted through invitations sent to student representatives. Participating students came from a broad range of academic disciplines: Engineering (31), Law (13), Economics and Statistics (35), Other Social Sciences (29), Humanities (28), Natural Sciences (30), Agricultural and Food Sciences (14), Business (11), Kiswahili (6), and Development Studies (3). A total of 200 students participated in over 10 sessions. Each session lasted approximately 2 hours, and the average payoff per respondent was approximately TZS 75,000 (27.8 USD), plus a show-up payment of TZS 10,000 (3.7 USD). These payoffs were significant and amounted, on average, to four times the average daily wage in Tanzania at the time of the experiment.

Participants to the two auxiliary opponent sessions described above were recruited from the same student population. They were selected in a stratified manner to match the eight possible opponent profiles in T1, i.e.: male/female, stem/non-stem, high-grade/low-grade.[16]

4.5 Empirical results

Since treatments vary within subject in a session, we can estimate equation (3) separately by session for each of the information treatments in that session, keeping in mind that the treatment mix varies across sessions (T1 & T3 for sessions 1-7 and T1 & T2 for sessions 8-10). This yields treatment-specific $\hat{\beta}_K^s$ estimates of $Pr(k_{it} = K|T_m)$ that vary across sessions – an approach that is equivalent to allowing session fixed effects in our analysis. While this method allows more refined estimates of $Pr(k_{it} = K|T_m)$, they result in too

[16]Specifically, out of the 16 opponents, we had two males and twos females enrolled in the Bachelor of Education (Science) with a first-class GPA in the previous year; two males and two females enrolled in the Bachelor of Education (Science) with a lower second class GPA in the previous year; two males and two females enrolled in the Bachelor of Education (Arts and Humanities) with a first-class GPA in the previous year; two males and two females enrolled in the Bachelor of Education (Arts and Humanities) with a lower second class GPA in the previous year. For one of the two profiles in each combination, we also had a video.

many estimates to present in a meaningful Table.[17] For this reason, we move directly to our main analysis, i.e., the effect of the treatments on the index of strategic sophistication in equation (8).

Results are presented in Table 3. As before, we drop choices that fall outside the allowed range $[a_{it}, b_{it}]$ since.[18] Individual participant fixed effects are included in the estimation. From the first column, we see that the factual information treatment has a have strong positive effects on the index of strategic sophistication $\hat{S}^r_{it}$, whether in isolation (T1) or combined with the video (T3). The video treatment (T2) on its own has not effect on the index. Pairwise t-tests between pairs of treatments indicate that estimated coefficients for T1 and T3 are significantly different at the 5% level from that of T2. The coefficient of T3 is larger in magnitude than that of T1, suggesting that the video may boost the effect of factual information alone. But this difference is not statistically significant (p-value of 0.402).

In columns 2 and 3, we split the sample between observations in which participants choose their own value before guessing their opponent's, and observations in which they choose after guessing. Factual information treatments T1 and T3 remain significant in both cases, but estimated treatment effects are much larger when participants choose after guessing. As shown in column 4, these differences are statistically different at the 5% level. This indicates that the priming effect of treatment Q reinforces the factual information treatment.

We also note a negative but small and marginally significant coefficient on the round variable, suggesting that the sophistication index falls slightly as participants progress through the experiment. This is the opposite of a learning effect. While we cannot be for certain, it could result from cognitive load depletion over the course of the

[17] A small number of session x treatment regressions were estimated with only L0, L1 and either L2 or L3, to eliminate negative coefficients in the calculation of the strategic sophistication index. Because negative coefficients tend to be very small in magnitude, this correction does not affect the results of the analysis. But they eliminate the possibility of data artifacts creeping into the calculation of the sophistication index.

[18] The proportion of missing guesses is much lower in the Tanzanian experiment (2%) than in the Ghana experiment (17%). Consequently, we omit the inverse probability weighting correction that we included in Table 2. We do, however, note that, as in the Ghana experiment, the information treatments also generate an increase in out-of-range choices. But the increase is very small in magnitude.

session.

Table 3: Treatment effects on the strategic thinking index $\hat{S}^r_{it}$ (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.121***	0.100*	0.143***	0.0243
	(0.027)	((0.042)	(0.033)	
T2 - Video	0.004	-0.012	0.021	0.1327
	(0.044)	(0.062)	(0.062)	
T3 - Factual + Video	0.147***	0.116*	0.179***	0.0160
	(0.030)	(0.047)	(0.038)	
Round	-0.004**	-0.005*	-0.004	
	(0.001)	(0.002)	(0.002)	
Constant	1.455***	1.454***	1.456***	
	(0.020)	(0.028)	(0.029)	
T1=T3 (p)	0.402	0.725	0.397	
T2=T3 (p)	0.008	0.109	0.032	
T1=T2 (p)	0.021	0.144	0.069	
N	4233	2111	2122	

Notes. The top panel of the Table shows the regression results. The dependent variable is the index of strategic thinking. It takes value 1 for random play (level-0), value 2 for level-1 thinking, etc. Guesses outside the allowed range are excluded from the regression. Individual participant fixed effects are included. Column 1 combines all within-range observations. Columns 2 and 3 include observations in which the participant chooses before or after being asked their belief about the opponent, respectively. A round variable is included to capture any learning effect occurring over the session. Standard errors are clustered at the individual player level. *p < 0.10; **p < 0.05; ***p < 0.01. Column 4 reports *p*-values of a Wald test comparing coefficients between column 2 and 3. The second panel of the Table shows the *p*-values of pairwise t-tests of one treatment vs. another.

Next we examine the effect of the treatments on the beliefs that participants report regarding the choice made by their opponent. In Appendix C we show that, as predicted by theory, participants make choices that are on average more sophisticated than the level of strategic thinking they ascribe to their opponent.

To estimate treatment effects on beliefs about the opponent, we proceed in the same way as we did for the participant's own choice, except that we mirror all the game parameters (i.e., a_j instead of a_i, etc). This gives the implied value of the sophistication index $\hat{S}^r_{jt}$ of the opponent j. Estimation results are presented in Table 4. The general pattern of results in column 1 resembles Table 3: the factual information treatments T1 and T3 have a significant effect on $\hat{S}^r_{jt}$ while the video treatment T2 has no effect. We also note that T3 has a stronger effect than T1, again suggesting that while the video treatment has no effect on its own, it reinforces the factual information treatment. This difference, however, is too small to be statistically significant, unlike in Table 3.

Columns 2 and 3 mirror Table 3: they compare the effect of the information treatments depending on whether beliefs about the opponent are elicited after or before the participant chooses his or her own action. As in 3, we find that the effects of treatments T1 an T3 are much stronger in columns 3, that is, when the participant is invited to reflect on the choice made by the opponent before making their own choice, and this difference is significant for T1. This suggests that, when participants are asked to guess the action of their opponent before making their own, they think more deeply about their opponent and ascribe them a choice that implies a higher level of strategic sophistication. This evidence is consistent with perspective taking being the channel by which the treatments increase strategic sophistication.

Table 4: Treatment effects on the implied strategic sophistication $\hat{S}^r_{jt}$ of opponent (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.066**	0.018	0.113***	0.0409
	(0.023)	(0.032)	(0.032)	
T2 - Video	0.009	-0.047	0.065	0.1864
	(0.042)	(0.058)	(0.059)	
T3 - Factual + Video	0.094**	0.066	0.122**	0.3285
	(0.028)	(0.043)	(0.037)	
Round	-0.002	-0.002	-0.003	
	(0.001)	(0.002)	(0.002)	
Constant	1.384***	1.419***	1.349***	
	(0.018)	(0.028)	(0.022)	
T1=T3 (p)	0.307	0.198	0.819	
T2=T3 (p)	0.077	0.088	0.408	
T1=T2 (p)	0.186	0.249	0.452	
N	4247	2116	2131	

NOTES. The dependent variable is the implied strategic sophistication $\hat{S}^r_{jt}$ of opponent. It takes value 1 for random play (level-0), value 2 for level-1 thinking, etc. The unit of analysis is the individual guess. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

If the information and priming treatments operate by inducing participants to think more deeply about the strategic implications of the game, we would also expect the effect of these treatments to be stronger for subjects who initially demonstrated a limited mechanistic understanding of the game. Put differently, participants who initially did not engage fully with the intricacies of the game may have been induced to think more deeply once the information and priming treatments triggers more perspective taking. To investigate this possibility, we re-estimate columns 2 and 3 of Table 3 separately for those participants who initially demonstrated a good mechanistic understanding of the game and those who did not.

Table 5: Treatment effects on index $\hat{S}^r_{it}$ by level of understanding of the game (Tanzania)

	(1)	(2)	(3)	(4)
	High underst.		Low underst.	
What comes first:	Choice	Guess	Choice	Guess
T1 - Factual Info	0.129***	0.151***	-0.009	0.113*
	(0.047)	(0.038)	(0.093)	(0.066)
T2 - Video	0.023	0.044	-0.170	-0.074
	(0.066)	(0.070)	(0.168)	(0.131)
T3 - Factual + Video	0.160***	0.200***	-0.045	0.098
	(0.054)	(0.045)	(0.091)	(0.058)
Round	-0.005**	-0.004	-0.005	-0.004
	(0.002)	(0.003)	(0.004)	(0.005)
Constant	1.434***	1.446***	1.526***	1.496***
	(0.031)	(0.032)	(0.061)	(0.063)
T1=T3 (p)	0.548	0.334	0.685	0.791
T2=T3 (p)	0.119	0.064	0.526	0.230
T1=T2 (p)	0.206	0.159	0.420	0.198
N	1675	1674	436	448

NOTES. The dependent variable is the index of strategic thinking $\hat{S}^r_{it}$. It takes value 1 for random play (level-0), value 2 for level-1 thinking, etc. The unit of analysis is the individual guess. High understanding means 5 or more correct responses on the quiz, low understanding means 4 or fewer correct responses. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

The results are shown in Table 5. Since the participants' level of mechanistic understanding of the game is not experimentally assigned, it should not be given a causal interpretation. We nonetheless expect that participants with a lower mechanistic understanding of the game would also demonstrate a lower level of strategic sophistication. This is indeed what we find. More importantly for our purpose, we also find that, among participants with a good understanding of the mechanics of the game, the factual information treatments T1 and T3 have a strong causal effect on strategic sophistication even without priming treatment Q: priming increases $\hat{S}^r_{it}$ a bit, but not by much. In contrast, among low understanding participants, we find no effect of treatments T1 and T3 on $\hat{S}^r_{it}$ without priming, but a large increase in sophistication with priming. We repeated the exercise with beliefs about the opponent's choice, and get a similar pattern of results (see Appendix Table A2). We however acknowledge that, in both cases, the four-way split of the data results in a number of observations that is too small for these differences to be statistically significant.

We also estimate the effect of the different treatments on participants' best response to their guess about their opponent's choice y_{jt}. For any guess about y_{jt}, i's payoff is highest when $y_{it} \in [a_{it}, b_{it}]$ is closest to $\gamma_{it} \times y_{jt}$. We calculate this value y^*_{it} for each observation and construct a dummy variable equal to 1 if the participant's choice y_{it} is within step size κ from y^*_{it}, and 0 otherwise. We then regress this variable on treatment dummies as in Table 3. This allows us to test whether the increases in strategic sophistication documented in 3 result entirely from ascribing a higher strategic sophistication to the opponent. If this is the sole causal mechanism, conditioning for i's belief about the opponent's choice of y_{jt} should absorb the reduced-form effect of the treatments on $\hat{S}^r_{it}$.

From the results presented in Table 6, this is what we find for participants whose choice of y_{it} was made *after* guessing their opponent's y_{jt}: for those participants, column (3) of Table 3 documented a large and significant effect of T1 and T3 on sophistication index $\hat{S}^r_{it}$. This effect entirely disappears in column 3 of Table 6, suggesting that, for those respondents, the effect of these treatments on strategic sophistication is entirely mediated by their effect on the implied strategic sophistication $\hat{S}^r_{jt}$ of their opponent. show no effect of T1 or T3 on the probability of choosing a best response to y_{jt}'s guess.

In contrast, a significant treatment effect of T1 and T3 remains for participants

Table 6: Treatment effects on best response to opponent's expected choice (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.039**	0.069**	0.010	-0.0502
	(0.015)	(0.022)	(0.021)	
T2 - Video	0.039	0.048	0.031	-0.0069
	(0.021)	(0.034)	(0.025)	
T3 - Factual + Video	0.013	0.045*	-0.019	-0.0545
	(0.017)	(0.023)	(0.025)	
Round	-0.004***	-0.005***	-0.003	
	(0.001)	(0.001)	(0.001)	
Constant	0.470***	0.379***	0.560***	
	(0.014)	(0.019)	(0.021)	
T1=T3 (p)	0.104	0.297	0.209	
T2=T3 (p)	0.281	0.946	0.097	
T1=T2 (p)	0.986	0.511	0.340	
N	4233	2111	2122	

NOTES. This table shows treatment effects on a dummy variable equal to 1 if the participant's own choice is within step size κ from the best response to their guess about the opponent's choice, and 0 otherwise. Column (2) and (3) split the sample by whether beliefs about opponents' guesses are elicited before or after subjects make their own choice. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

who choose their own y_{it} before guessing their opponent's y_{jt}. This suggests that, when deciding y_{it} these participants do not first form a belief about y_{jt} and then pick a best response to that. To do that, they need to be prompted to guess y_{jt} first, even in within-subject design such as ours. While our experiment does not allow us to identify precisely why this is the case, one possible explanation is that forming a belief about the opponent *before* choosing y_{it} is cognitively costly. Hence participants economize on this task in rounds when they are not forced to go through the mental process. This may also explain why, in Table 3, the effects of T1 and T3 on strategic sophistication are smaller when participants can eschew the mental process before making a choice for themselves: economizing on mental effort results in a smaller improvement in strategic sophistication.

5 Robustness analysis

To test the robustness of our results, we replicate the analysis presented in Tables 2 to 5 and Appendix Table A2 with two alternative indices of strategic sophistication. These estimates are a direct extension of the way data from level-k experiments have been used in the past (e.g., Costa-Gomes and Crawford, 2006) and Fragiadakis et al. (2016), i.e., interpreting a choice close to an L^k value as indicating a level-k sophistication, in spite of the fact that participants who play at random (are L^0) can also choose an L^k value by chance. For comparison purposes, we report in Appendix Table B1 a categorization of choices by their naively implied level of strategic reasoning. Model predictions other than level-k (i.e., iterated dominance and Nash) represent a fraction of all observed choices that is not noticeably different from what would occur by chance by participants playing L_0, which is why they are omitted thereafter.

The first alternative index of strategic sophistication, denoted A^r_{it} attributes to a high level of sophistication any guess within distance κ from it. A^r_{it} takes integer values 0 to 3. A guess within distance κ from L^3 yields $A^r_{it} = 3$; a guess close to L^2 but not L^3 takes value 2; and a guess close to L^1 but not L^2 or L^3 takes value 1. Other guesses take value 0.

As is immediately apparent from subsection 2.2, this index overestimates strategic sophistication since it ignores the possibility that a random choice may within κ of L^k by chance. It also overestimates strategic sophistication because, it attributes fully to $L^{k'}$ any choice for which $L^k = L^{k'}$ with $k < k'$. To counteract this second source of overestimation, we introduce a second alternative measure of strategic sophistication B^r_{it} that attributes such cases to L^k instead. More precisely, a guess close to L^1 yields a value 1 for B^r_{it}; a guess close to L^2 but not L^1 takes value 2. A guess close to L^3 but not L^1 or L^2 takes value 3. Other guesses take value 0.

Appendix Tables B2 and B3 present regression estimates for A^r_{it} and B^r_{it} in Ghana. These can be compared to those presented in Table 2. We see that coefficient point estimates for the Team member treatment are virtually identical to those reported in magnitude for $\hat{S}^r_{it}$, but they are not statistically significant. Coefficient estimates for the Team captain treatment are smaller in magnitude – except in Column (4) – and they are generally not statistically significant – except in Column (4) for B^r_{it}. Given the magnitude of the standard errors, however, we could not reject that the estimates are the same. From this we conclude that, as could be expected, our proposed methodology implemented in Table 2 does provide an improvement in precision.

A similar exercise for Table 3 is shown in Appendix Tables B4 and B5. Results for alternative indices A^r_{it} and B^r_{it} massively confirm what we found for $\hat{S}^r_{it}$ in Tanzania. A similar conclusion if we compare Table 4 with Appendix Tables B6 and B7: if anything, the significant results shown with $\hat{S}^r_{jt}$ are amplified when using alternative indices A^r_{jt} and B^r_{jt}: even the video treatment T2 on its own is shown to have a significant effect in some of the regressions. The same general conclusion holds for Table 5, which are repeated for A^r_{it} and B^r_{it} in Appendix Tables B8 and B9.[19] From this we conclude that our findings are not an artifact of our novel methodology: they are robust to more standard (but biased) approaches.

6 Conclusions

This paper studies how strategic sophistication responds to two distinct mechanisms: social interaction on the decision-maker's side and information about the opponent. Using two complementary experiments based on level-k guessing games, we examine whether teamwork and opponent-specific signals can foster deeper strategic reasoning.

First, we run an experiment with entrepreneurs in Ghana, which focuses on teamwork and responsibility. We find that team captains display substantially greater strategic sophistication than individuals making decisions alone, while non-captain team members show much smaller improvements. This pattern suggests that the main effect of teamwork operates through responsibility for the team's outcome rather than through

[19]We also obtain similar results if we split the sample between sessions 1-7 and 8-10, or if we expand the sophistication index to include other strategies discussed in Costa-Gomes and Crawford (2006) and Fragiadakis et al. (2016), namely, iterated dominance and Nash equilibrium. Results are available on request.

broad-based learning or information sharing within the team. Being accountable for a collective decision appears to induce greater cognitive effort and deeper reasoning, highlighting the importance of incentives and responsibility in shaping strategic behavior.

Second, we run an experiment with university students in Tanzania. We find that providing factual information about the opponent, such as gender, field of study, and academic performance, leads participants to attribute higher levels of strategic reasoning to their opponent and to increase their own depth of reasoning accordingly. This is consistent with the idea that cognitively relevant information helps players form more structured beliefs about others' behavior, thereby improving best responses. By contrast, exposure to a short introductory video of the opponent, designed to convey personal and relatable information without signaling strategic ability, has no measurable effect when provided in isolation. This suggests that affective or emotive cues alone are insufficient to enhance strategic reasoning in this context. However, when combined with factual information, the video significantly strengthens its impact. This pattern indicates that affective exposure can complement cognitively relevant signals, potentially by increasing attention to the opponent or facilitating the processing of information about their strategic traits. While our design allows us to document this complementarity, it does not fully disentangle the underlying mechanisms. One possibility is that affective cues increase engagement and effort in processing factual information. Another is that they make it easier for participants to adopt the perspective of a more strategically sophisticated opponent. Distinguishing between these channels remains an important direction for future research.

Taken together, our results highlight that strategic sophistication can be shaped both by how decisions are made and by how opponents are perceived. Improvements in reasoning do not arise from generic social or emotional cues, but rather from interventions that either increase accountability or provide structured, cognitively relevant information. These findings contribute to the literature on strategic thinking by moving beyond the documentation of heterogeneity and identifying conditions under which deeper reasoning can be fostered.

References

Agranov, M., E. Potamites, A. Schotter, and C. Tergiman (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior 75*(2), 449–463.

Andries, M., L. Bursztyn, T. Chaney, and M. Djourelova (2024). In their shoes. Technical report, CEPR Discussion Paper No. 19197.

Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In *J. Decety and W. Ickes (Eds.), The social neuroscience of empathy*, pp. 3–15. Boston Review.

Blanchette, I., S. Caparos, and B. Trémolière (2017). Emotion and reasoning. In *International handbook of thinking and reasoning*, pp. 57–70. Routledge.

Blanchette, I. and A. Richards (2004). Reasoning about emotional and neutral materials: Is logic affected by emotion? *Psychological science 15*(11), 745–752.

Blanchette, I. and A. Richards (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion 24*(4), 561–595.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics 119*(3), 861–898. Cited by: 990.

Casanelles, T. E. and D. Gonçalves (2026). Incentives and strategic behaviour: An experiment. Technical report, CEPR DP21338.

Costa-Gomes, M. A. and V. P. Crawford (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review 96*(5), 1737–1768.

Costa-Gomes, M. A. and G. Weizsäcker (2008). Stated beliefs and play in normal-form games. *Review of Economic Studies 75*(3), 729–762.

Crawford, V. P., M. A. Costa-Gomes, and N. Iriberri (2013). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature 51*(1), 5–62.

Damasio, A. R. (2006). *Descartes' Error*. Random House.

Davis, M. H. (1996). Non-affective outcomes. In *Empathy: A social psychological approach*. Routledge.

Dean, J. T., M. R. Dillon, E. Duflo, H. Kannan, and E. Spelke (2025). Cognitive science at scale: Classroom games played with peers, pairing symbolic and intuitive training, durably enhance young children's learning of mathematics. Technical report, PsyArXiv.

Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science 298*(5596), 1191–1194.

Enke, B., R. Rodriguez-Padilla, and F. Zimmermann (2022). Moral universalism and the structure of ideology. *Review of Economic Studies 90*, 1934–1962.

Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica 73*(5), 1623–1672.

Fafchamps, M., D. McKenzie, S. Quinn, and C. Woodruff (2014). Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in ghana. *Journal of Development Economics 106*, 211–226.

Fershtman, C. and U. Gneezy (2001). Discrimination in a segmented society: An experimental approach. *Quarterly Journal of Economics 116*(1), 351–377.

Fragiadakis, D. E., D. T. Knoepfle, and M. Niederle (2016). Who is strategic? Technical report, Stanford University.

Gans, J. S. (2026). Belief dispersion and entrepreneurial entry. Working Paper 35091, National Bureau of Economic Research.

Georganas, S., P. J. Healy, and R. A. Weber (2015). On the persistence of strategic sophistication. *Journal of Economic Theory 159*, 369–400.

Gill, D. and Y. Rosokha (2024). Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma. *American Economic Journal: Microeconomics 16*.

Grosskopf, B. and R. Nagel (2008). The two-person beauty contest. *Games and Economic Behavior 62*, 93–99.

Hoff, K. and P. Pandey (2006). Discrimination, social identity, and durable inequalities. *American Economic Review Papers & Proceedings 96*(2), 206–211.

Jung, N., C. Wranke, K. Hamburger, and M. Knauff (2014). How emotions affect logical reasoning: evidence from experiments with mood-manipulated participants, spider phobics, and people with exam anxiety. *Frontiers in psychology 5*, 570.

Kalla, J. and D. Broockman (2023). Which narrative strategies durably reduce prejudice? evidence from field and survey experiments supporting the efficacy of perspective-getting. *American Journal of Political Science 67*, 185–204.

Kocher, M. G., S. Strauss, and M. Sutter (2006). Individual or team decision-making: Causes and consequences of self-selection. *Games and Economic Behavior 56*(2), 259–270.

Kocher, M. G. and M. Sutter (2005). The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *Economic Journal 115*(500), 200–223.

Krebs, D. (1975). Empathy and altruism. *Journal of Personality and Social Psychology 32*, 1134–1146.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes 65*(3), 272–292.

Loewenstein, G. and J. Lerner (2003). The role of emotion in decision making. In R. Davidson, H. Goldsmith, and K. Scherer (Eds.), *Handbook of affective science*, pp. 619–642. Oxford University Press.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review 85*(5), 1313–1326.

Paluck, E. L., S. A. Green, and D. P. Green (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy 3*, 129–158.

Pettigrew, T. F. and L. R. Tropp (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology 90*, 751–783.

Rick, S. and G. Loewenstein (2008). The role of emotion in economic behavior. In M. Lewis, J. M. Haviland-Jones, and L. F. Barrett (Eds.), *Handbook of Emotions, Third Edition*, pp. 138–156. The Guilford Press.

Sbriglia, P. (2008). Revealing the depth of reasoning in p-beauty contest games. *Experimental Economics 11*, 107–121.

Scacco, A. and S. S. Warren (2018). Can social contact reduce prejudice and discrimination? evidence from a field experiment in nigeria. *American Political Science Review 112*(3), 654–677.

Schwarz, N. (2000). Emotion, cognition, and decision making. *Cognition & emotion 14*(4), 433–440.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics 87*(3), 355–374.

Stahl, D. O. and P. W. Wilson (1995). On players models of other players: Theory and experimental evidence. *Games and Economic Behavior 10*(1), 218–254.

Stahl II, D. O. and P. W. Wilson (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization 25*(3), 309–327.

Train, K. (2002). *Discrete Choice Methods with Simulation*. Cambridge University Press.

Online Appendix

A Additional Tables and Figures

Table A1: Treatment effects on $Pr(k_{it} = K)$ (Ghana)

	T0	T1	T2
Level-0			
Coefficient	0.843***	0.795***	0.794***
Std. error	0.035	0.036	0.049
Level-1			
Coefficient	0.120***	0.124***	0.137***
Std. error	0.033	0.031	0.043
Level-2			
Coefficient	0.038***	0.052***	0.081***
Std. error	0.012	0.010	0.016
Level-3			
Coefficient	-0.001	0.030**	-0.012
Std. error	0.021	0.014	0.021

NOTES. The chosen interval step size κ is 20. Standard errors are clustered at the individual player level. *p < 0.10; **p < 0.05; ***p < 0.01.

Table A2: Treatment effects on the implied index $\hat{S}^r_{jt}$ of the opponent by level of understanding of the game (Tanzania)

	(1)	(2)	(3)	(4)
	High underst.		Low underst.	
What comes first:	Choice	Guess	Choice	Guess
T1 - Factual Info	0.050	0.116***	-0.102	0.102
	(0.034)	(0.034)	(0.085)	(0.087)
T2 - Video	-0.003	0.061	-0.265	0.082
	(0.063)	(0.064)	(0.133)	(0.155)
T3 - Factual + Video	0.104*	0.137**	-0.079	0.071
	(0.047)	(0.042)	(0.097)	(0.083)
Round	-0.005*	-0.003	0.009	-0.001
	(0.002)	(0.002)	(0.006)	(0.005)
Constant	1.425***	1.339***	1.393***	1.387***
	(0.033)	(0.025)	(0.059)	(0.051)
T1=T3 (p)	0.210	0.639	0.753	0.763
T2=T3 (p)	0.137	0.329	0.244	0.944
T1=T2 (p)	0.370	0.448	0.299	0.891
N	1684	1685	432	446

Notes. The dependent variable is the implied index of strategic thinking $\hat{S}^r_{jt}$ of the opponent. It takes value 1 for random play (level-0), value 2 for level-1 thinking, etc. The unit of analysis is the individual guess. High understanding means 5 or more correct responses on the quiz, low understanding means 4 or fewer correct responses. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

B Robustness analysis

Table B1: Categorization of Guesses by Naively Implied Level of Strategic Reasoning

Level of strategic reasoning	N	Pct.	Cum. pct.
L1 strategy	1520	34.5	34.5
L2 strategy	624	14.2	48.7
L3 strategy	97	2.2	50.9
D1 first iterated dominance	82	1.9	52.8
D2 second iterated dominance	0	0	52.8
D3 third iterated dominance	40	.9	53.7
Nash equilibrium strategy	76	1.7	55.4
Unknown (includes L0)	1512	34.4	89.8
Confused	449	10.2	100
Total	4400		

Notes. This table shows the level of strategic reasoning (k-level) naively implied by each guess. The experiment had 200 participants. Each participant made 22 guesses. This produced 4400 guesses in total. When the guess is too far from a known level of strategic reasoning, it is categorized as ”Unknown” or ”Confused”.

Table B2: Treatment effects on alternative index $\hat{A}^r_{it}$ (Ghana)

	1	2	3	4
Team member				
Coefficient	0.0208	0.0284	0.0301	0.0284
Std. error	0.0477	0.0469	0.0469	0.0469
Team captain				
Coefficient	0.0591	0.0666	0.0663	0.0962
Std. error	0.0559	0.0552	0.0552	0.0660
N.missing guesses when team captain				
Coefficient				-0.0480
Std. error				0.0526
Intercept				
Coefficient	0.5049***	0.4968***	0.4965***	0.4968***
Std. error	0.0407	0.0397	0.0396	0.0397
N	2991	2991	2991	2991

Note: The dependent variable is the alternative index of strategic thinking $\hat{A}^r_{it}$. Guesses outside the allowed range are excluded from the regression. Column 1 includes no weighting. Column 2 includes Inverse Probability Weights (IPW) for missing guesses. Column 3 includes an IPW correction with more predictors. Column 4 adds a regressor for the number of missing guesses by team members when captain. Standard errors are clustered at the individual player level. *p $<$ 0.10; **p $<$ 0.05; ***p $<$ 0.01.

Table B3: Treatment effects on alternative index $\hat{B}^r_{it}$ (Ghana)

	1	2	3	4
Team member				
Coefficient	0.0202	0.0263	0.0278	0.0263
Std. error	0.0421	0.0416	0.0415	0.0416
Team captain				
Coefficient	0.0736	0.0762	0.0751	0.1087*
Std. error	0.0506	0.0501	0.0500	0.0584
N.missing guesses when team captain				
Coefficient				-0.0528
Std. error				0.0425
Intercept				
Coefficient	0.4515***	0.4454***	0.4451***	0.4454***
Std. error	0.0357	0.0348	0.0348	0.0348
N	2991	2991	2991	2991

Note: The dependent variable is the alternative index of strategic thinking $\hat{B}^r_{it}$. Guesses outside the allowed range are excluded from the regression. Column 1 includes no weighting. Column 2 includes Inverse Probability Weights (IPW) for missing guesses. Column 3 includes an IPW correction with more predictors. Column 4 adds a regressor for the number of missing guesses by team members when captain. Standard errors are clustered at the individual player level. *p $<$ 0.10; **p $<$ 0.05; ***p $<$ 0.01.

Table B4: Treatment effects on alternative index A^r_{it} (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.172***	0.132	0.212***	0.0830
	(0.045)	(0.070)	(0.055)	
T2 - Video	0.080	0.060	0.100	0.0424
	(0.069)	(0.094)	(0.100)	
T3 - Factual + Video	0.203***	0.134	0.272***	0.1408
	(0.051)	(0.076)	(0.067)	
Round	-0.011***	-0.011**	-0.010*	
	(0.003)	(0.004)	(0.004)	
Constant	0.854***	0.841***	0.867***	
	(0.036)	(0.049)	(0.053)	
T1=T3 (p)	0.517	0.971	0.369	
T2=T3 (p)	0.139	0.532	0.146	
T1=T2 (p)	0.238	0.525	0.304	
N	4233	2111	2122	

Notes. This table shows treatment effects on alternative measure of strategic sophistication A^r_{it} that attributes to a high level of sophistication any guess within step size κ from it. The unit of analysis is the individual guess. The dependent variable takes integer values 0 to 3. A guess close to L3 takes value 3. A guess close to L2 but not L3 takes value 2. A guess close to L1 but not L2 or L3 takes value 1. Other guesses take value 0. This measure overestimates strategic sophistication, as explained in the text. Guesses outside the allowed range are omitted from the analysis. Column (2) and (3) split the sample by whether beliefs about opponents' guesses are elicited before or after subjects make their own choice. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

Table B5: Treatment effects on alternative index B^r_{it} (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.184***	0.156***	0.212***	0.0669
	(0.025)	(0.040)	(0.032)	
T2 - Video	0.145**	0.099	0.191**	0.1042
	(0.044)	(0.060)	(0.065)	
T3 - Factual + Video	0.194***	0.161***	0.227***	0.0772
	(0.030)	(0.047)	(0.036)	
Round	-0.005***	-0.007**	-0.004	
	(0.002)	(0.002)	(0.002)	
Constant	0.440***	0.465***	0.415***	
	(0.021)	(0.029)	(0.029)	
T1=T3 (p)	0.739	0.913	0.712	
T2=T3 (p)	0.356	0.412	0.631	
T1=T2 (p)	0.431	0.422	0.765	
N	4233	2111	2122	

NOTES. This table shows treatment effects on alternative measure of strategic sophistication B^r_{it} that attributes to a high level of sophistication any guess within step size κ from it. The unit of analysis is the individual guess. The dependent variable takes integer values 0 to 3. A guess close to L1 takes value 1. A guess close to L2 but not L1 takes value 2. A guess close to L3 but not L1 or L2 takes value 3. Other guesses take value 0. This measure overestimates strategic sophistication, but by less than A^r_{it}. Guesses outside the allowed range are omitted from the analysis. Column (2) and (3) split the sample by whether beliefs about opponents' guesses are elicited before or after subjects make their own choice. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

Table B6: Treatment effects on the alternative index $\hat{A}^r_{jt}$ of the opponent (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.208***	0.167***	0.248***	0.2493
	(0.036)	(0.051)	(0.052)	
T2 - Video	0.174*	0.059	0.286**	0.0913
	(0.069)	(0.098)	(0.094)	
T3 - Factual + Video	0.206***	0.181**	0.231***	0.5829
	(0.048)	(0.070)	(0.065)	
Round	-0.009***	-0.009*	-0.008**	
	(0.002)	(0.004)	(0.003)	
Constant	0.681***	0.756***	0.606***	
	(0.032)	(0.050)	(0.041)	
T1=T3 (p)	0.968	0.832	0.807	
T2=T3 (p)	0.700	0.294	0.637	
T1=T2 (p)	0.652	0.295	0.725	
N	4247	2116	2131	

Notes. The dependent variable is the implied strategic sophistication $\hat{A}^r_{jt}$ of opponent. The unit of analysis is the individual guess. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$.

Table B7: Treatment effects on the alternative index $\hat{B}^r_{jt}$ of the opponent (Tanzania)

	(1) All	(2) Choice before guess	(3) Choice after guess	(4) col.2 vs col.3
T1 - Factual Info	0.213***	0.167***	0.259***	0.0767
	(0.025)	(0.035)	(0.035)	
T2 - Video	0.163***	0.130*	0.195**	0.5076
	(0.046)	(0.064)	(0.066)	
T3 - Factual + Video	0.224***	0.193***	0.256***	0.3428
	(0.031)	(0.044)	(0.044)	
Round	0.001	0.002	0.000	
	(0.002)	(0.002)	(0.002)	
Constant	0.328***	0.361***	0.295***	
	(0.020)	(0.030)	(0.026)	
T1=T3 (p)	0.736	0.577	0.954	
T2=T3 (p)	0.271	0.410	0.463	
T1=T2 (p)	0.327	0.587	0.411	
N	4247	2116	2131	

NOTES. The dependent variable is the implied strategic sophistication $\hat{B}^r_{jt}$ of opponent. The unit of analysis is the individual guess. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

Table B8: Treatment effects on alternative index $\hat{A}^r_{jt}$, by level of understanding of the game (Tanzania)

	(1)	(2)	(3)	(4)
	High underst.		Low underst.	
What comes first:	Choice	Guess	Choice	Guess
T1 - Factual Info	0.171*	0.210***	-0.015	0.218
	(0.077)	(0.061)	(0.166)	(0.131)
T2 - Video	0.101	0.123	-0.118	-0.003
	(0.099)	(0.113)	(0.278)	(0.206)
T3 - Factual + Video	0.215*	0.288***	-0.155	0.216
	(0.086)	(0.078)	(0.148)	(0.123)
Round	-0.012**	-0.010*	-0.008	-0.013
	(0.004)	(0.005)	(0.008)	(0.009)
Constant	0.819***	0.846***	0.920***	0.949***
	(0.056)	(0.060)	(0.103)	(0.113)
T1=T3 (p)	0.589	0.312	0.304	0.984
T2=T3 (p)	0.372	0.227	0.906	0.366
T1=T2 (p)	0.554	0.484	0.754	0.348
N	1675	1674	436	448

NOTES. This table shows treatment effects on alternative measure of strategic sophistication $\hat{A}^r_{jt}$ that attributes to a high level of sophistication any guess within step size κ from it. The unit of analysis is the individual guess. The dependent variable takes integer values 0 to 3. A guess close to L3 takes value 3. A guess close to L2 but not L3 takes value 2. A guess close to L1 but not L2 or L3 takes value 1. Other guesses take value 0. This measure overestimates strategic sophistication, as explained in the text. High understanding means 5 or more correct responses on the quiz, low understanding means 4 or fewer correct responses. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

Table B9: Treatment effects on alternative index $\hat{B}^r_{jt}$, by level of understanding of the game (Tanzania)

	(1)	(2)	(3)	(4)
	High underst.		Low underst.	
What comes first:	Choice	Guess	Choice	Guess
T1 - Factual Info	0.179***	0.204***	0.068	0.241**
	(0.045)	(0.036)	(0.087)	(0.067)
T2 - Video	0.110	0.207**	0.058	0.124
	(0.064)	(0.071)	(0.164)	(0.157)
T3 - Factual + Video	0.205***	0.233***	0.002	0.204**
	(0.053)	(0.043)	(0.096)	(0.069)
Round	-0.007**	-0.004	-0.007	-0.005
	(0.002)	(0.003)	(0.005)	(0.005)
Constant	0.440***	0.403***	0.559***	0.461***
	(0.032)	(0.033)	(0.067)	(0.063)
T1=T3 (p)	0.622	0.545	0.408	0.579
T2=T3 (p)	0.257	0.747	0.769	0.643
T1=T2 (p)	0.374	0.970	0.957	0.489
N	1675	1674	436	448

NOTES. This table shows treatment effects on an alternative measure of strategic sophistication $\hat{B}^r_{jt}$ that attributes to a high level of sophistication any guess within step size κ from it. The unit of analysis is the individual guess. The dependent variable takes integer values 0 to 3. A guess close to L1 takes value 1. A guess close to L2 but not L1 takes value 2. A guess close to L3 but not L1 or L2 takes value 3. Other guesses take value 0. This measure overestimates strategic sophistication, as explained in the text, but by less than measure 1. High understanding means 5 or more correct responses on the quiz, low understanding means 4 or fewer correct responses. The analysis omits choices outside the range allowed by game parameters. Participant fixed effects are included throughout. Standard errors are clustered at the participant level. *p < 0.10; **p < 0.05; ***p < 0.01.

C Relation between own choice and expected choice of opponent

Table C1: Regressing own level of strategic sophistication on opponent's expected strategic sophistication

	OLS	Participant FE
	(1)	(2)
Expected opponent's level of strategic sophistication	.536	.502
	(.030)***	(.030)***
Const.	.833	.891
	(.059)***	(.058)***
Player FE	No	Yes
N	1404	1404

NOTES. This table shows the results of an OLS regression of own naive expanded index of strategic sophistication (a categorical variable taking values from 1 to 7) on the expected level of strategic sophistication of the opponent (a categorical variable taking the same values 1 to 7). The unit of analysis is an individual choice (each subject makes 22 choices, one for each of their 22 tasks). The regression excludes observations in which the subject's own choice or the expected choice of the opponent is classified as "unknown" or "confused". *p < 0.10; **p < 0.05; ***p < 0.01.

Table C1 provides a detailed examination of the correlation between players' level of strategic reasoning and the expectation of their opponent's level. We estimate an Ordinary Least Squares (OLS) regression model in which each observation is a subject's choice in one task and the dependent variable is the level of strategic reasoning implied by that choice. For the purpose of this analysis, we expand our naive sophistication index B_{it}^{k} to include the seven different strategies include in Table B1, with 1 denoting L^1 and 7 representing Nash equilibrium. The independent variable is constructed from the subject's reported belief about their opponent and is the expected level of strategic reasoning of the opponent that is implied by that belief; it is also categorized from 1 to 7. Beliefs about an opponent's choice are handled in the same way as subjects' choices, that is, a belief is classified as falling in one of the seven categories if it is within 10 units of the corresponding benchmark. Beliefs that are out of bounds or cannot be rationalized by one of our seven levels of strategic reasoning are excluded from the analysis. If a choice is within 10 units of two or more benchmarks, it is assigned to the strategy whose benchmark is closest; if several benchmarks are closest, they are assigned to the lowest strategy, as in B^r_{it}.

Level-k thinking and iterated dominance would imply an intercept equal to one: if the level of strategic reasoning ascribed to the opponent is k, the subject's best response is to choose the level-$k+1$ strategy. Furthermore, if the level of strategic reasoning ascribed to the opponent increases by one unit, the best response of a level-k subject should increase by one unit as well.

These predictions are put to the test in Table C1, with two specifications: (1) OLS and (2) Participant Fixed Effects (FE). We first note that, in both regressions, the intercept is close to 1 – and not statistically different from 1 at the 5% level in the Participant FE regression. Secondly, the opponent's expected level of strategic reasoning is large and statistically significant, but it is less than 1 in magnitude and statistically different from 1. These findings confirm that participants select a higher level of strategic thinking for themselves than they implicitly assign to their opponent.

D Task Parameters

Table B1: Original Costa-Gomez & Crawford (2006) parameters and implied strategies

	Game parameters						Strategies for i			Strategies for j		
Game	L_i	U_i	μ_i	L_j	U_j	μ_j	$L1$	$L2$	$L3$	$L1$	$L2$	$L3$
1	300	900	0.7	100	900	1.3	350	546	318.5	780	455	709.8
2	100	900	1.3	300	900	0.7	780	455	709.8	350	546	318.5
3	100	900	0.5	100	500	1.5	150	250	112.5	500	225	375
4	100	500	1.5	100	900	0.5	500	225	375	150	250	112.5
5	100	500	0.7	100	500	1.5	210	315	220.5	450	315	472.5
6	100	500	1.5	100	500	0.7	450	315	472.5	210	315	220.5
7	300	500	0.7	100	900	1.5	350	420	367.5	600	525	630
8	100	900	1.5	300	500	0.7	600	525	630	350	420	367.5
9	100	500	0.7	100	900	0.5	350	105	122.5	150	175	100
10	100	900	0.5	100	500	0.7	150	175	100	350	105	122.5
11	300	900	1.3	300	900	1.3	780	900	900	780	900	900
12	300	900	1.3	300	900	1.3	780	900	900	780	900	900
13	300	500	1.5	300	900	1.3	500	500	500	520	650	650
14	300	900	1.3	300	500	1.5	520	650	650	500	500	500
15	100	900	0.5	300	500	0.7	200	175	150	350	300	300
16	300	500	0.7	100	900	0.5	350	300	300	200	175	150

Table B2: Ghana parameters and implied strategies

	Game parameters						Strategies for i			Strategies for j		
Game	L_i	U_i	μ_i	L_j	U_j	μ_j	$L1$	$L2$	$L3$	$L1$	$L2$	$L3$
1	200	800	0.7	100	900	1.2	350	420	294	600	420	504
2	100	500	1.1	100	900	0.6	500	198	330	180	300	118.8
3	300	900	0.8	100	900	1.3	400	624	416	780	520	811.2
4	100	1000	1.3	300	800	0.7	715	500.5	650.65	385	500.5	350.35
5	100	900	0.5	100	600	1.5	175	300	131.25	600	262.5	450
6	100	500	1.3	100	900	0.7	500	273	455	210	350	191.1
7	100	500	0.5	100	500	1.5	150	225	112.5	450	225	337.5
8	200	600	1.4	200	700	0.7	600	392	588	280	420	274.4
9	100	700	0.7	100	900	1.2	350	336	294	480	420	403.2
10	200	900	1.4	300	900	0.7	840	539	823.2	385	588	377.3
11	100	900	0.5	100	500	1.5	150	250	112.5	500	225	375
12	100	500	1.4	100	900	0.6	500	252	420	180	300	151.2
13	200	600	0.7	100	600	1.3	245	364	222.95	520	318.5	473.2
14	100	500	1.5	100	500	0.7	450	315	472.5	210	315	220.5
15	200	800	1.5	200	700	0.7	675	525	708.75	350	472.5	367.5
16	200	600	0.6	100	700	1.5	240	360	216	600	360	540

Table B3: Tanzania parameters and implied strategies

	Game parameters						Strategies for i			Strategies for j		
Game	L_i	U_i	μ_i	L_j	U_j	μ_j	$L1$	$L2$	$L3$	$L1$	$L2$	$L3$
1	0	100	1	0	100	1	50	50	50	50	50	50
2	0	100	1	0	100	2	50	100	100	100	100	100
3	0	100	2	0	100	1	100	100	100	50	100	100
4	0	100	1	0	100	0.5	50	25	25	25	25	13
5	0	100	1	0	100	1	50	50	50	50	50	50
6	0	100	1	0	100	2	50	100	100	100	100	100
7	0	100	2	0	100	1	100	100	100	50	100	100
8	0	100	1	0	100	0.5	50	25	25	25	25	13
9	0	100	0.5	0	100	1	25	25	13	50	25	25
10	0	100	1	0	200	1	100	50	100	50	100	50
11	40	100	1	0	100	1	50	70	50	70	50	70
12	40	200	0.5	0	100	1	40	50	40	100	40	50
13	0	100	1	40	200	1	100	50	100	50	100	50
14	0	100	1	40	200	0.5	100	40	50	40	50	40
15	0	100	1	40	100	1	70	50	70	50	70	50
16	0	200	1	0	100	1	50	100	50	100	50	100
17	0	200	0.5	0	100	1	25	50	13	100	25	50
18	40	200	1	0	100	1	50	100	50	100	50	100
19	0	100	1	0	200	0.5	100	25	50	25	50	13
20	0	100	0.5	100	400	2	100	50	100	100	200	100
21	0	200	2	0	200	1	200	200	200	100	200	200
22	0	200	1	0	200	0.5	100	50	50	50	50	25

E Details of the Ghana Experiment

E.1 Game instructions

Experimental Study of Strategic Interaction using Guessing Games

EXPERIMENT INSTRUCTIONS

Welcome!

You are about to participate in an experiment to test how you make decisions.

Today you will receive the instructions to participate in the experiment, which will continue over the next 16 days. In addition to participation fee (5 GhCedi) and some mobile phone credit (10 GhCedi) that you will receive at the end of today's session, you may earn a considerable additional amount of money, depending on the decisions you make. On average, participants in this game win about GhC 60. All the money that you earn is yours to keep, and will be paid to you in cash, at the end of the experiment, after 16 days of play.

Your earnings will be determined by your decisions and by the decisions of other participants. We now explain in detail how the game works and how your winnings are exactly determined. We kindly ask you to pay attention to the explanation and if you have any questions, please raise your hand. At the end of this training session you will be asked to sit an understanding test to evaluate if you have understood the rules of the game. It is important, therefore, that you pay attention and raise any questions you might have during this training session.

The experiment has 16 rounds.

You will be divided into groups of 4 and will play as a team.

In each round your team will be matched randomly with one other team (a new one in each round), who will be your opponent in that round. You will not know which of the other teams you are matched with. It could be a team from this training session, or a team from another training session.

In each round your team and the opponent team separately and independently make decisions called GUESSES. Together, your team's GUESS and your opponent's GUESS determine the number of POINTS that you earn in a round.

Earning more points increases your money payment at the end of the experiment, as explained below.

NOTE: Neither your team's guess nor your opponent's guess in a round affects how your team and other teams are matched or the games they face in the rest of the experiment.

To make a good guess, you must understand how your guess and your opponent's guess determines the number of points that you and they earn in every decision situation. Here is how.

In each decision situation, each team has a TARGET, a LOWER LIMIT and an UPPER LIMIT. These targets and limits are typically different for your team and your opponent team, and they change from round to round. In all other respects, the games are identical in all 16 rounds.

In each round each team is asked to guess a number between their lower limit and upper limit. Their opponent team is asked to also guess a number. If you make a mistake and guess outside your limit, your guess is automatically adjusted down to the upper limit (if it is above it), or up to the lower limit (if it is below it).

The aim of the game is to guess a number that is as close as possible to your opponent's guess *times* your target. The closer your guess is to your opponent's guess times your target, the higher the number of points you earn in each game.

EXAMPLE:

A: YOUR LOWER LIMIT = 100 and YOUR UPPER LIMIT = 600
B: YOUR TARGET = 1.5
C: YOUR OPPONENT'S LOWER LIMIT = 100 and YOUR OPPONENT'S UPPER LIMIT = 400
D: YOUR OPPONENT TARGET = 2

YOU:

LOWER LIMIT **UPPER LIMIT**

100 **600**

TARGET = 1.5

YOUR OPPONENT

LOWER LIMIT **UPPER LIMIT**

100 **400**

TARGET = 2

You receive the most points if you can guess a number that is equal to your opponent's guess times 1.5. In order to win the game, you must try to understand what your opponent is going to guess and multiply it by your target.

NOTE: Your guesses should not have decimal places.

LET'S TRY:

[The experimenter plays the mock game above between himself and another collaborator. They both make their guesses and at the end the payoffs are worked out as follows – referring to the graph below].

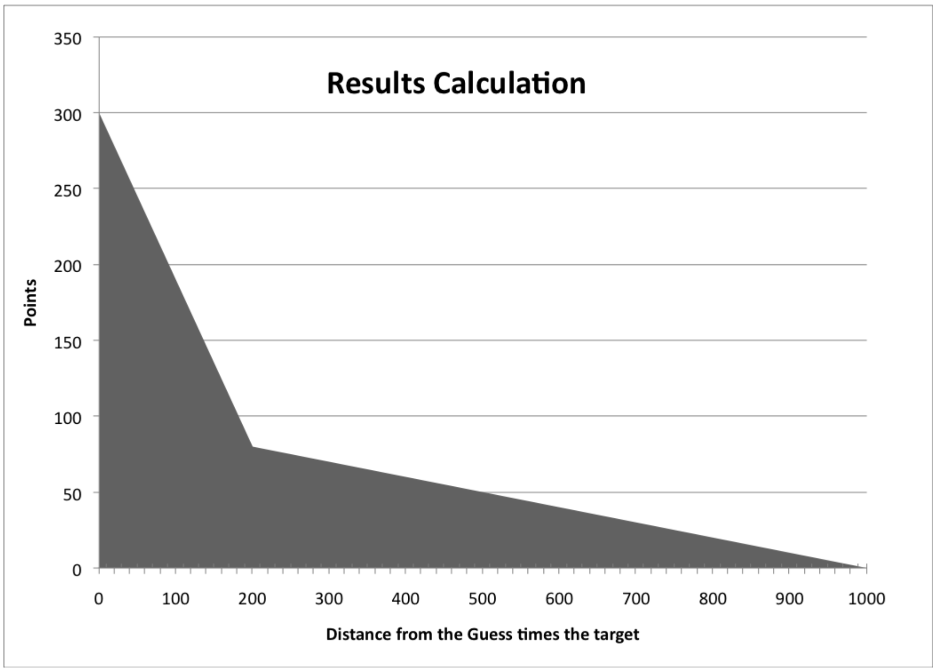
[Player A picks 150; Player B picks 450 and gets adjusted to 400]

*[Player A would get the most points if he had guessed the perfect guess of 400*1.5 = 600. Instead, he guessed 150 and is therefore 450 units away from the perfect guess]*

*[Player B would get the most points if he had guessed the perfect guess 150*2 = 300. Instead, he guessed 400 and is therefore 100 units away from the perfect guess]*

The following graph shows exactly how points are calculated, based on how far your guess is from the perfect guess, which is your opponent's guess times your target.

In the example above, Player A would have earned 55 points, since his guess is 450 units away from the perfect guess of 600 and Player B would have earned 190 points, since his guess is 100 units away from his perfect guess of 300.



You will play this type of game once a day for 16 days. At the end of the 16 days, we will randomly select 5 games out of the 16 you played and we will convert every 15 points you make on each game into 1 GhCedi.

So, for example, if you earn 90 points on every game, we will pay each member of the team an amount equal to (450 points / 15) = 30 GhCedi. The perfect guess in just ONE of the five selected games earns every member of the team 300 points, that is, 20 GhCedi.

GROUP PLAY:

You will play these games in teams.

To play the game your team needs four pieces of information:

A: Your lower and upper limit
B: Your target
C: Your opponent's lower and upper limit
D: Your opponent's target

Each day is a different game, and these four pieces of information are different for each game. Each of the 16 days over which the games are played, each team member will receive ONE of the four pieces of information that, together, are needed to make a guess. It is VERY important, therefore, that you leave your correct phone number at the end of this training.

TEAM CAPTAINS

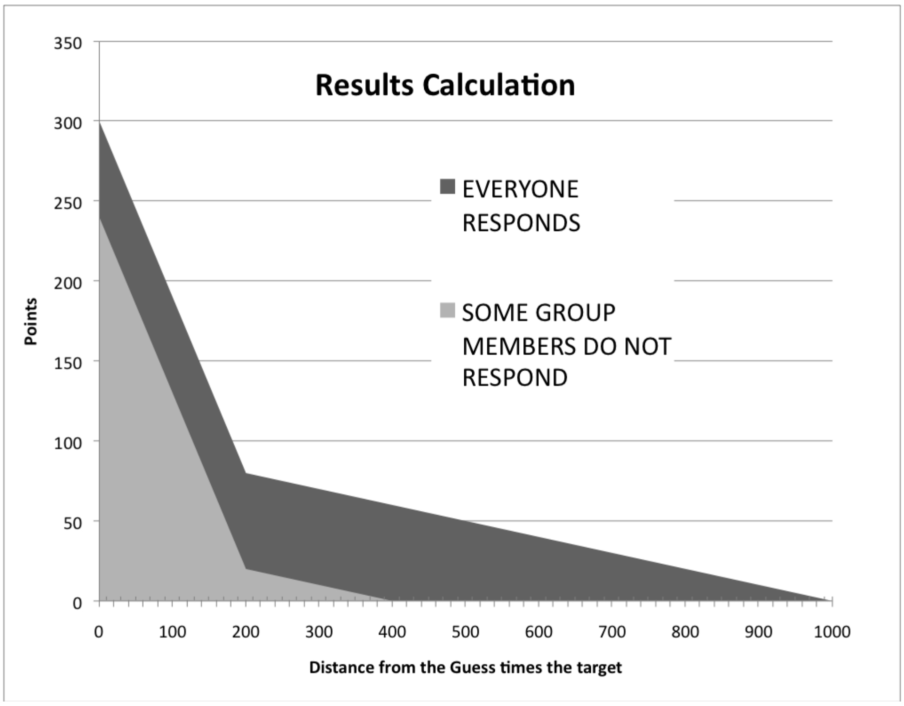
Each team will have a team captain for each game. The team captain is responsible for sending the answer of the team in an SMS to us.

We will rotate the team captains and each team member will be team captain 4 times out of 16 games. We will specify in our SMS who is team captain for the game played on that day.

You will receive the SMS with all the necessary info by 9 a.m. on the day of play and the team captain is required to send back the answer of the team by 7 p.m. at the latest. Late answers may cause the answer to be disqualified.

In addition to the captain's answer, each team member must also send their own preferred guess by SMS to us. This guess may differ from the captain's guess if they do not agree with it. Only the teams where the captain and every other team member sends their guesses will be entitled to the full payment of the winnings from the game.

If at least one of the team members does not send their guess, a penalty of 60 (equal to 4 Cedis) is deducted from the game score, as shown in the following graph.



This is the end of our training. Any questions?

E.2 Protocol for communication with participants

Communication with participants was conducted via a structured SMS exchange program that operated daily over the 16 days of the experiment. Each morning (by 9 a.m.), every team member received an individual SMS containing one piece of information required to play the game. Specifically, each participant was sent one of the following: (i) their teamâs lower and upper limits, (ii) their teamâs target, (iii) the opponent teamâs lower and upper limits, or (iv) the opponent teamâs target.

Participants were required to share the information they received with their teammates in order to reconstruct the full set of game parameters and jointly determine the teamâs guess. The team captain, identified in the same SMS, was responsible for submitting the teamâs final guess via SMS by 7 p.m. on the same day.

In addition, all team members were required to submit their own individual preferred guess via SMS (by 9 p.m.), regardless of whether it coincided with the teamâs final decision. This served both as a compliance check and as an individual-level measure of beliefs or preferences. Teams in which all members submitted their individual guesses were eligible for full payment; failure by any member to respond resulted in a penalty applied to the teamâs score.

This daily SMS exchange protocol ensured decentralized information, required active communication within teams, and enabled the researchers to monitor participation and responses in real time.

E.3 Protocol for team assignments

Participants were randomly assigned to teams of four within each experimental session (i.e., each day of play). Team composition was fixed for the duration of the experiment, so that the same four individuals interacted repeatedly over the 16 rounds.

Within each team, the role of team captain was randomly assigned in every round. As a result, over the course of the 16 days of play, each participant served as team captain exactly four times.

The identity of the team captain for each round was communicated to participants via the daily SMS message. The captain was responsible for submitting the teamâs final guess for that round, ensuring a clear allocation of responsibility while maintaining random rotation of decision authority across team members.

E.4 Participant engagement

Since participants to that study was selected from a random number of enumerations areas across a city of approximately 3 million inhabitants at the time, we expect respondents not to know each other. To confirm this, each participant assigned to the team treatment was asked whether he or she knew any of their of team members. Since there are 38 teams of 4, this represents 228 possible pairs. In only five of those (2.2%) is a pre-existing link reported, and in only one of them is the link reciprocal. This confirms that participants were, in their overwhelming majority, unknown to each other.

Non-participation to the experiment is rampant but mostly occasional. On average over all games, in the individual play treatment 31% have no reported guess in the individual treatment, but in the team treatment this figure falls to 25% for team members and 20% for captains. Only 18.4% of participants in the individual treatment report an individual guess for all 16 games. This number is 27.9% in team play. The median number of missed games is 2.5 for the individual treatment and 2 for the team treatment. 11.2% of individual participants miss all games vs. 5.9% of team participants.

One of the reason for missing guesses is the loss of phone signal which, as shown at the top of Table E.4, was relatively frequent during the experiment.[20] Participants did not, however, find the games themselves particularly difficult, particularly those in teams, and they perceived the games as getting easier over time. The vast majority stated that they used some reasoning when making a guess, with a higher reported use of reasoning by team members: some 60% of participants in teams report using reasoning all the time, compared to only 33% among individual players. Participants also report changing their

[20]Regression analysis, not shown here to save space, confirms that the probability of a missing guess increases with the frequency of problems participants reported with the phone network, but also with poor performance of the team and finding that the game is not getting easier. It also confirms the statistically significant reduction in missing guesses in the team treatment.

reasoning across games.

Table C1: Perception of games by participants

	Treatment assignment	
	Individual play	Team play
Frequency of phone network problems during the experiment		
Never	(42.6%)	(38.2%)
Less than 5	(48.9%)	(52.8%)
5 to 10	(7.4%)	(5.5%)
10 to 15	(1.1%)	(0.5%)
Every day	(0.0%)	(3.0%)
How difficult did you find the games?		
Very easy	(7.4%)	(10.2%)
Easy	(13.8%)	(21.8%)
Average	(64.9%)	(50.3%)
Difficult	(12.8%)	(12.2%)
Very difficult	(1.1%)	(5.6%)
Did the game get easier over time?		
Yes	(66.0%)	(70.6%)
No	(34.0%)	(29.4%)
Did you guess or reason?		
Just guessed	(3.2%)	(3.0%)
Used reasoning a little bit	(13.8%)	(6.6%)
Used reasoning sometimes	(13.8%)	(8.6%)
Used reasoning a lot	(36.2%)	(22.2%)
Used reasoning all the time	(33.0%)	(59.6%)
Was your reasoning always the same?		
Yes	(23.1%)	(19.0%)
No	(76.9%)	(81.0%)

Note: Summary of questions asked to participants at the end of the experiment about the difficulty of the games and how they picked a guess.

We also find ample evidence of cooperation in teams. The vast majority of participants reported listening to other team members and influencing them, at least part of the time. The also perceived that their opinion was reflected in the guesses made by the team captain and, in 61% of the cases, reported no disagreement within the team. The majority of participants perceived that their team performed well (71%) and cooperated well (76%).

Participants also report a high amount of communication among them, primarily by phone call and SMS (Table E.4). Only 3-4% of participants report no phone calls received or made per day with members of their team, with an average of 4.3 calls made and 4.1 calls received from team members. The was also a sizable use of SMS messages – around

Table C2: Cooperation in teams

	Factor-variable percent
Did you listen to others?	
Never	(1.5%)
A little bit	(2.5%)
Sometimes	(9.5%)
A lot	(13.1%)
All the time	(73.4%)
Did you influence others?	
Never	(6.6%)
A little bit	(3.5%)
Sometimes	(21.7%)
A lot	(15.7%)
All the time	(52.5%)
Was you opinion reflected?	
Never	(5.6%)
A little bit	(4.1%)
Sometimes	(29.6%)
A lot	(24.5%)
All the time	(36.2%)
Was there disagreement?	
Never	(60.6%)
A little bit	(11.1%)
Sometimes	(14.1%)
A lot	(5.1%)
All the time	(9.1%)
How well team performed?	
Very poorly	(1.5%)
Poorly	(2.5%)
Average	(25.1%)
Well	(39.2%)
Very well	(31.7%)
How well cooperated?	
Very poorly	(2.0%)
Poorly	(4.0%)
Average	(17.6%)
Well	(35.2%)
Very well	(41.2%)

Note: Summary of questions asked to participants at the end of the experiment about the level of cooperation in teams.

2 per day received and sent – but with a large fraction of participants (38-39%) not using them at all. Some participants also reported meeting in person with their team, but the overwhelming majority (87%) did not.

Table C3: Communication in teams

SMS sent per day		*SMS received per day*	
0	(37.8%)	0	(38.6%)
1	(9.6%)	1	(5.6%)
2	(8.5%)	2	(11.7%)
3	(28.2%)	3	(26.4%)
4	(10.6%)	4	(8.6%)
5	(1.1%)	5	(2.0%)
6	(3.2%)	6	(4.6%)
16	(1.1%)	7	(0.5%)
		8	(0.5%)
		10	(1.5%)
Phone calls made per day		*Phone calls received per day*	
0	(3.6%)	0	(3.0%)
1	(3.0%)	1	(4.0%)
2	(12.2%)	2	(16.6%)
3	(26.4%)	3	(29.6%)
4	(19.3%)	4	(14.6%)
5	(7.1%)	5	(7.5%)
6	(17.8%)	6	(13.6%)
7	(2.5%)	7	(1.0%)
8	(3.0%)	8	(5.0%)
9	(1.0%)	9	(2.0%)
10	(2.0%)	10	(1.0%)
12	(0.5%)	13	(0.5%)
15	(0.5%)	14	(0.5%)
16	(1.0%)	15	(0.5%)
		16	(0.5%)
Number of team meetings during the experiment			
0	(86.8%)	4	(1.0%)
1	(2.0%)	5	(0.5%)
2	(4.1%)	10	(1.0%)
3	(1.0%)	14	(2.0%)
		16	(1.5%)

Note: Summary of questions asked to participants at the end of the experiment about the amount of communication between team members.

F Details of the Tanzania Experiment

F.1 Game Instructions

Experimental Study of Strategic Interaction using Guessing Games

INSTRUCTIONS TO BE READ TO PARTICIPANTS

Welcome!

You are about to participate in an experiment to test how you make decisions.

You will shortly receive some simple instructions and complete the experiment over the course of the next 2 hours.

For participating in today's session, you will receive **10,000 TSh**.

In addition, you may earn a considerable additional amount, depending on the decisions you make.

On average, participants in this game win about **65,000 TSh on top of the participation fee**.

Your earnings will be determined by your decisions and by the decisions of other participants. We now explain in detail how the game works and how your winnings are exactly determined. We kindly ask you to pay attention to the explanation and if you have any questions, please raise your hand.

At the end of the session, you will be asked to sit an understanding test to evaluate if you have understood the rules of the game. It is important, therefore, that you pay attention and raise any questions you might have during this training session.

INSTRUCTIONS

The experiment has 22 rounds.

In each round, you will be matched randomly with another player (a new one in each round), who will be your opponent in that round.

Your opponent does not know who you are. They only know that they are playing against a student from the University of Dar es Salaam.

In each round, you and your opponent separately and independently make decisions called GUESSES. Together, your GUESS and your opponent's GUESS determine the number of POINTS that you earn in a round.

Earning more points increases your money payment at the end of the experiment, as explained below.

NOTE: Neither your guess nor your opponent's guess in a round affects how you and your opponents are matched or the games you face in the rest of the experiment.

To make a good guess, you must understand how your guess and your opponent's guess determines the number of points that you and they earn in every decision situation. Here is how.

In each decision situation, each player has a **MULTIPLIER**, a **LOWER LIMIT** and an **UPPER LIMIT**. These multipliers and limits are typically different for you and your opponent, and they change from round to round. In all other respects, the games are identical in all 22 rounds.

In each round, each player is asked to guess a number between their lower limit and upper limit. The opponent is asked to also guess a number.

1

If you guess outside your limit, your guess is automatically adjusted down to the upper limit (if it is above it), or up to the lower limit (if it is below it).

The aim of the game is to guess a number that is as close as possible to your opponent's guess *times* your multiplier.

The closer your guess is to your opponent's guess times your multiplier, the higher the number of points you earn in each game.

EXAMPLE:

A: YOUR LOWER LIMIT = 100 and YOUR UPPER LIMIT = 900
B: YOUR MULTIPLIER = 2
C: YOUR OPPONENT'S LOWER LIMIT = 100 and YOUR OPPONENT'S UPPER LIMIT = 600
D: YOUR OPPONENT MULTIPLIER = 2

YOU:

LOWER LIMIT	UPPER LIMIT
100	**900**

MULTIPLIER = 2

YOUR OPPONENT

LOWER LIMIT	UPPER LIMIT
100	**600**

MULTIPLIER = 2

You receive the most points if you can guess a number that is equal to your opponent's guess times 2. In order to win the game, you must try to understand what your opponent is going to guess and multiply it by your multiplier.

The further away you are from the perfect guess, the fewer points you make.

NOTE: Your guesses should not have decimal places.

LET'S TRY:

[The experimenter plays the mock game above between himself and another collaborator. They both make their guesses and at the end, the payoffs are worked out as follows – referring to the graph below].

[Player A picks 300; Player B picks 200]

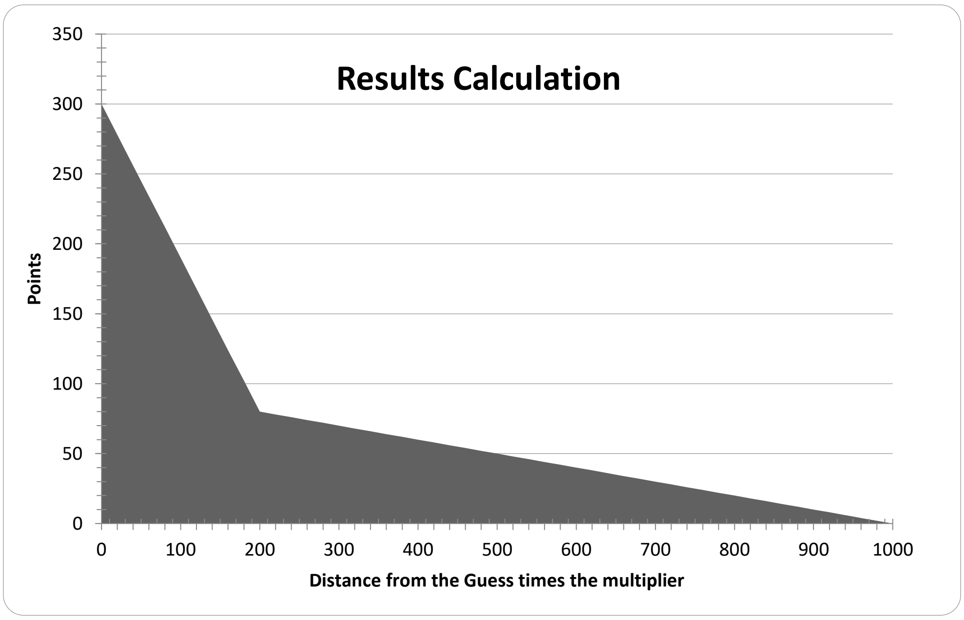
*[Player A would get the most points if s/he guessed the perfect guess of 200*2 = 400. Instead, s/he guessed 300 and is therefore 100 units away from the perfect guess]*

*[Player B would get the most points if s/he guessed the perfect guess 300*2 = 600. Instead, s/he guessed 200 and is therefore 400 units away from the perfect guess]*

The graph on the board (ENUMERATOR: Please point to the graph below, which should be printed and placed on the board) shows exactly how points are calculated, based on how far your guess is from the perfect guess, which is your opponent's guess times your multiplier.

In the example above, Player A earned 180 points, since the guess is 100 units away from the perfect guess and Player B earned 70 points, since the guess is 400 units away from the perfect guess.

You don't need to worry about the details in the graph. Just know that the further you are from the perfect guess, the fewer points you make!



You will play this type of game 22 times over the course of the experiment.

At the end, we will randomly choose one game and we will convert every point you made in that game into 300 TSh.

So, for example, if you earn 100 points in that game, we will pay you 30,000 TSh.

Please play every game carefully, as every game could be the one that determines your earnings at the end.

This is the end of the instructions. Any questions?

YOU CAN NOW START. AS PART OF THE EXPERIMENT, YOU WILL RECEIVE SOME INFORMATION AND YOU WILL BE ASKED SOME QUESTIONS ALONG THE WAY, PLEASE READ CAREFULLY AND PAY ATTENTION TO THE INFORMATION PROVIDED.

3

F.2 Questionnaire prompts

Figure D1: Introductory Message


UDSM Project Questionnaire

You will now play a series of 22 games, each against a different opponent. Your opponents are students from the University of Dar es Salaam.

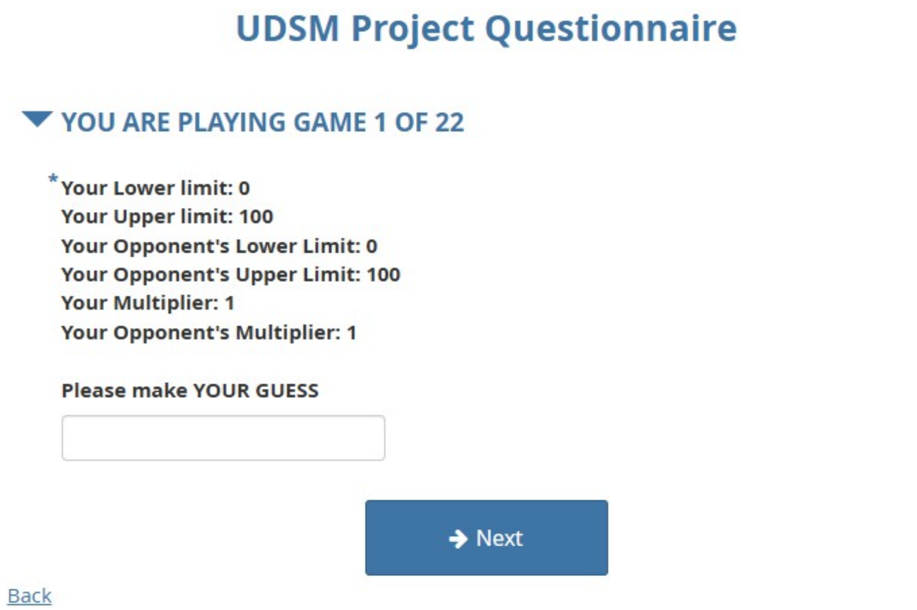
In each game you will know YOUR LOWER & UPPER LIMIT, YOUR OPPONENT'S LOWER AND UPPER LIMIT, YOUR MULTIPLIER, YOUR OPPONENT'S MULTIPLIER.

Please pay attention to this information.

Next

Back

Figure D2: Elicitation of Own Guess


UDSM Project Questionnaire

YOU ARE PLAYING GAME 1 OF 22

*Your Lower limit: 0
Your Upper limit: 100
Your Opponent's Lower Limit: 0
Your Opponent's Upper Limit: 100
Your Multiplier: 1
Your Opponent's Multiplier: 1

Please make YOUR GUESS

Next

Back

Figure D3: Belief Elicitation of Opponent's Guess

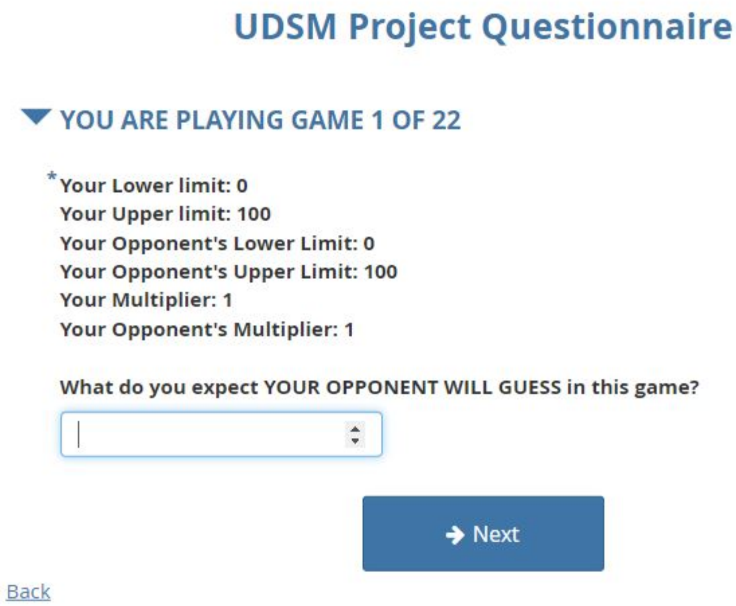


Figure D4: Screenshots from two videos