Paternalistic Persuasion

Alexandra Ballyk*

September 19, 2024

Abstract

Paternalistic experts ("Advisors") often seek to make decision-makers ("Choosers") better off by persuading them to change their behavior. Given that Choosers may be unwilling to make certain behavioral changes, an optimal recommendation accounts for Choosers' response to the recommended change. In a setting where Choosers are wary of recommenders' incentives, I experimentally investigate whether Advisors send optimal recommendations and why they may fail to do so. I find that up to 79% of Advisors fail to send optimal recommendations. Instead, up to two-thirds send recommendations that would only be optimal if Choosers were required to follow them. When it comes to correcting this mistake, I show that prompting Advisors to think about Choosers' likely response to a recommended change is much more effective than repeating the interaction and providing feedback about Choosers' behavior. The initial mistake is thus consistent with a failure to focus on recommendations that Choosers are likely to accept.

JEL Classification Numbers: C91, D82, D83, D91.
Keywords: Paternalism, Strategic Communication, Failure of Contingent Thinking, Laboratory Experiment.

*Department of Economics, University of Toronto (email: alexandra.ballyk@mail.utoronto.ca). I am thankful to my supervisor Yoram Halevy for his support and guidance throughout all stages of this project. I also wish to recognize my committee members Tanjim Hossain and Anne-Katrin Roesler, who respectively provided invaluable feedback on the experimental design and theoretical framework. This paper has benefited from helpful suggestions from Billur Gorgulu and Stanton Hudja, as well as seminar participants at the University of Toronto and the CIREQ Ph.D. students' conference. Xiaoya Gao and Shanshan Liu provided outstanding research assistance. Financial support from the University of Toronto's economics department is gratefully acknowledged. The experiments were pre-registered on AsPredicted.org.

1 Introduction

From financial advisors and lawyers to doctors and coaches, many paternalistic experts seek to persuade decision-makers to change their behavior. One challenge these paternalists face, however, is that decision-makers are often reluctant to follow through with a recommended change. This reluctance may stem from intrinsic biases[1] that inhibit behavioral change, or extrinsic threats to the decision-maker's trust in the quality of the recommendation. In order to successfully persuade a decision-maker to change their behavior, a paternalist must therefore strategically account for the decision-maker's response to the recommended change. That said, it is unclear whether paternalists actually do so. If paternalists fail to send their recommendations strategically, diagnosing and correcting this failure may drastically improve their ability to steer decision-makers towards better choices.

This paper experimentally studies whether paternalists strategically account for decision-makers' responses to their recommendations, and why they may fail to do so. In naturally-occurring settings, paternalists are often uncertain about decision-makers' preferences and information sets. This makes it difficult to pin down paternalists' optimal recommendations; that is, the recommendations that maximize decision-makers' wellbeing, subject to the constraint of being accepted. Moreover, a failure to send optimal recommendations may be caused by uncertainty, limited strategic thinking, or both; it may be difficult to separate the two effects. Conducting a laboratory experiment allows me to overcome these challenges. First, I develop a model of a recommendation game, which provides testable predictions about paternalists' optimal recommendations. I then test those predictions in a controlled laboratory setting that provides paternalists with complete certainty about decision-maker's incentives, action sets and information set.

The recommendation game captures the notion that decision-makers may have doubts about advisors' expertise, or worry about them having ulterior motives. For instance, doctors have been shown to recommend more highly-compensated procedures, which are not always best for their patient's health (Gruber and Owings, 1994). Similar behavior is observed among commissions-based financial advisors, whose recommendations may embody their commission incentives, their limited knowledge of certain financial products, or both (Anagol et al., 2017). As a result, decision-makers tend to be more willing to follow recommendations from individuals they trust (Alsan and Eichmeyer, 2024; Cole et al., 2013).

The recommendation game features multiple states, each with a corresponding action. A Chooser with a non-uniform prior over states aims to minimize the distance between the action they take and the action corresponding to the true state. They must decide between stick-

[1] For instance, decision-makers are often impatient (Laibson, 1997) or tempted by inferior alternatives (Gul and Pesendorfer, 2001), and are in general reluctant to deviate from their status-quo behavior (Samuelson and Zeckhauser, 1988). Changing one's behavior thus often requires overcoming these intrinsic biases.

ing with the action that is most likely to correspond to the true state according to their prior (their "status-quo"), or taking an alternative action that has been recommended to them by an Advisor. However, the Chooser is unsure whether the recommended action is in their best interest, as the Advisor may be one of two types. A "paternalistic" type knows the true state and has the same objective as the Chooser. A "commitment" type always recommends an action that is far from the Chooser's status quo. It is therefore not always optimal for the Chooser to accept a recommendation to take that action. This in turn means that it is not always optimal for the paternalistic type to recommend that action when it corresponds to the true state. More precisely, the recommendation game can have a "trusting" equilibrium where the Chooser accepts all recommendations, and the paternalistic type always recommends the action corresponding to the true state (a "full truth-telling" strategy). But in order for that equilibrium to exist, the Chooser's expected payoff to accepting the recommendation that a commitment type sends must be sufficiently high. Otherwise, the game only has a "wary" equilibrium where the Chooser rejects the recommendation that a commitment type sends. Thus, when the action that the commitment type recommends corresponds to the true state, it is optimal for the paternalistic type to recommend the action adjacent to it, which the Chooser will accept (a "partial truth-telling" strategy).

Experimental participants play as Choosers and paternalistic Advisors. The commitment type is replaced by a "computer" that always recommends the action most distant from the Chooser's status-quo. As I am interested in studying the recommendations Advisors send in situations where Choosers do not accept all recommendations, I select parameters of the recommendation game for which only a wary equilibrium exists. Furthermore, to make the game easy for participants to understand, I frame it as a "Plant Pot" Game where Choosers aim to produce the largest possible plant by turning on a sprinkler that is as close as possible to a pot with a plant seed in it.

The Plant Pot Game is the core task in the main experiment, which seeks to establish whether Advisors send optimal recommendations. The experiment's secondary objective is to determine why Advisors may fail to do so, focusing on two potential mechanisms. First, Advisors may not believe that Choosers are playing their wary equilibrium strategy, and may be best responding to this belief. This would be consistent with a model of level-k thinking (see, e.g., Camerer et al. (2004)): for instance, Advisors may be level-1 thinkers who are best-responding to the belief that level-0 Choosers accept all recommendations. To explore the role of Advisors' beliefs about Choosers' strategies, participants play twenty rounds of the Plant Pot Game, and receive feedback about their co-player's strategy at the end of each round. Second, Advisors may suffer from a failure of contingent thinking (Niederle and Vespa, 2023): they may fail to focus on the contingency that is actually realized, which in this case would be the strategy that Choosers actually play. To explore the role of a failure of contingent thinking, I allow Advisors to condition their strategy on the Chooser's in the final round. This effectively reverses the sequence of the game, which may help Advisors focus on the relevant Chooser strategy contingency.

I begin by establishing that the recommendations specified by Advisors' partial truth-telling strategy are empirically optimal. In every round, a clear majority of Choosers reject the recommendation that the computer always sends. Nevertheless, only 21% of Advisors send optimal recommendations in the first round. Instead, nearly two-thirds (63%) play their full truth-telling strategy. That is, most Advisors send recommendations that *would* be optimal *if* Choosers accepted all recommendations. Repetitions and feedback have a modest effect on Advisors' strategies: 45% send optimal recommendations in the penultimate round. Reversing the sequence of the game has a much larger effect. In the final round, where Advisors can condition their strategy on the Chooser's, 78% play their partial truth-telling strategy when Choosers reject the computer's recommendation. As a result, Advisors' failure to send optimal recommendation appears driven by a failure of contingent thinking.

I run a second experiment, the beliefs experiment, to further explore the explanatory power of the failure of contingent thinking mechanism. Decision-makers are often able to overcome a failure of contingent thinking if encouraged to focus on the relevant contingency (Esponda and Vespa, 2023). The design of the beliefs experiment is inspired by this finding. Specifically, before certain rounds of the Plant Pot Game, I elicit Advisors' beliefs about the modal Chooser strategy in the upcoming round. If the failure to send optimal recommendations is driven by a failure of contingent thinking, Advisors who focus on the relevant Chooser strategy contingency should be more likely to send optimal recommendations.

The beliefs experiment's results support the failure of contingent thinking mechanism. In all rounds that Advisors' beliefs are elicited, most Advisors believe Choosers will reject the recommendation that the computer always sends. Stated otherwise, the belief elicitation task prompts most Advisors to focus on the relevant contingency. As a result, a much larger proportion of Advisors send optimal recommendations in the beliefs experiment. The proportion doing so begins at 32%, and reaches 68% by the final round.

To summarize the results of both experiments, encouraging Advisors to focus on Choosers' likely response to their recommendations drastically improves their ability to steer Choosers towards better decisions. These results have important implications for assessments of the efficacy of paternalistic interventions. Globally, there are over 200 government units dedicated to the design and implementation of "nudges" (OECD, 2017), or light-touch behavioral interventions that seek to improve decision-makers' wellbeing (Thaler and Sunstein, 2009). The effectiveness of such an intervention, which may determine whether it is implemented at a larger scale, is often measured using its effect on decision-makers' behavior (DellaVigna and Linos, 2022). My results suggest, however, that a null effect on behavior does not necessarily indicate that the intervention has no potential. Rather, redesigning the intervention so as to better account for how decision-makers respond to it may make it much more effective. Training nudge unit policymakers to put themselves "in decision-makers' shoes" may be a promising way to optimize the design and impact of paternalistic interventions.

This paper links an emerging literature on paternalistic behavior to the literature on strategic communication. The existing paternalism literature has established when and why paternalists prefer removing alternatives from a Chooser's choice set (a "hard" intervention) over preserving their freedom of choice (a "soft" intervention). While Advisors generally prefer soft interventions (Bartling et al., 2023), when they do implement hard interventions, they aim to align Choosers' choices with their aspirations (Ambuehl et al., 2021). In my experiment, Advisors must recommend an alternative to Choosers' status-quo. Advisors are thus limited to a soft intervention, which is the only type of intervention that most paternalists have at their disposal. For instance, an employer may be able to set a default or recommended retirement savings contribution rate, but may not have the power to enforce that rate. Furthermore, prior work on the value of decision rights suggests that Choosers prefer soft interventions to hard ones (Bartling et al., 2014; Fehr et al., 2013). However, unlike a hard intervention, a soft intervention is only effective if the paternalist has strategically accounted for the Chooser's response to it. This paper shows that paternalists struggle to do so on their own, establishes the root of this mistake, and identifies interventions that help them overcome it. In doing so, I also add to the literature on failures of contingent thinking (see Niederle and Vespa (2023) for a review) by identifying a novel environment where such failures occur. These failures have been documented in several other strategic environments, including auctions (Kagel and Levin, 1986; Thaler, 1988), voting problems (Esponda and Vespa, 2014, 2023; Ali et al., 2021), public good contribution games (Calford and Cason, 2024) and market interactions (Ngangoué and Weizsäcker, 2021).

Previous experimental studies of strategic communication have focused on the interaction between an uninformed receiver and a single type of informed sender, whose incentives are imperfectly aligned with the receiver's (see Blume et al. (2020) for a review). Senders in such games tend to "over-communicate": they reveal the true state more frequently than theoretically predicted (Cai and Wang, 2006; Kawagoe and Takizawa, 2009; Wang et al., 2010). These findings are in line with my own, which suggests that over-communication is a common deviation from equilibrium in strategic communication environments. This paper is most closely related to Altmann et al. (2022), who study how default-setters' incentives impact the informativeness of the defaults they set, and thus whether decision-makers benefit from following them. In a binary-choice setting, the authors find that default-setters whose incentives are more closely aligned with decision-makers' select defaults that benefit decision-makers more. In this paper, I keep the proportion of Advisors of each type - and thus the expected degree of alignment of players' preferences - constant. What's more, Choosers' action set includes a "middle ground" action that is between their status-quo and an action they're wary of taking.

These design choices allow me to identify whether Advisors strategically account for Choosers' wariness when sending recommendations.

The rest of the paper is structured as follows. In Section 2, I develop a model that characterizes players' optimal strategies in the recommendation game. Sections 3 and 4 describe the design and results of the main and beliefs experiments, respectively. Section 5 concludes by providing a discussion of the experimental and theoretical results.

2 Theoretical Framework

To formalize the idea that paternalistic Advisors should send recommendations strategically, I develop a model of a recommendation game. The model is prescriptive: its purpose is to outline the strategies that players *should* play, which may not be those that they *actually* play empirically. The recommendation game is based on a simplified version of the classic strategic communication model (Crawford and Sobel, 1982), which features a single type of Sender whose preferences are known to the Receiver. A paternalistic Advisor, however, may not be able to signal to the Chooser that their preferences are fully aligned. Choosers may thus be unsure whether an Advisor's recommendation is truly in their best interest. The recommendation game incorporates this uncertainty by introducing two types of Advisors, where an Advisor's type captures the alignment of their preferences with the Chooser's. Furthermore, paternalistic Advisors often seek to persuade Choosers to change their behavior. As a result, they often search for, or provide an alternative to, an action the Chooser is already taking. Thus, while Receivers in the classic model are not biased towards any action, Choosers in the recommendation game must decide whether to retain a status-quo action.

2.1 The Recommendation Game

There are three states, $\theta \in \{1,2,3\}$, each with a corresponding action $a \in \{1,2,3\}$. Let $\hat{a}(\theta)$ denote the action corresponding to the true state. There are two players, a *Chooser* and an *Advisor*, who have different information about the true state. The Chooser's prior belief assigns probability $q_\theta > 0$ to state θ, where $q_1 \geq q_3$ and $q_1 + q_2 + q_3 = 1$.

The Advisor may be one of two types. The information the Advisor has, as well as their preferences, depends on which type they are. With probability p, the Advisor is a *paternalistic* type: they know the true state, and their utility function is the same as the Chooser's. With probability $1-p$, the Advisor is a non-strategic *commitment* type: they always recommend action 3. The paternalistic type represents an expert who has the Chooser's best interest at heart, such as highly-trained financial advisors and lawyers who have a fiduciary duty[2] to their

[2]An individual with a fiduciary duty to their client is legally required to act in their client's best interest.

clients. The commitment type's behavior has two possible interpretations. First, they may be an Advisor who knows the true state, but whose incentives are not fully aligned with the Chooser's. An example may be a financial advisor who is pressured to sell certain financial products in order to meet the bank's sales goals. Second, suppose that before the beginning of the game, the Advisor holds a prior belief that assigns the greatest weight to state 3, and must exert effort to learn the true state. The commitment type could be interpreted an Advisor who has not exerted the effort to learn the true state. For instance, a coach may recommend that an athlete follow a specific training plan because it works well for the *average* athlete, without determining whether the plan will work well for that *particular* athlete.

The Chooser's goal is to minimize the distance between the action they take and the action corresponding to the true state. Their preferences are described by the utility function $u(\alpha|a_c(r) - \hat{a}(\theta)|)$, where $\alpha > 0$, $u'(\cdot) < 0$ and $u''(\cdot) = 0$. A paternalistic Advisor has the same preferences as the Chooser. When a paternalistic Advisor is indifferent between recommendations, I assume that they recommend the action closest to $\hat{a}(\theta)$. This assumption effectively states that paternalistic Advisors prefer to tell the truth when indifferent between recommendations, which is in line with evidence on the psychological cost of lying (Abeler et al., 2019).

The game proceeds in three stages. First, Nature determines the state and the Advisor's type. Second, the Advisor recommends an action $r(\theta)$ to the Chooser. If the Advisor is a commitment type, they recommend $r(\theta) = 3$ for all θ. If the Advisor is a paternalistic type, they observe the realized state and can recommend any action; that is, $r(\theta) \in \{1, 2, 3\}$. Third, the Chooser observes the recommendation r, and either accepts it or rejects it. Accepting entails taking the recommended action, while rejecting entails taking action 1; that is, $a_C(r) \in \{1, r\}$. Recall that action 1 is most likely to correspond to the true state according to the Chooser's prior. I call action 1 the Chooser's *status-quo*,[3] which can be interpreted as an action the Chooser is already taking, or one that they have independently determined to be most likely to correspond to the true state. The structure of the Chooser's prior thus captures a form of motivated reasoning (Epley and Gilovich, 2016; Kunda, 1990) where the Chooser believes that actions far from their status-quo are weakly less likely to be in their best interest. The Chooser's action space captures situations where the Advisor makes actions available to the Chooser, either by providing the opportunity to take the action or by alleviating search costs. For instance, a prospective home-buyer may hire a realtor to alleviate the costs of searching for a new home, and the buyer may only be able to purchase a new home if their realtor has arranged for them to make an offer on it.

[3]The fact that action 1 is the Chooser's status-quo is a behavioral assumption. That is, under the assumed structure of the Chooser's prior ($q_1 \geq q_3$), action 1 is not necessarily the Chooser's optimal action based on their prior. Action 1 is the Chooser's ex ante optimal action for priors satisfying the slightly stronger condition $q_1 \geq q_2 + q_3$.

2.2 Equilibria

Depending on the values of certain parameters, the recommendation game has up to two pure-strategy[4] weak perfect Bayesian equilibria. Propositions 1 and 2 describe the equilibria, and Proposition 3 shows that no other equilibria exist. In Section 5, I discuss the robustness of these results to different variations of the model's assumptions. All proofs are provided in Appendix A.

It is natural to think that the recommendation game may have an equilibrium where the paternalistic Advisor always recommends the action corresponding to the true state. Proposition 1 shows that such an equilibrium can exist, but only for certain parameter values.

PROPOSITION 1: *When $q_3 \geq q_1(1-p)$, the recommendation game has a "trusting" equilibrium where...*

- *The paternalistic Advisor plays a "full truth-telling" strategy: they always recommend the action corresponding to the true state.*
- *The Chooser plays an "accept all" strategy: they accept all recommendations.*

Intuitively, when the Chooser accepts all recommendations, it is optimal for the paternalistic Advisor to recommend the action corresponding to the true state. However, the Chooser will only want to accept a recommendation of 3 if their expected payoff to doing so is sufficiently high. This occurs when the probability of state 3 occurring is at least as high as the Chooser's likelihood of receiving the commitment type's recommendation in state 1.

It may be surprising that a trusting equilibrium does not always exist. That said, the recommendation game does have an equilibrium that exists for all parameter values, which is described in Proposition 2.

PROPOSITION 2: *For all parameter values, the recommendation game has a "wary" equilibrium where...*

- *The paternalistic Advisor plays a "partial truth-telling" strategy: they recommend 2 when the state is 3, and otherwise recommend the action corresponding to the true state.*
- *The Chooser plays an "accept-reject" strategy: they accept a recommendation of 2 and reject a recommendation of 3.*

Proposition 2 outlines the paternalistic Advisor's optimal strategy when Choosers reject a

[4]For the purpose of specifying testable hypotheses for the experiments, I restrict my attention to pure-strategy equilibria. This is because experimental participants do not often use mixed strategies. For instance, in Prisoner's Dilemma experiments, most participants tend to play one of several pure strategies (Dal Bó and Fréchette, 2018), and many participants play pure strategies even when they have the opportunity to explicitly randomize between actions (Romero and Rosokha, 2023).

recommendation that they are wary of. Intuitively, since the Chooser accepts a recommendation of 2 but rejects a recommendation of 3, it is optimal for the paternalistic Advisor to recommend action 2 when the state is 3. Proposition 3 shows that the equilibria described in Propositions 1 and 2 are the unique pure strategy equilibria of the recommendation game.

PROPOSITION 3: *The trusting equilibrium and the wary equilibrium are the unique pure-strategy weak perfect Bayesian equilibria of the recommendation game.*

3 Main Experiment

My main research objective is to determine whether, in situations where Choosers are reluctant to make certain behavioral changes, Advisors account for this reluctance when issuing recommendations. The main experiment thus features a version of the recommendation game where Advisors are required to account for such reluctance: one where only a wary equilibrium exists. Furthermore, I framed the recommendation game as a "Plant Pot" Game to make it easier to understand. Choosers aimed to produce the largest possible plant by turning on a sprinkler (i.e., taking an action) that was as close as possible to a pot with a plant seed in it (i.e., the true state). They did so by deciding whether to accept a recommendation regarding which sprinkler to turn on. Recommendations to turn on sprinkler 3 could either be from the computer (the commitment-type Advisor) or an Advisor (the paternalistic-type Advisor).

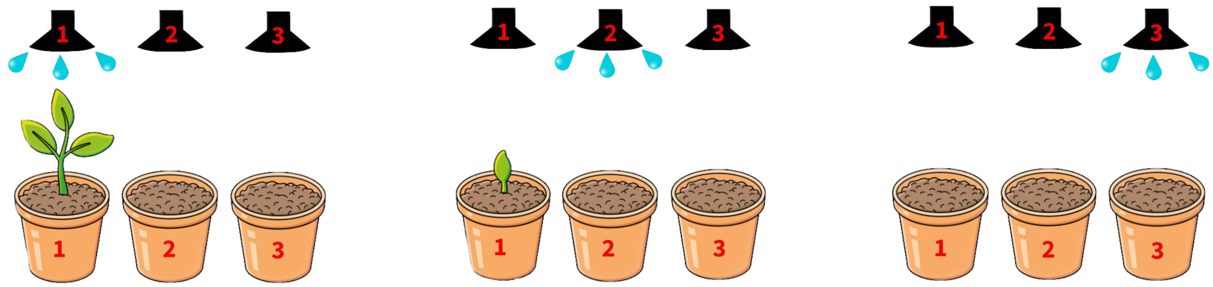
My secondary research objective is to shed light on why Advisors may fail to account for Choosers' reluctance to change their behavior. The main experiment assesses the explanatory power of two potential mechanisms. To investigate the role of uncertainty about Choosers' strategies, participants play multiple rounds of the Plant Pot Game, and receive feedback about their co-player's strategy at the end of each round. To investigate a possible failure of contingent thinking, Advisors play a round of the Plant Pot Game where they can condition their strategy on the Chooser's. This intervention, which effectively reverses the sequence of the game, is akin to an intervention that has been shown to mitigate failures of contingent thinking in committee voting problems (Esponda and Vespa, 2014).

Section 3.1 describes the design of the main experiment. Its results are presented and discussed in Section 3.2.

3.1 Experimental Design

The Plant Pot Game

The Plant Pot Game is the core task in the main experiment. The game features three pots arranged in a line. One of the pots - the "plant pot" - has a plant seed in it, whereas the other two do not. Each pot has a sprinkler above it. When a sprinkler is turned on, it sprays water directly below it, as well as below to the right and left of it. The pot closest to it receives the most water, while the pot farthest from it receives the least. As illustrated in Figure 1, a seed will grow into a larger plant if it receives more water.



(a) Pot 1 is the plant pot; sprinkler 1 is turned on

(b) Pot 1 is the plant pot; sprinkler 2 is turned on

(c) Pot 1 is the plant pot; sprinkler 3 is turned on

Figure 1: Illustration of the Plant Pot Game

Participants playing the role of Chooser are incentivized to produce the largest possible plant: they receive \$16 if they produce a large plant, \$13 if they produce a small plant, and \$10 if they do not produce a plant. Choosers know that pot 1 has a 9-in-20 chance of being the plant pot, pot 2 has a 7-in-20 chance and pot 3 has a 4-in-20 chance.

While sprinkler 1 will be turned on by default, Choosers may receive a recommendation to switch to one of the other sprinklers. If recommended to switch from the default, a Chooser must decide between accepting or rejecting the proposed switch, where rejecting entails sticking with the default sprinkler. If they are recommended to stick with the default, they automatically do so.

Choosers know that they have a 1-in-3 chance of receiving an Advisor's recommendation, and a 2-in-3 chance of receiving the computer's recommendation. Advisors are participants who have the same incentives as Choosers[5] and who know which pot is the plant pot. Advisors are also aware of all the information Choosers have when deciding whether to accept a recommendation. The computer recommends for the Chooser to switch to sprinkler 3 regardless of which pot is the plant pot.

In summary, the Plant Pot Game's structure and incentives are identical to those of the recommendation game, with a linear payoff function for the Chooser and paternalistic Advisor. It also features parameters for which only a wary equilibrium exists. As a result, it is optimal for Choosers to accept a recommendation to switch to sprinkler 2 and reject a recommendation to switch to sprinkler 3. Advisors should thus recommend the sprinkler corresponding to the

[5]If a Chooser received the computer's recommendation, the Advisor they were paired with received \$10 regardless of the size of the plant the Chooser produced.

plant pot when pots 1 and 2 are the plant pot, but recommend sprinkler 2 when pot 3 is the plant pot.

Structure of the Experiment

Experimental sessions were conducted with even numbers of participants. At the beginning of a session, half of the participants were randomly assigned to each role (Chooser or Advisor). Participants retained their role for the entire experiment.

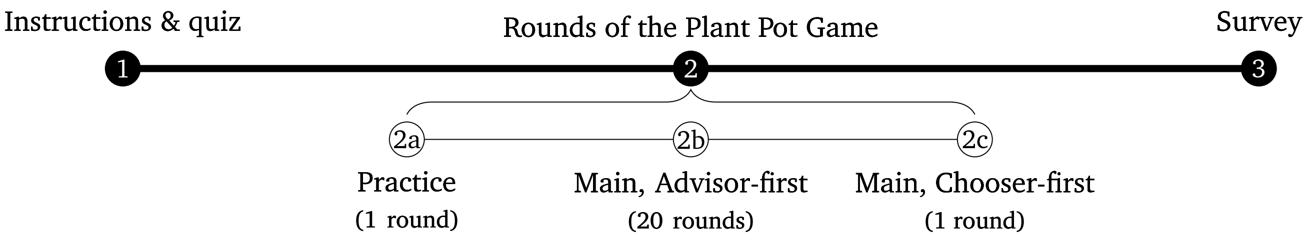


Figure 2: Structure of the experiment

Figure 2 summarizes the structure of the experiment. After reading the instructions, participants completed a comprehension quiz. They were required to complete all quiz questions correctly before proceeding to the first round. As an incentive to read the instructions carefully, participants received a $1 bonus payment if they answered all questions correctly on their first try.

Next, participants played 22 rounds of the Plant Pot Game. The first was an unpaid practice round, and the rest were "main" rounds that were eligible to count for payment. At the beginning of each round, Choosers and Advisors were randomly re-matched, and the plant pot and source of the recommendation shown to the Chooser were determined according to a draw from the relevant probability distribution. In all rounds, participants' strategies were elicited using the strategy method. Advisors specified the sprinkler they would recommend for each possible plant pot (i.e., pot 1, 2 and 3), and Choosers specified whether they would accept or reject each recommendation that involved switching from the default sprinkler (i.e., recommendations to switch to sprinkler 2 and to switch to sprinkler 3). At the end of each round, participants received feedback on the outcome of the round and their co-player's strategy. Thus, an Advisor learned how the Chooser they were paired with would have responded to all recommendations that involved switching sprinklers, regardless of which recommendation that Chooser actually received.

In the first 21 rounds, the game was played out in the sequence described in Section 2: Advisors sent their recommendations, then Choosers responded. In the final round, the sequence was reversed: Choosers committed to a response to each possible recommendation, then Advisors sent their recommendations. To implement this reversed sequence using the strategy method, I allowed Advisors to specify the sprinkler they would recommend for each possible plant pot and each possible strategy the Chooser could play (i.e., accept all, accept-reject, reject-accept and reject all). I refer to the first 20 main rounds as Advisor-first rounds, and the final main round as the Chooser-first round.

After completing all rounds of the Plant Pot Game, participants completed an end-of-experiment survey, which included questions about the rationale behind their strategies and a demographic questionnaire.

Procedures

Appendix B contains the experiment's instructions. The experiment was pre-registered on AsPredicted.org, programmed and deployed in oTree (Chen et al., 2016), and conducted at the Toronto Experimental Economics Laboratory (TEEL). Participants were recruited from the University of Toronto student body using the ORSEE online recruitment system (Greiner, 2015). I conducted seven sessions, each with 16 to 20 participants, for a total of 136 participants (68 Choosers, 68 Advisors).

Participants' total payment included a $5 show-up fee, their earnings from the comprehension quiz, and their earnings from one round of the game. All rounds except the practice round were eligible to count for payment. The round that counted for each participant was randomly selected at the beginning of the experiment and revealed to them at the end of the experiment. The average participant earned $18.69 and took 40 minutes to complete the experiment.

3.2 Results

I begin by establishing whether, given the empirical distribution of Choosers' strategies, it is indeed optimal for Advisors to play their wary equilibrium strategy. Next, I investigate whether Advisors play their empirically-optimal strategy. Unless otherwise specified, all analyses focus exclusively on the 20 Advisor-first rounds. I exclude one participant who withdrew from the experiment,[6] which yields a final sample of 135 participants (68 Choosers, 67 Advisors).

Which recommendations are optimal?

The model of the recommendation game predicts that, given the parameters used in the experiment, it is optimal for Advisors to play their partial truth-telling strategy; that is, to recommend the action corresponding to the state when the state is 1 or 2, and to recommend action 2 when the state is 3. This prediction relies on the assumption that a majority of Choosers play

[6]This participant was playing the role of Advisor, and withdrew during the survey.

their wary equilibrium strategy, which is to accept a recommendation of 2 and reject a recommendation of 3. Prior to analyzing Advisors' strategies, I therefore determine whether partial truth-telling is indeed a best-response to the empirical distribution of Choosers' strategies, and that Advisors are able to learn this through the feedback they receive.

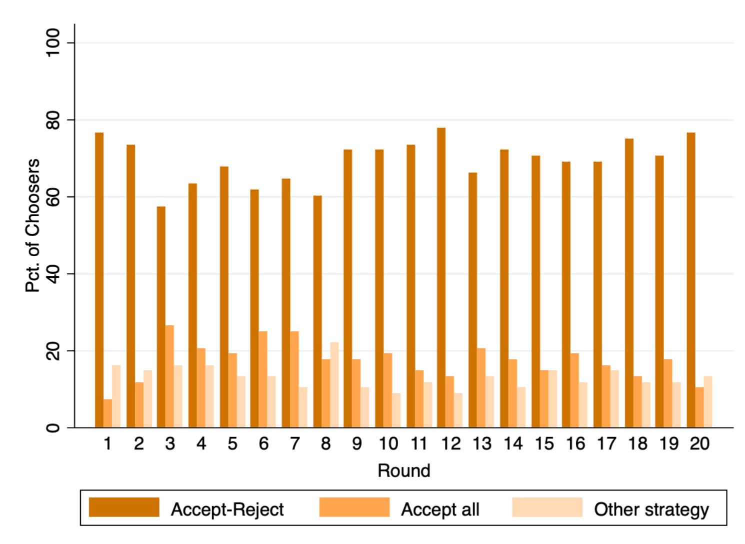


Figure 3: Distribution of Choosers' strategies

Notes. This figure displays the percent of Choosers in the main experiment playing accept-reject (their wary equilibrium strategy), accept-all (their trusting equilibrium strategy), and other strategies in a given round.

Consistent with the theoretical prediction, a clear majority of Choosers play their wary equilibrium strategy. Figure 3 displays the distribution of Choosers' strategies in each round. In the average round, 70% of Choosers play accept-reject, while at most 27% of Choosers play any other strategy in any given round. As a result, in every round, the modal Chooser plays accept-reject.

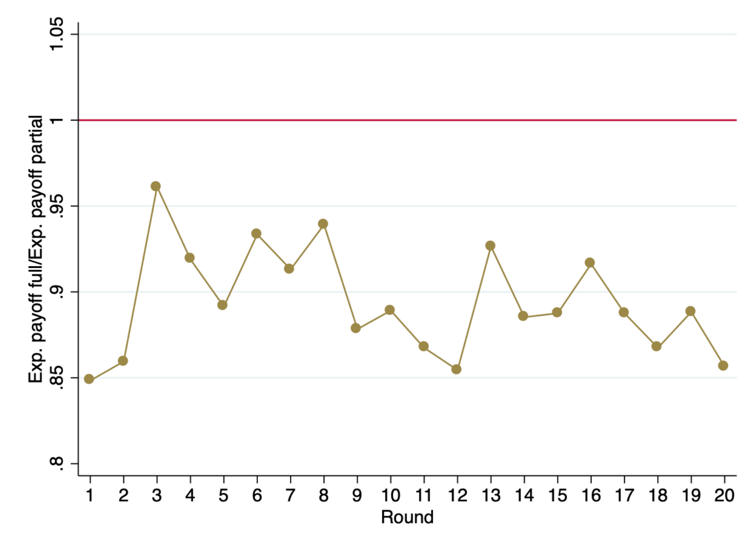


Figure 4: Expected state-3 payoff of Advisors' equilibrium strategies

Notes. This figure displays the expected state-3 payoff of Advisors' full truth-telling strategy as a fraction of the expected state-3 payoff of their partial truth-telling strategy. Expected payoffs are calculated using the empirical distribution of Choosers' strategies in the main experiment.

Partial truth-telling is indeed the best response to this empirical distribution of Choosers' strategies. Since Advisors' equilibrium strategies differ only in the action recommended in state 3, Figure 4 compares the expected state-3 payoff of these strategies,[7] given the empirical distribution of Choosers' strategies. In all rounds, partial truth-telling yields a higher expected payoff than full truth-telling: the fraction $\frac{\text{Exp. payoff of full truth-telling}}{\text{Exp. payoff of partial truth-telling}}$ is always below 1. Across all rounds, the expected payoff of partial truth-telling is up to 16% higher, with an average payoff difference of 12%.

RESULT 1: *Partial truth-telling is the best response to the empirical distribution of Choosers' strategies.*

[7]Let $P_{a_c(2)a_c(3)}$ be the percent of Choosers who play strategy $(a_c(2), a_c(3))$. The expected state-3 payoff to full truth-telling is $(P_{21} + P_{11}) \cdot (M - 2d) + (P_{23} + P_{13}) \cdot M$. The expected state-3 payoff to partial truth-telling is $(P_{21} + P_{23}) \cdot (M - d) + (P_{11} + P_{13}) \cdot (M - 2d)$.

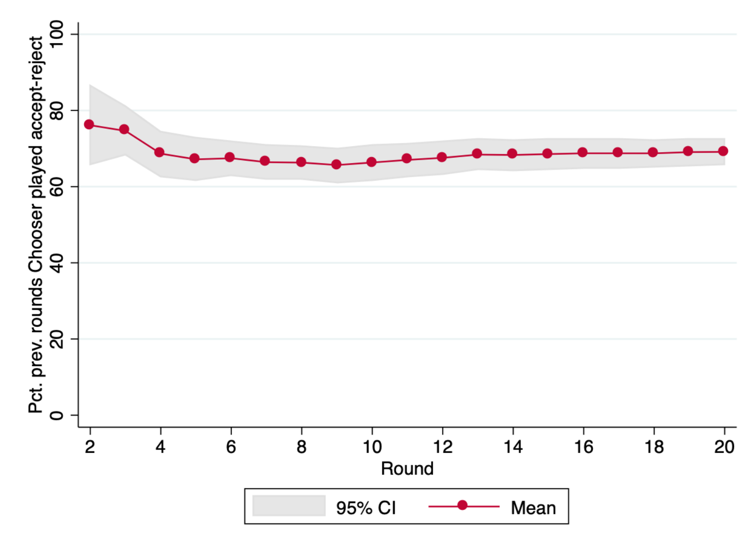


Figure 5: Distribution of feedback received by Advisors

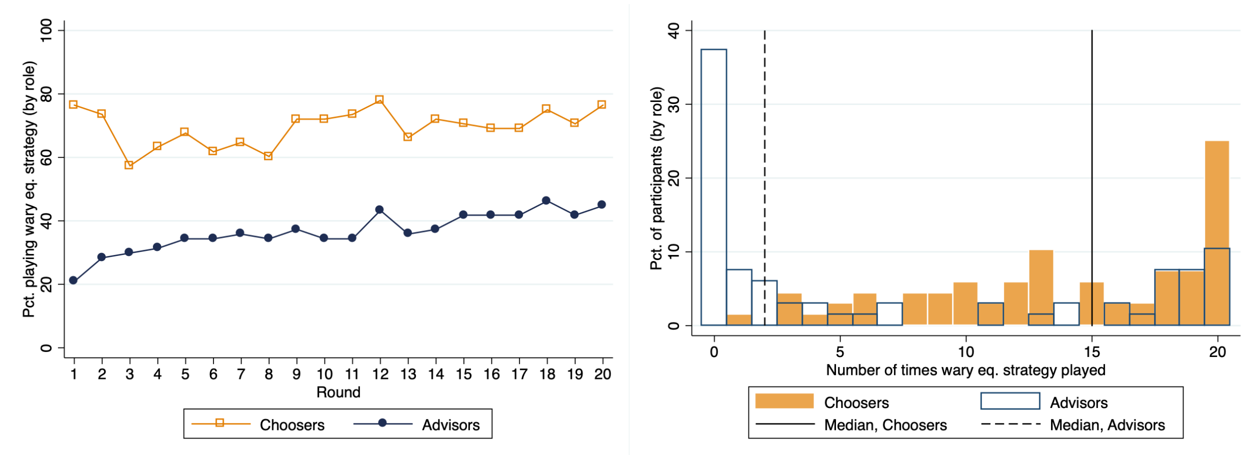
Notes. This figure summarizes the feedback Advisors in the main experiment have received (on the y-axis) when choosing their strategy in round x (on the x-axis). More precisely, it displays the percent of previous rounds in which Advisors received the feedback that the Chooser they were paired with played accept-reject. The pink line indicates the mean (across Advisors), and the grey area indicates the 95% confidence interval of that mean.

Furthermore, the fact that Choosers consistently play accept-reject is reflected in the feedback Advisors receive. Figure 5 displays, for rounds 2 through 20, the percent of previous rounds in which the average Advisor received the feedback that they were paired with a Chooser who played accept-reject. The average Advisor has received this feedback in 66% to 76% of prior rounds. The average frequency across all rounds is 69%, which is only one percentage point lower than the true average frequency with which Choosers play accept-reject.

RESULT 2: *Advisors learn that Choosers consistently play accept-reject.*

Result 1 confirms that the optimal recommendations are those specified by Advisors' wary equilibrium strategy, and Result 2 indicates that Advisors learn this. Next, I investigate whether Advisors send their optimal recommendations.

Do Advisors send optimal recommendations?



(a) Round-level **(b)** Player-level

Figure 6: Frequency of optimal strategy play

Notes. This figure displays the frequency with which main experiment participants in each role play their wary equilibrium strategies. In Figure 6a, frequency is defined at the round-level: the Figure displays the percent of participants in each role playing their wary equilibrium strategy in each round. In Figure 6b, frequency is defined at the player-level: the Figure displays the percent of participants in each role playing their wary equilibrium strategy in x out of the 20 rounds.

Figure 6 displays the frequency with which participants in each role play their wary equilibrium strategies. Figure 6a defines frequency at the round-level, while Figure 6b defines frequency at the player-level. Regardless of how frequency is defined, it is clear that only a minority of Advisors send optimal recommendations.

As shown in Figure 6a, in any given round, less than half of Advisors play their partial truth-telling strategy. The proportion doing so begins at 21% in the first round, reaches a maximum of 46% in round 18, and settles at 45% in the final round. As a result, the proportions of Advisors and Choosers playing their wary equilibrium strategies are significantly different in all rounds (proportions test, $p < 0.01$ in all rounds).

Similar results emerge when comparing the distributions of wary equilibrium strategy play at the player level. These distributions differ significantly by role (Wilcoxon rank-sum test, $p < 0.01$), with Choosers tending to play optimally in a larger number of rounds. As shown in Figure 6b, the median Chooser plays accept-reject in 15 rounds, and the modal Choosers - who make up 25% of all Choosers - do so in *all* rounds. In contrast, the median Advisor plays partial truth-telling in 2 rounds, and the modal Advisors - who make up 37% of all Advisors - do so in *none* of the rounds. Finally, whereas 63% of Advisors play optimally in fewer than half of all rounds, only 23% of Choosers do so.

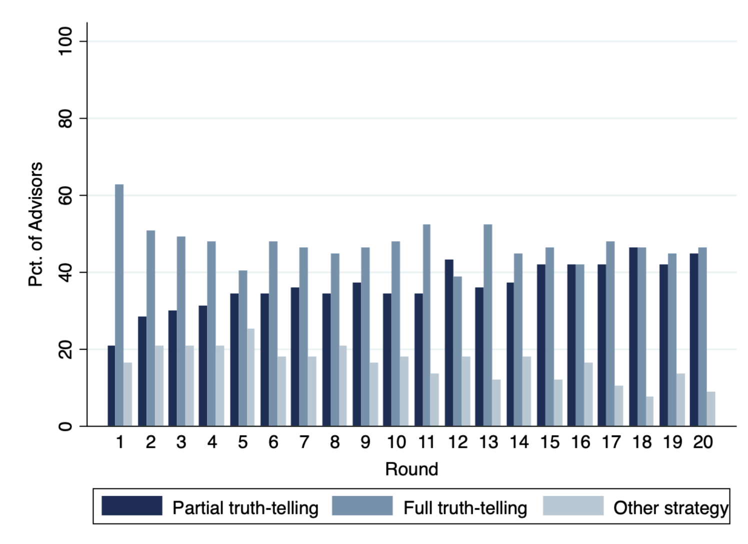


Figure 7: Distribution of Advisors' strategies

Notes. This figure displays the percent of Advisors in the main experiment playing partial truth-telling (their wary equilibrium strategy), full truth-telling (their trusting equilibrium strategy), and other strategies in a given round.

Having established that Advisors *do not* often send optimal recommendations, I now examine which recommendations they *do* send. Figure 7, which displays the distribution of Advisors' strategies in each round, shows that full truth-telling is consistently the most popular strategy. It is the modal Advisor strategy in 19 out of the 20 rounds, and 40% to 63% of Advisors play it in any given round.

RESULT 3: *A minority of Advisors play partial truth-telling in any given round. Full truth-telling is the modal Advisor strategy in almost all rounds.*

While full truth-telling is almost always the modal strategy, Advisors' strategies evolve as rounds progress. More precisely, partial truth-telling becomes more common. The proportion of Advisors playing partial truth-telling increases significantly between the first and last rounds (proportions test, $p < 0.01$), while the proportion playing full truth-telling significantly decreases (proportions test, $p = 0.06$).

Rounds	All	1-10	11-20
	(1)	(2)	(3)
Round	0.009^{***}	0.013^{**}	0.009^{**}
	(0.003)	(0.006)	(0.004)
Constant	0.226^{**}	0.157	0.280^{*}
	(0.115)	(0.111)	(0.144)
Observations	1340	670	670
Mean	0.37	0.32	0.41

Notes. Coefficients from OLS regressions conducted among Advisors in the main experiment, with standard errors in parentheses. The dependent variable is an indicator for playing partial truth-telling (Advisors' wary equilibrium strategy) in a given round, and the independent variable is the round number. All regressions include controls for the session and standard errors clustered at the participant-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 1: Evolution of Advisors' likelihood of playing partial truth-telling (main experiment)

Table 1 reports the results of regressions that further explore the evolution of Advisors' strategies. As seen in column 1, playing an additional round of the game increases an Advisor's likelihood of playing partial truth-telling by 2.4% (0.9 p.p.). The results from columns 2 and 3 reveal that this effect is larger in early rounds. In rounds 1 through 10 (column 2), playing an additional round increases an Advisor's likelihood of playing partial truth-telling by 4.1% (1.3 p.p.). However, in rounds 11 through 20 (column 3), an additional round only has a 2.2% (0.9 p.p.) effect.

RESULT 4: *Advisors become slightly more likely to play partial truth-telling as rounds progress, especially in early rounds.*

Result 4 suggests that gaining experience with the game, and receiving feedback about Choosers' strategies, helps Advisors learn which recommendations are optimal. However, by the final round, more than half of Advisors still fail to send optimal recommendations. The Chooser-first round is designed to help Advisors do so by allowing them to directly respond to Choosers' wary equilibrium strategy. Next, I examine how this reversal of the sequence of the game impacts Advisors' likelihood of sending optimal recommendations.

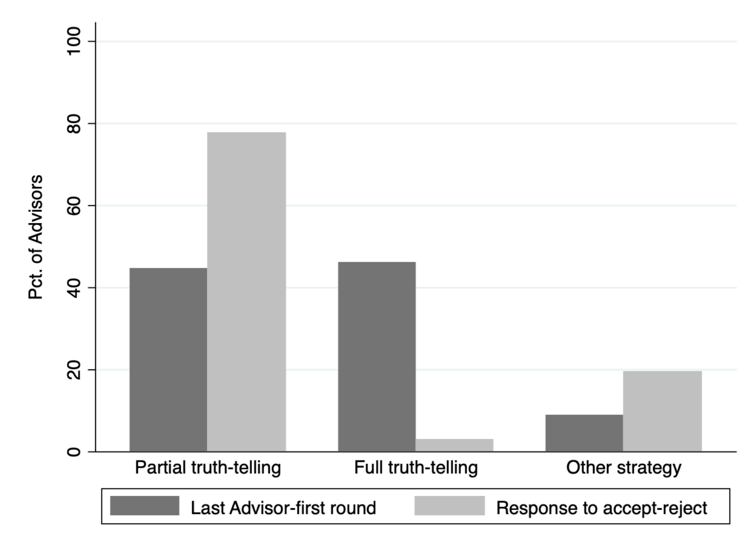


Figure 8: Advisors' strategies: last Advisor-first round vs. Chooser-first round

Notes. This figure displays the percent of Advisors in the main experiment playing partial truth-telling (their wary equilibrium strategy), full truth-telling (their trusting equilibrium strategy), and other strategies in two different scenarios. The darker bars indicate the percent of Advisors playing a given strategy in the last Advisor-first round. The lighter bars indicate the percent of Advisors playing a given strategy when responding to accept-reject (Choosers' wary equilibrium strategy) in the Chooser-first round.

Figure 8 compares Advisors' strategies in the final Advisor-first round to their responses to accept-reject in the Chooser-first round. Recall that the Chooser-first round occurs immediately after the final Advisor-first round, so Advisors do not gain more experience with the game between these two rounds. Nevertheless, Advisors are much more likely to send optimal recommendations when directly responding to Choosers' wary equilibrium strategy. 78% of Advisors play partial truth-telling in the latter scenario, in contrast to the 45% who do so in the final Advisor-first round. Similarly, the proportion playing full truth-telling decreases from 46% to 3%. Both of these changes are statistically significant (proportions test, $p < 0.01$ for both).

RESULT 5: *Advisors are much more likely to play their partial truth-telling strategy when directly responding to accept-reject.*

To summarize the results of the main experiment, only a minority of Advisors send optimal recommendations. Instead, they tend to send recommendations that *would* be optimal if Choosers accepted all recommendations, which Choosers do not. While Advisors learn their optimal recommendations as rounds progress, this learning effect is much smaller than the effect of reversing the sequence of the game. This version of the game effectively allows Advisors to send their recommendations after learning which recommendations Choosers accept.

Why might Advisors fail to send optimal recommendations? The learning effect rules out mechanisms that would require strategies to be unaffected by repetitions and feedback, such as a high psychological lying cost or the presence of a group of confused Advisors.[8] Moreover, the relative weakness of the learning effect implies that a lack of experience with the game, and/or uncertainty about Choosers' strategies, can only partially explain Advisors' behavior.

Instead, Advisors' behavior appears consistent with a failure of contingent thinking (FCT). Such failures are prevalent in decision problems where many contingencies (i.e., states) *could* be realized. One type of FCT occurs when decision-makers fail to focus on the contingency that is *actually* realized. In the game studied in this experiment, while there are four possible strategies Choosers could play, a clear majority of Choosers play one of these strategies. Advisors may fail to think contingently by failing to focus their best response to that strategy.

Decision-makers who fail to think contingently tend to exhibit several behavioral hallmarks (Niederle and Vespa, 2023). First, they tend to play a "naïve" strategy instead of the optimal one. As summarized by Result 3, many Advisors did exactly that: instead of playing a strategy that accounts for the fact that Choosers reject a specific recommendation, they play a strategy that *would* be optimal *if* Choosers accepted all recommendations. Second, repetitions and feedback do little to fix the FCT. The weak learning effect summarized by Result 4 is consistent with this trend. Third, many decision-makers overcome the FCT if placed in the relevant contingency or encouraged to focus on it. Changing the timing of a strategic game is an effective way of placing decision-makers in the relevant contingency (Esponda and Vespa, 2014). Result 5 indicates that Advisors react quite strongly to such an intervention, becoming much more likely to send optimal recommendations.

The results of the main experiment thus suggest that Advisors fail to send optimal recommendations because they fail to think contingently. If so, simply prompting Advisors to focus on the relevant contingency - that is, the recommendations Choosers accept - should make them more likely to send optimal recommendations. As an additional test of the mechanism, I therefore conduct a second experiment that allows me to study the effect of such an intervention.

[8] Beyond the presence of a learning effect, there are several pieces of evidence that rule out lying costs and confusion. If it were costly to play non-truth-telling strategies, Advisors would play full truth-telling in all contingencies in Chooser-contingent strategy round. However, only 6% of Advisors do so. Furthermore, the experiment was designed so as to eliminate, or at least significantly mitigate, confusion. As mentioned in Section 3.1, participants were incentivized to read the instructions carefully, and were required to answer all comprehension questions correctly before proceeding to the first round. Additionally, Advisors who never send optimal recommendations do not make more mistakes on the comprehension quiz (Wilcoxon rank-sum test, $p = 0.29$), which suggests that this group was not exceptionally confused.

4 Beliefs Experiment

In the beliefs experiment, Advisors are occasionally prompted to predict the Chooser's strategy before submitting their own. This design choice is inspired by the finding that many participants overcome FCTs if they focus on the relevant contingency, even if not directly placed in it (Esponda and Vespa, 2023). Furthermore, such strategy-prediction tasks have been shown to increase Nash equilibrium strategy play (Croson, 2000, 1999), presumably because they prompt participants to think about which of their opponent's actions they should best respond to. The beliefs experiment thus provides an additional test of the failure of contingent thinking mechanism, which predicts that Advisors who focus on the relevant Chooser strategy contingency (accept-reject) will be more likely to send optimal recommendations. Section 4.1 describes the design of the beliefs experiment. Its results are presented and discussed in Section 4.2.

4.1 Experimental Design

Structure of the Experiment

The basic structure of the beliefs experiment was identical to that of the main experiment (see Figure 2 for a reminder). The only difference was that, before certain rounds, I elicited participants' beliefs about their co-player's strategy in the upcoming round. More precisely, participants indicated their belief about the modal strategy played by participants in the opposite role in previous sessions of the experiment. As mentioned earlier, prior work suggests that eliciting Advisors' beliefs about Choosers' strategies could make them more likely to play their optimal strategy. To ensure a similar effect on Choosers, Choosers' belief elicitation task entailed specifying their beliefs about Advisors' strategies and re-stating the recommendation the computer always sent.

Beliefs were elicited before every sixth Advisor-first round, starting with the first one. There were thus four belief elicitation tasks, which occurred before the first, seventh, thirteenth and nineteenth Advisor-first rounds. These tasks were incentivized using a simple scheme that paid \$5 if the stated beliefs were correct,[9] and nothing otherwise.[10] Choosers were additionally required to correctly re-state the computer's recommendation in order to receive the \$5 payment.

[9]For payment purposes, the correct modal strategies were the modal strategies in the main experiment. As a reminder, in all rounds of the main experiment, the modal Chooser played accept-reject, and the modal Advisor played their full truth-telling strategy.

[10]I chose this simple incentive scheme over a more complex one (e.g., a scoring rule) for two reasons. First, recent evidence suggests that more complex incentive schemes may confuse participants (Danz et al., 2022). Second, the primary goal of this experiment was to prompt Advisors to focus on the relevant Chooser strategy contingency. Using a scoring rule would have involved eliciting Advisors' beliefs about the probability of Choosers playing each of their strategies, which would have made Advisors less likely to focus on the relevant Chooser strategy.

Procedures

Appendix B contains the experiment's instructions. Similar to the main experiment, the beliefs experiment was pre-registered on AsPredicted.org, programmed and deployed in oTree (Chen et al., 2016), and conducted at the TEEL with participants recruited via ORSEE (Greiner, 2015). I conducted seven sessions, each with 16 to 22 participants, for a total of 130 participants (65 Choosers, 65 Advisors).

Participants' total payment included a \$5 show-up fee, their earnings from the comprehension quiz, and their earnings from one main round of the game and one belief elicitation task. The round and belief elicitation task that counted for each participant was randomly selected at the beginning of the experiment and revealed to them at the end of the experiment. The average participant earned \$21.71 and took 41 minutes to complete the experiment. In Appendix C, I show that participants' characteristics did not differ significantly between the two experiments.

4.2 Results

I first establish whether Advisors focused on the relevant contingency. That is, did most Advisors believe that Choosers played accept-reject? Next, I investigate whether participants' strategies differ across experiments. Recall that if Advisors failed to send optimal recommendations because they failed to think contingently, Advisors who focus on Choosers' modal strategy should be more likely to send optimal recommendations. As with the main experiment, I focus exclusively on the 20 Advisor-first rounds and drop participants who withdrew from the experiment.[11] The final sample includes 129 participants (64 Choosers, 65 Advisors). All reported p-values are from two-sided tests.

[11] One participant withdrew from the beliefs experiment. They were playing the role of Chooser, and withdrew after round 11. Given that one-to-one matching of Choosers and Advisors requires an even number of participants to be present in all rounds, to avoid having to stop the session, a research assistant played at this participant's terminal for the remainder of the rounds. The research assistant played the modal Chooser strategy in all rounds of all prior experimental sessions, which was accept-reject.

Do Advisors focus on the relevant contingency?

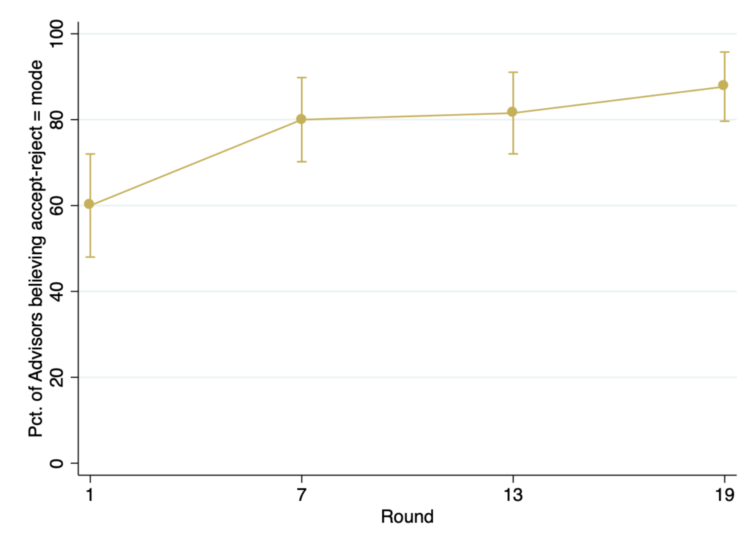


Figure 9: Advisors' beliefs about Choosers' strategies

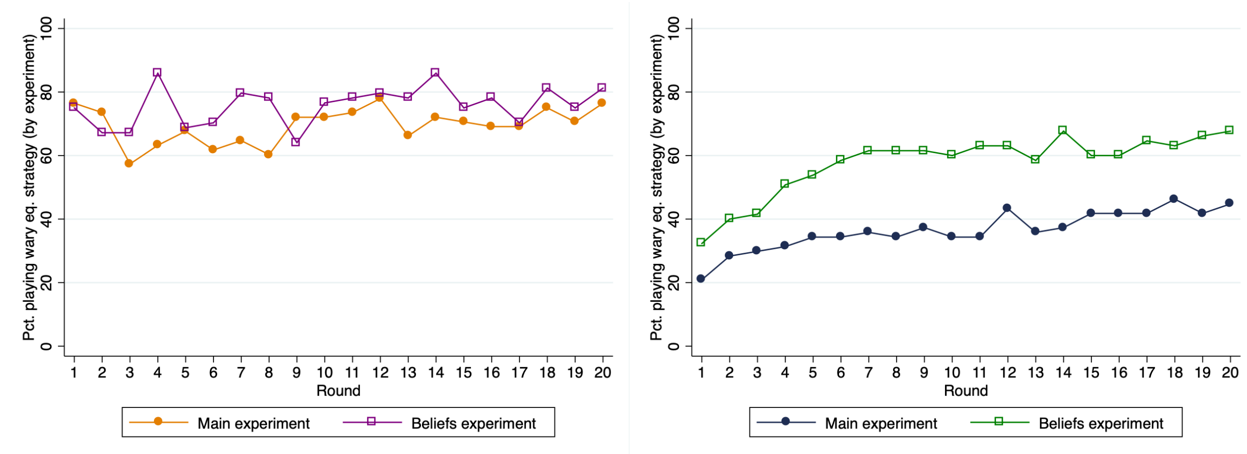
Notes. This figure displays the percent of Advisors in the beliefs experiment who believe that, in prior experimental sessions, the modal Chooser strategy in round x was accept-reject. Bars indicate the 95% confidence interval.

Figure 9 displays the percent of Advisors who believe accept-reject is the modal Chooser strategy in a given round. In every round that beliefs are elicited, a majority of Advisors believe accept-reject is the most-commonly played Chooser strategy. Moreover, more Advisors believe this as rounds progress, presumably due to the feedback they receive. Between rounds 1 and 19, the percentage increases from 60% to 88%. This increase, as well as the increase from rounds 1 to 7, are significant (proportions tests, $p \leq 0.01$ for both).

RESULT 6: *A majority of Advisors believe that accept-reject is the modal Chooser strategy in all rounds where beliefs are elicited.*

Result 6 indicates that eliciting Advisors' beliefs prompted most of them to focus on the relevant contingency. Next, I investigate whether, as predicted by a failure of contingent thinking, this intervention increases Advisors' likelihood of sending optimal recommendations.

How does eliciting Advisors' beliefs affect their strategies?



(a) Choosers

(b) Advisors

Figure 10: Frequency of optimal strategy play (round-level), by experiment

Notes. This figure displays the percent of participants in each experiment playing their optimal strategy in each round. Choosers and Advisors are plotted in separate sub-figures.

Figure 10 compares the frequency of wary equilibrium strategy play across the two experiments. Figure 10a shows that Choosers' strategies do not change much across the two experiments. In almost all rounds, the percent of Choosers playing accept-reject in the beliefs experiment is not significantly different from the main experiment (proportions tests, $p > 0.13$).[12]

Despite the lack of difference in Choosers' strategies, Figure 10b shows that Advisors are much more likely to send optimal recommendations in the beliefs experiment. The aggregate trend in Advisors' strategies is the same in both experiments: the percent of Advisors playing partial truth-telling increases as rounds progress. However, the percent doing so in the beliefs experiment is consistently larger than in the main experiment, and significantly so in all rounds after and including round 4 (proportions tests, $p < 0.05$). Even in round 1 - when Advisors have not received any feedback about Choosers' strategies - the percent of Advisors playing partial truth-telling is 57% larger in the beliefs experiment. As a result of this change in Advisors' strategies, Choosers and Advisors play their wary equilibrium strategies at more similar rates in the beliefs experiment. Averaging across all rounds, the percent difference in the proportion of Advisors and Choosers playing optimally is significantly smaller in the beliefs experiment

[12]Proportions tests indicate that significantly more Choosers in the beliefs experiment play accept-reject in rounds 4 ($p < 0.01$), 8 ($p < 0.05$), 7 and 14 ($p < 0.1$ for both). This means that, starting in round 5, Advisors in the beliefs experiment may have been matched with more Choosers who played accept-reject than were Advisors in the main experiment. However, beliefs experiment Advisors are more likely to play partial truth-telling starting in round 1. Given that the change in Advisors' strategies begins prior to the change in Choosers' strategies, it is unlikely that it was driven by an increased number of pairings with Choosers playing accept-reject.

(Wilcoxon rank-sum test, $p < 0.01$).

RESULT 7: *Advisors are much more likely to play partial truth-telling in the beliefs experiment than in the main experiment.*

Rounds	All (1)	1-10 (2)	11-20 (3)
Round	0.013^{***}	0.033^{***}	0.005
	(0.003)	(0.008)	(0.003)
Constant	0.258^{*}	0.201	0.338^{**}
	(0.150)	(0.148)	(0.161)
Observations	1300	650	650
Mean	0.578	0.522	0.634

Notes. Coefficients from OLS regressions conducted among Advisors in the beliefs experiment, with standard errors in parentheses. The dependent variable is an indicator for playing partial truth-telling (Advisors' wary equilibrium strategy) in a given round, and the independent variable is the round number. All regressions include controls for the session and standard errors clustered at the participant-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 2: Evolution of Advisors' likelihood of playing optimal strategy (beliefs experiment)

In addition to being more likely to send optimal recommendations, Advisors learn to send these recommendations more quickly. Table 2 reports the results of regressions that explore the evolution of Advisors' strategies in the beliefs experiment. Across all rounds (column 1), playing an additional round increases an Advisor's likelihood of playing partial truth-telling by 2.3% in the beliefs experiment. While this aggregate effect is essentially identical to that in the main experiment (see column 1 of Table 1), columns 2 and 3 reveal that Advisors learn their optimal strategy more quickly in the beliefs experiment. In early rounds (column 2), the effect of playing an additional round in the beliefs experiment is 6.3% (3.3 p.p.) - 50% larger than the early-round effect in the main experiment (see column 2 of Table 1). In late rounds (column 3), Advisors' likelihood of playing partial truth-telling does not change significantly in the beliefs experiment. In the main experiment, however, Advisors' strategies still evolve in late rounds, suggesting that they are still learning which strategy is optimal.

RESULT 8: *Advisors learn their optimal strategy more quickly in the beliefs experiment than in the main experiment.*

Results 7 and 8 support the claim that Advisors' failure to send optimal recommendations is driven by a failure of contingent thinking. Next, I provide an additional test of this mechanism. Recall that, in every round, some Advisors hold incorrect beliefs about the modal Chooser strategy. In other words, eliciting Advisors' beliefs does not prompt all of them to focus on the relevant contingency. I therefore next examine whether Advisors who hold accurate beliefs are more likely to send optimal recommendations.

Rounds	All	Rounds 1 & 7	Rounds 13 & 19
	(1)	(2)	(3)
Believes optimal is mode	0.446^{***}	0.377^{***}	0.541^{***}
	(0.075)	(0.096)	(0.096)
Constant	0.017	0.017	-0.068
	(0.881)	(0.897)	(0.666)
Observations	260	130	130
Mean	0.546	0.469	0.623

Notes. Coefficients from OLS regressions conducted among Advisors in the beliefs experiment, with standard errors in parentheses. The dependent variable is an indicator for playing partial truth-telling (Advisors' wary equilibrium strategy) in a given round, and the independent variable is an indicator for believing accept-reject (Choosers' wary equilibrium strategy) is the modal Chooser strategy in a given round. All regressions include controls for the round and session. Standard errors clustered at the participant-level. * denotes $p < 0.1$, ** denotes $p < 0.05$, and *** denotes $p < 0.01$.

Table 3: Effect of beliefs on Advisors' likelihood of playing optimal strategy

Table 3 reports the results of regressions that identify the impact of Advisors' beliefs about the modal Chooser strategy in a given round on the strategy they play in that round. These results confirm that Advisors who focus on the relevant contingency - that is, who believe Choosers play accept-reject - are more likely to send optimal recommendations. As shown in column 1, Advisors who hold correct beliefs about the modal Chooser strategy in the upcoming round are 81.7% (44.6 p.p.) more likely to send optimal recommendations in that round. Columns 2 and 3 indicate that the effect is similar in early and late rounds.

RESULT 9: *Advisors who hold correct beliefs about the modal Chooser strategy, and thus focus on the relevant contingency, are more likely to send optimal recommendations.*

In summary, despite little-to-no change in Choosers' strategies across experiments, Advisors in the beliefs experiment are much more likely to send optimal recommendations. Results 6 and 9 point to the reason for this discrepancy: the belief elicitation task prompts most Advisors to focus on Choosers' modal strategy, and those who do are much more likely to send optimal recommendations. Thus, in both experiments, directly or indirectly placing Advisors in the relevant Chooser strategy contingency makes them less likely to send "naïve" recommendations and more likely to send optimal ones. This behavior is consistent with a failure of contingent thinking.

5 Discussion and Conclusion

When attempting to persuade Choosers to change their behavior, paternalistic Advisors should account for how Choosers may respond to the recommended change. In a setting where Choosers are wary of recommenders' incentives, I show that many Advisors fail to account for this wariness. Despite repeated interactions with Choosers and feedback about their responses to recommendations, most Advisors send recommendations that would only be optimal if Choosers were required to follow them. Interventions that prompt Advisors to put themselves "in the Chooser's shoes" help correct this mistake.

In reality, the settings in which Advisors issue recommendations may be even more complex than in my experiment. First, Choosers may consider alternatives other than their status-quo and the recommended alternative. Second, non-paternalistic Advisors may also be strategic, which would require paternalistic Advisors to account for *their* strategies as well. Finally, states and their corresponding actions may not be equidistant from each other. For instance, a bank may offer a basic credit card with no annual fees, and two premium credit cards with high yet similar fees. The optimality of the paternalistic Advisor's partial truth-telling strategy is robust to these variations of the recommendation game, and it would conceivably be even more challenging for Advisors to send optimal recommendations in these more complex environments. Determining exactly how much more challenging may be a promising avenue for future research.

That said, the optimality of a partial truth-telling strategy requires certain restrictions on the Chooser's beliefs and action set. In order for the recommendation game to have a wary equilibrium, the Chooser must be able to take an action that is in-between their status-quo and the action they are wary of taking. Put differently, a partial truth-telling strategy can only be optimal situations where the Chooser benefits from making a large behavioral change. Such situations are very common empirically. For instance, whereas the Canadian government recommends drinking no more than two alcoholic beverages per week (Paradis et al., 2023), the average Canadian adult drinks nearly five times that amount (Statistics Canada, 2024).[13] The optimality of a partial truth-telling strategy is also limited to situations where the Chooser believes that a large behavioral change is unlikely to be in their best interest, as modelled by the

[13]Similarly, whereas the WHO recommends that adults perform at least 150 minutes of moderate exercise per week (World Health Organization, 2020), up to one-third of adults worldwide do not exercise at all (Ipsos Global Advisor, 2021). Finally, despite the importance of saving for retirement, 36% of unretired Canadians have no retirement savings (Abacus Data, 2024). In all of these scenarios, it is thus possible to recommend an action that is in-between many Choosers' status-quo and the action that is in their best interest.

assumption $q_1 \geq q_3$. This assumption captures a common form of motivated reasoning, where a Chooser convinces themselves that they don't need to change their behavior. For instance, Choosers often persuade themselves that they do not need to drink less alcohol, exercise more or save more, even when presented evidence that such changes are in their best interest.

In my model and experiment, Choosers' mistrust in recommenders' incentives is the only barrier that may prevent them from changing their behavior. However, other such barriers exist. For instance, Advisors may need to account for the difficulties Choosers may face in breaking pre-existing habits, which impact many important consumption decisions (Charness and Gneezy, 2009; Atkin, 2013; Acland and Levy, 2015; Yakovlev, 2018; Harris and Kessler, 2019). In future work, it may thus be valuable assess Advisors' ability to account for other or multiple barriers to behavioral change.

References

Abacus Data (2024). Canadian Retirement Study.

Abeler, J., D. Nosenzo, and C. Raymond (2019). Preferences for Truth-Telling. *Econometrica 87*(4), 1115–1153.

Acland, D. and M. R. Levy (2015). Naiveté, Projection Bias, and Habit Formation in Gym Attendance. *Management Science 61*(1), 146–160.

Ali, S. N., M. Mihm, L. Siga, and C. Tergiman (2021). Adverse and Advantageous Selection in the Laboratory. *American Economic Review 111*(7), 2152–78.

Alsan, M. and S. Eichmeyer (2024). Experimental Evidence on the Effectiveness of Nonexperts for Improving Vaccine Demand. *American Economic Journal: Economic Policy 16*(1), 394–414.

Altmann, S., A. Falk, and A. Grunewald (2022). Communicating through Defaults. *Review of Economics and Statistics*, forthcoming.

Ambuehl, S., B. D. Bernheim, and A. Ockenfels (2021). What Motivates Paternalism? An Experimental Study. *American Economic Review 111*(3), 787–830.

Anagol, S., S. Cole, and S. Sarkar (2017). Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market. *Review of Economics and Statistics 99*(1), 1–15.

Atkin, D. (2013). Trade, Tastes, and Nutrition in India. *American Economic Review 103*(5), 1629–1663.

Bartling, B., A. W. Cappelen, H. Hermes, M. Skivenes, and B. Tungodden (2023). Free to Fail? Paternalistic Preferences in the United States. *DICE Discussion Paper No. 400.*

Bartling, B., E. Fehr, and H. Herz (2014). The Intrinsic Value of Decision Rights. *Econometrica 82*(6), 2005–2039.

Blume, A., E. K. Lai, and W. Lim (2020). Strategic Information Transmission: A Survey of Experiments and Theoretical Foundations. In C. M. Capra, R. T. Croson, M. L. Rigdon, and T. S. Rosenblat (Eds.), *Handbook of Experimental Game Theory*, pp. 311–347. Edward Elgar Publishing.

Cai, H. and J. T.-Y. Wang (2006). Overcommunication in Strategic Information Transmission Games. *Games and Economic Behavior 56*(1), 7–36.

Calford, E. M. and T. N. Cason (2024). Contingent Reasoning and Dynamic Public Goods Provision. *American Economic Journal: Microeconomics 16*(2), 236–266.

Camerer, C. F., T.-H. Ho, and J.-K. Chong (2004). A Cognitive Hierarchy Model of Games. *The Quarterly Journal of Economics 119*(3), 861–898.

Charness, G. and U. Gneezy (2009). Incentives to Exercise. *Econometrica 77*(3), 909–931.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery (2013). Barriers to Household Risk Management: Evidence from India. *American Economic Journal: Applied Economics 5*(1), 104–135.

Crawford, V. P. and J. Sobel (1982). Strategic Information Transmission. *Econometrica 50*(6), 1431–1451.

Croson, R. T. (1999). The Disjunction Effect and Reason-Based Choice in Games. *Organizational Behavior and Human Decision Processes 80*(2), 118–133.

Croson, R. T. (2000). Thinking Like a Game Theorist: Factors Affecting the Frequency of Equilibrium Play. *Journal of Economic Behavior & Organization 41*(3), 299–314.

Dal Bó, P. and G. R. Fréchette (2018). On the Determinants of Cooperation in Infinitely Repeated Games: A Survey. *Journal of Economic Literature 56*(1), 60–114.

Danz, D., L. Vesterlund, and A. J. Wilson (2022). Belief Elicitation and Behavioral Incentive Compatibility. *American Economic Review 112*(9), 2851–2883.

DellaVigna, S. and E. Linos (2022). RCTs to Scale: Comprehensive Evidence from Two Nudge Units. *Econometrica 90*(1), 81–116.

Epley, N. and T. Gilovich (2016). The Mechanics of Motivated Reasoning. *Journal of Economic Perspectives 30*(3), 133–140.

Esponda, I. and E. Vespa (2014). Hypothetical Thinking and Information Extraction in the Laboratory. *American Economic Journal: Microeconomics 6*(4), 180–202.

Esponda, I. and E. Vespa (2023). Contingent Thinking and the Sure-Thing Principle: Revisiting Classic Anomalies in the Laboratory. *Review of Economic Studies*, forthcoming.

Fehr, E., H. Herz, and T. Wilkening (2013). The Lure of Authority: Motivation and Incentive Effects of Power. *American Economic Review 103*(4), 1325–1359.

Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association 1*(1), 114–125.

Gruber, J. and M. Owings (1994). Physician Financial Incentives and Cesarean Section Delivery. *NBER Working Paper #4933*.

Gul, F. and W. Pesendorfer (2001). Temptation and Self-Control. *Econometrica 69*(6), 1403–1435.

Harris, M. C. and L. M. Kessler (2019). Habit Formation and Activity Persistence: Evidence from Gym Equipment. *Journal of Economic Behavior & Organization 166*, 688–708.

Ipsos Global Advisor (2021). Global views on exercise and team sports.

Johnson, E. (2017, March). 'I will do anything I can to make my goal': TD teller says customers pay price for 'unrealistic' sales targets. *CBC News*. Available at: https://www.cbc.ca/news/canada/british-columbia/td-tellers-desperate-to-meet-increasing-sales-goals-1.4006743.

Johnson, E., J. McDonald, M. McNair, K. Ivany, and M. McCann (2024, March). Hidden cameras capture bank employees misleading customers, pushing products that help sales targets. *CBC News*. Available at: https://www.cbc.ca/news/business/marketplace-hidden-camera-banks-1.7142427.

Kagel, J. H. and D. Levin (1986). The Winner's Curse and Public Information in Common Value Auctions. *American Economic Review 76*(5), 894–920.

Kawagoe, T. and H. Takizawa (2009). Equilibrium Refinement vs. Level-k Analysis: An Experimental Study of Cheap-Talk Games with Private Information. *Games and Economic Behavior 66*(1), 238–255.

Kunda, Z. (1990). The Case for Motivated Reasoning. *Psychological Bulletin 108*(3), 480.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *The Quarterly Journal of Economics 112*(2), 443–478.

Ngangoué, M. K. and G. Weizsäcker (2021). Learning from Unrealized versus Realized Prices. *American Economic Journal: Microeconomics 13*(2), 174–201.

Niederle, M. and E. Vespa (2023). Cognitive Limitations: Failures of Contingent Thinking. *Annual Review of Economics 15*, 307–328.

OECD (2017). Behavioural Insights and Public Policy: Lessons from around the World. *OECD Publishing*.

Paradis, C., P. Butt, K. Shield, N. Poole, S. Wells, T. Naimi, A. Sherk, and the Low-Risk Alcohol Drinking Guidelines Scientific Expert Panels (2023). Canada's Guidance on Alcohol and Health: Final Report. Ottawa: Canadian Centre on Substance Use and Addiction.

Romero, J. and Y. Rosokha (2023). Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma. *Econometrica 91*(6), 2295–2331.

Samuelson, W. and R. Zeckhauser (1988). Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty 1*, 7–59.

Statistics Canada (2024). Control and sale of alcoholic beverages and cannabis, April 1, 2022, to March 31, 2023.

Thaler, R. H. (1988). Anomalies: The Winner's Curse. *Journal of Economic Perspectives 2*(1), 191–202.

Thaler, R. H. and C. R. Sunstein (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Penguin Books.

Wang, J. T.-Y., M. Spezio, and C. F. Camerer (2010). Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games. *American Economic Review 100*(3), 984–1007.

World Health Organization (2020). Who guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020. Licence: CC BY-NC-SA 3.0 IGO.

Yakovlev, E. (2018). Demand for Alcohol Consumption in Russia and its Implication for Mortality. *American Economic Journal: Applied Economics 10*(1), 106–149.

Appendix

A Proofs

In all proofs, I use $t \in \{Pat, Com\}$ to denote the Advisor's type, where Pat indicates the paternalistic type and Com indicates the commitment type. I use $\mu(\theta, t|r)$ to denote the Chooser's belief about the probability of the history (θ, t) having occurred, conditional on receiving recommendation r.

A.1 Proofs of Propositions

Note that if the Chooser is recommended to take action 1 (i.e., to stick with their status-quo), their action set is the singleton set $\{1\}$. It is therefore without loss to restrict one's attention to the Chooser's responses to recommendations of 2 and 3, which is the approach I take in all proofs in Section A.1. Furthermore, given that the commitment Advisor is non-strategic, the Chooser and paternalistic Advisor are the only players in the recommendation game. When describing each equilibrium, I therefore include players' strategies, as well as a reminder that the commitment Advisor plays $r(\theta) = 3$ for all θ.

Proof of Proposition 1

Suppose the Chooser plays $a_C(r) = r$ for all $r \in \{2, 3\}$; that is, they accept all recommendations to take an action other than their status-quo. The paternalistic Advisor's unique best response is $r(\theta) = \hat{a}(\theta)$ for all θ, given that they want the Chooser's action to be as close as possible to $\hat{a}(\theta)$. If the paternalistic Advisor uses this strategy, the Chooser forms the following beliefs:

$$\mu(1, Com|3) = \frac{q_1(1-p)}{q_1(1-p) + q_2(1-p) + q_3} \tag{1}$$

$$\mu(2, Com|3) = \frac{q_2(1-p)}{q_1(1-p) + q_2(1-p) + q_3} \tag{2}$$

$$\mu(3, Com|3) = \frac{q_3(1-p)}{q_1(1-p) + q_2(1-p) + q_3} \tag{3}$$

$$\mu(3, Pat|3) = \frac{q_3 p}{q_1(1-p) + q_2(1-p) + q_3} \tag{4}$$

$$\mu(1, Pat|1) = 1 \tag{5}$$

$$\mu(2, Pat|2) = 1 \tag{6}$$

$$\mu(\theta, Pat|r) = 0 \text{ for all } r \neq \theta \tag{7}$$

Given these beliefs, it is always better for the Chooser to accept a recommendation of 2 than to stick with their status-quo. Accepting a recommendation of 3 is better than sticking with their status-quo if

$$\frac{1}{q_1(1-p)+q_2(1-p)+q_3}[q_1(1-p)u(2\alpha)+q_2(1-p)u(\alpha)+q_3u(0)] \geq$$
$$\frac{1}{q_1(1-p)+q_2(1-p)+q_3}[q_1(1-p)u(0)+q_2(1-p)u(\alpha)+q_3u(2\alpha)]$$

which simplifies to

$$q_3 \geq q_1(1-p) \tag{8}$$

Thus, if Equation 8 holds, there exists a "trusting" weak perfect Bayesian equilibrium where the paternalistic Advisor plays $r(\theta) = \hat{a}(\theta)$ for all θ, the Chooser plays $a_C(r) = r$ for all r, and the Chooser holds the beliefs described in Equations 1 through 7. Additionally, by assumption, the commitment Advisor recommends $r(\theta) = 3$ for all θ. ■

Proof of Proposition 2

Suppose the Chooser plays $a_C(2) = 2$ and $a_C(3) = 1$; that is, they accept a recommendation of 2 and reject a recommendation of 3. The paternalistic Advisor's unique best response is $r(1) = 1$ and $r(2) = r(3) = 2$, given that they want the Chooser's action to be as close as possible to $\hat{a}(\theta)$. The uniqueness of the best response for $\theta = 1$ arises from the assumption that the paternalistic Advisor recommends the action closest to $\hat{a}(\theta)$ when indifferent between recommendations. If the paternalistic Advisor uses this strategy, the Chooser forms the following beliefs:

$$\mu(\theta, Com|3) = q_\theta \text{ for all } \theta \tag{9}$$
$$\mu(\theta, Pat|3) = 0 \text{ for all } \theta \tag{10}$$
$$\mu(1, Pat|2) = 0 \tag{11}$$
$$\mu(2, Pat|2) = \frac{q_2}{q_2+q_3} \tag{12}$$
$$\mu(3, Pat|2) = \frac{q_3}{q_2+q_3} \tag{13}$$
$$\mu(1, Pat|1) = 1 \tag{14}$$
$$\mu(\theta, Pat|1) = 0 \text{ for all } \theta \neq 1 \tag{15}$$

According to these beliefs, accepting a recommendation of 2 will always bring the Chooser closer to the action corresponding to the true state than will rejecting it (i.e., sticking with their status-quo). Thus, it is always best for the Chooser to accept a recommendation of 2. Rejecting a recommendation of 3 is better than accepting it if

$$q_1 u(2\alpha) + q_2 u(\alpha) + q_3 u(0) \leq q_1 u(0) + q_2 u(\alpha) + q_3 u(2\alpha)$$

which simplifies to

$$q_3 \leq q_1 \tag{16}$$

Equation 16 holds by assumption. Thus, for all parameter values, there exists a "wary" weak perfect Bayesian equilibrium where the paternalistic Advisor plays $r(3) = 2$ and $r(\theta) = \hat{a}(\theta)$ for all $\theta \neq 3$, the Chooser plays $a_C(2) = 2$ and $a_C(3) = 1$, and the Chooser holds the beliefs described in Equations 9 through 15. Additionally, by assumption, the commitment Advisor recommends $r(\theta) = 3$ for all θ. ■

Proof of Proposition 3

In Propositions 1 and 2, I showed that the paternalistic Advisor has a unique best response to each of the Chooser's strategies that involve accepting a recommendation of 2.[14] To show that the trusting and wary equilibria are the unique equilibria of the recommendation game, I therefore prove that there can be no equilibrium where the Chooser rejects a recommendation of 2.

Suppose the Chooser plays a strategy that involves rejecting a recommendation of 2. The paternalistic Advisor's unique best response is $r(\theta) = \hat{a}(\theta)$ for all θ, given that they want the Chooser's action to be as close as possible to $\hat{a}(\theta)$ and recommend the action closest to $\hat{a}(\theta)$ when indifferent[15] between recommendations. If the paternalistic Advisor uses this strategy, the Chooser would hold the belief $\mu(2, Pat|2) = 1$, $\mu(\theta, Pat|2) = 0$ for all $\theta \neq 2$. Stated otherwise, receiving a recommendation of 2 indicates that the state is 2. The Chooser could thus profitably deviate to accepting a recommendation of 2. Thus, there is no weak perfect Bayesian equilibrium where the Chooser rejects a recommendation of 2. As a result, the trusting and wary equilibria are the unique equilibria of the recommendation game. ■

[14] Recall that the Chooser has two strategies that involve accepting a recommendation of 2: accept all (i.e., accept 2, accept 3) and accept-reject (i.e., accept 2, reject 3).

[15] When the Chooser rejects a recommendation of 2, in state 2, the paternalistic Advisor is indifferent between all possible recommendations. When the Chooser rejects all recommendations, the paternalistic Advisor is additionally indifferent between all possible recommendations in states 1 and 3.

A.2 Proofs of Model's Robustness to Alternative Assumptions

In the model presented in Section 2, I impose a restriction on the Chooser's action space, assume states are equidistant and assume that the commitment Advisor is non-strategic. With these assumptions in place, I find that the recommendation game has an equilibrium where the paternalistic Advisor plays a partial truth-telling strategy. In this section, I show that the optimality of the partial truth-telling strategy does not depend on these assumptions.

Chooser can take any action

Modification to main model: The Chooser's action space becomes $a_c(r) \in \{1,2,3\}$ instead of $a_c(r) \in \{1,r\}$.

Proof: Suppose the Chooser plays the following strategy, which is similar to their accept-reject strategy in the original model: $a_c(1) = 1$, $a_c(2) = 2$ and $a_c(3) = 1$. The paternalistic Advisor's unique best response is their partial truth-telling strategy; that is, $r(1) = 1$ and $r(2) = r(3) = 2$. The Chooser thus forms the beliefs outlined in Equations 9 through 15.

I next find the conditions under which it is optimal for the Chooser to play the strategy specified above, given their beliefs. Since the Chooser knows the state is 1 if they are recommended 1, it is always optimal for them to play $a_c(1) = 1$. If the Chooser is recommended 2, they know the state is either 2 or 3. $a_c(2) = 2$ is optimal if $q_2 \geq q_3$. Finally, a recommendation of 3 reveals no new information about the true state. In order for $a_c(3) = 1$ to be optimal, it must yield a greater expected payoff than $a_c(3) = 3$ and $a_c(3) = 2$, which is true under the following conditions:

$$(1-p)[q_1u(0)+q_2u(\alpha)+q_3u(2\alpha)] \geq (1-p)[q_1u(2\alpha)+q_2u(\alpha)+q_3u(0)] \tag{17}$$
$$(1-p)[q_1u(0)+q_2u(\alpha)+q_3u(2\alpha)] \geq (1-p)[q_1u(\alpha)+q_2u(0)+q_3u(\alpha)] \tag{18}$$

Equation 17 simplifies to $q_1 \geq q_3$, and Equation 18 simplifies to the stronger condition $q_1 \geq q_2 + q_3$. Thus, when the Chooser can take any action, the paternalistic Advisor can play a partial truth-telling strategy in equilibrium if $q_2 \geq q_3$ and $q_1 \geq q_2 + q_3$. ■

Non-equidistant states

Modification to main model: Instead of there being distance 1 between adjacent states, there is distance d_{12} between states 1 and 2 and d_{23} between states 2 and 3, where $d_{12},\ d_{23} \geq 0$.

Proof: Suppose the Chooser plays their accept-reject strategy; that is, $a_c(2) = 2$ and $a_c(3) = 1$. The paternalistic Advisor's unique best response is their partial truth-telling strategy; that is,

$r(1) = 1$ and $r(2) = r(3) = 2$. The Chooser thus forms the beliefs outlined in Equations 9 through 15.

Given these beliefs, it is always optimal for the Chooser to accept a recommendation of 2, since doing so will always bring them closer to the action corresponding to the true state than will rejecting it. It is optimal for the Chooser to reject a recommendation of 3 if

$$(1-p)[q_1 u(0) + q_2 u(\alpha d_{12}) + q_3 u(\alpha(d_{12}+d_{23}))] \geq$$
$$(1-p)[q_1 u(\alpha(d_{12}+d_{23})) + q_2 u(\alpha d_{23}) + q_3 u(0)]$$

which simplifies to

$$q_1 \geq q_2 \frac{u(\alpha d_{23}) - u(\alpha d_{12})}{u(0) - u(\alpha(d_{12}+d_{23}))} + q_3 \tag{19}$$

When $d_{23} \geq d_{12}$, $u(\alpha d_{23}) \leq u(\alpha d_{12})$ (recall that $u'(\cdot) < 0$). For those relative distances, the left-hand side of Equation 19 is thus strictly smaller than q_3, which means Equation 19 is always satisfied (recall that $q_1 \geq q_3$ by assumption). Thus, when $d_{23} \geq d_{12}$, the recommendation game always has an equilibrium where the Advisor plays a partial truth-telling strategy. When $d_{23} > d_{12}$, the recommendation game has such an equilibrium if Equation 19 is satisfied. ■

Strategic self-interested Advisor

Modification to main model: The non-strategic commitment Advisor is replaced by a strategic, self-interested player. This player either (1) receives a bonus if the Chooser takes action 3, or (2) places some weight on each of two outcomes: maximizing the Chooser's wellbeing and recommending action 3. More precisely, the self-interested Advisor's preferences are described by one of the following payoff functions:

$$U_1(a_c(r)) = \begin{cases} \pi & \text{if } a_c(r) \neq 3 \\ \pi + B & \text{if } a_c(r) = 3, \text{ where } B > 0 \end{cases}$$
$$U_2(a_c(r)) = \gamma \cdot v(\alpha |a_c(r) - \hat{a}(\theta)|) + (1-\gamma) \cdot B \cdot \mathbb{1}\{r = 3\}, \text{ where } B > 0 \text{ and } \gamma \in [0,1]$$

$U_1(\cdot)$ captures the preferences of an Advisor who stands to benefit significantly more when the Chooser takes one particular action. Physicians, for instance, may receive a much larger compensation from performing a particular procedure. $U_2(\cdot)$ captures the preferences of an Advisor who is pressured to recommend certain products, which is common in the financial industry.[16]

[16]For instance, a major Canadian news outlet has documented how financial advisors at large Canadian banks

Proof: I wish to show that, for each of the self-interested Advisor's possible payoff functions, the recommendation game can have an equilibrium where the Chooser plays accept-reject, the paternalistic Advisor plays their partial truth-telling strategy and the self-interested Advisor always recommends action 3. If the Advisors use these strategies, the Chooser forms the beliefs outlined in Equations 9 through 15. As shown in the proof of Proposition 2, given these beliefs, it is optimal for the Chooser to accept a recommendation of 2 and reject a recommendation of 3.

I next check that each Advisor's strategy is a best response to the Chooser's strategy. As shown in the proof of Proposition 2, the paternalistic Advisor's unique best response to the Chooser's accept-reject strategy is their partial truth-telling strategy. When the self-interested Advisor has the preferences described by $U_1(\cdot)$, they are indifferent between the Chooser accepting a recommendation of 2 and rejecting any given recommendation. Thus, given the Chooser's strategy, a self-interested Advisor with these preferences cannot profitably deviate from recommending 3 in all states. As a result, when the self-interested Advisor has the preferences described by $U_1(\cdot)$, the recommendation game always has an equilibrium where the paternalistic Advisor plays a partial truth-telling strategy.

When the self-interested Advisor has the preferences described by $U_2(\cdot)$, they may have a profitable deviation if recommending an action other than 3 leads the Chooser to take an action that is closer to the true state. This can only be the case in state 2, where recommending 2 leads the Chooser to take the action corresponding to the true state. The self-interested Advisor prefers recommending 3 in state 2 if

$$\gamma \cdot v(\alpha) + (1-\gamma)B \geq \gamma \cdot v(0)$$

which simplifies to

$$\frac{B}{u(0) - u(\alpha) + B} \geq \gamma \tag{20}$$

Intuitively, the self-interested Advisor must place a sufficiently high weight on their bonus (relative to the Chooser's wellbeing). Thus, when the self-interested Advisor has the preferences described by $U_2(\cdot)$, the recommendation game has an equilibrium where the paternalistic Advisor plays a partial truth-telling strategy if Equation 20 is satisfied. ■

are pressured - and sometimes even "coached" - to recommend certain financial products to their clients (Johnson, 2017; Johnson et al., 2024).

B Experiment instructions

Below is a transcription of the instructions and tasks in both experiments. Any instructions or tasks that are specific to a particular role (Chooser or Advisor) or a particular experiment (Main or Beliefs) are indicated by text boxes.

Welcome to the experiment!

General instructions. This is an experiment designed to study decision-making. If you pay close attention to the instructions, you can earn a significant amount of money. Please ensure all of your electronics (cell phones, smart watches, etc.) are put away, and do not talk with others during the experiment. If you have a question, please raise your hand and an experimenter will come answer it in private.

MAIN EXPERIMENT ONLY

> ***Structure of the experiment.*** The experiment consists of multiple rounds of a game, followed by a short survey. You will receive detailed instructions prior to each part of the experiment.

BELIEFS EXPERIMENT ONLY

> ***Structure of the experiment.*** The experiment consists of multiple rounds of a game. Before several of those rounds, you will complete a guessing task. Finally, you will complete a short survey. You will receive detailed instructions prior to each part of the experiment.

Payment. Your total earnings from the experiment will consist of several components.

- *Show-up fee.* You will receive a $5.00 payment for completing the experiment.
- *Quiz.* After reading the experiment's instructions but before proceeding to the main task, you will complete a quiz that will test your understanding of the instructions. You will receive $1.00 if you answer all quiz questions correctly on your first try.

MAIN EXPERIMENT ONLY

> - *Randomly-selected round of the game.* One round of the game has been randomly selected to count for your payment.

BELIEFS EXPERIMENT ONLY

> - *Randomly-selected round and guessing task.* One round of the game, as well as one guessing task, have each been randomly selected to count for your payment.

The Plant Pot Game - Instructions

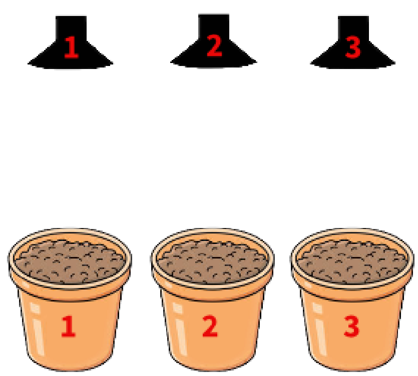
There are three pots, each with a sprinkler above it. The “plant pot” has a plant seed in it; the other two do not.

CHOOSERS ONLY

You are playing the role of **Chooser**. You will not know for sure which pot is the plant pot, but you will know each pot’s chance of being the plant pot. Specifically, pot 1 has a 9-in-20 chance, pot 2 has a 7-in-20 chance and pot 3 has a 4-in-20 chance.

ADVISORS ONLY

Participants playing the role of Chooser know that pot 1 has a 9-in-20 chance of being the plant pot, pot 2 has a 7-in-20 chance and pot 3 has a 4-in-20 chance. You, however, are playing the role of **Advisor**. As we’ll explain later, Advisors have more information about the plant pot than Choosers do.



9-in-20 7-in-20 4-in-20

CHOOSERS ONLY

The pots, the sprinklers and each pot’s chance of being the plant pot.

ADVISORS ONLY

The pots, the sprinklers and the information Choosers have (each pot’s chance of being the plant pot).

When a sprinkler is turned on, it sprays water directly below it, as well as below to the right and left of it. The pot closest to it receives the most water; the pot farthest from it receives the least. A seed will grow into a larger plant if it receives more water.

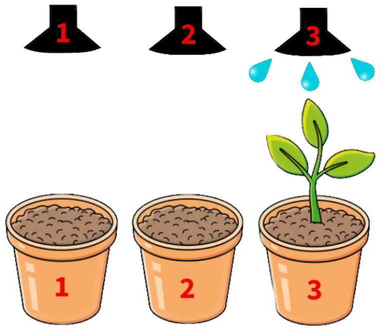


Figure 1

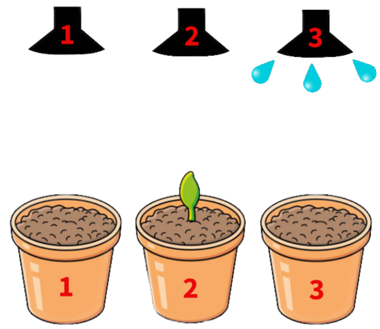


Figure 2

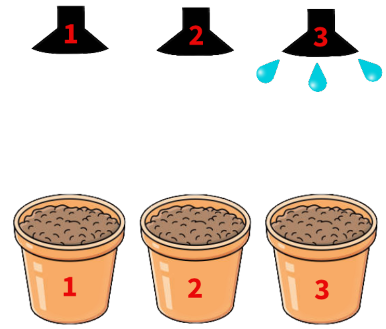


Figure 3

Example: Turning on sprinkler 3 will produce a large plant if pot 3 is the plant pot (as in Figure 1), a small plant if pot 2 is the plant pot (as in Figure 2), and no plant if pot 1 is the plant pot (as in Figure 3).

Your goal

CHOOSERS ONLY

As a Chooser, your goal is to turn on the sprinkler that's as close as possible to the plant pot, and thus to produce the largest possible plant.

ADVISORS ONLY

As an Advisor, you will be paired with a participant playing the role of Chooser. Choosers and Advisors have the same goal: for the Chooser to turn on the sprinkler that's as close as possible to the plant pot, and thus to produce the largest possible plant.

Your task

CHOOSERS ONLY

While sprinkler 1 will be turned on by default, you may receive a recommendation to switch to one of the other sprinklers. Your task requires you to consider two scenarios: the scenario where you receive a recommendation to switch to sprinkler 2, and the one where you receive a recommendation to switch to sprinkler 3. You must decide whether you will accept the proposed switch or stick with the default in each of these scenarios. If you are recommended to stick with the default, you will automatically do so.

ADVISORS ONLY

Your task requires you to consider three scenarios: the scenario where pot 1 is the plant pot, the one where pot 2 is the plant pot, and the one where pot 3 is the plant pot. You must choose a sprinkler to recommend to the Chooser in each scenario. Think of this task as if you know a specific pot (pot 1, 2 or 3) is the plant pot, and with this knowledge, must recommend a sprinkler to the Chooser.

Sprinkler 1 will be turned on by default; you may recommend that the Chooser either stick with this default or switch to one of the other sprinklers. While you are completing this task, the Chooser will decide whether, if recommended to switch to a given sprinkler, they will accept the proposed switch or stick with the default. If they are recommended to stick with the default, they will automatically do so.

Sources of recommendations

CHOOSERS ONLY

The recommendation you receive will be from one of two sources: an Advisor or the computer. An Advisor is another participant who has the same goal as you, and can recommend sticking with the default or switching to one of the other sprinklers. While you are completing your task, they will choose a sprinkler to recommend for each possible plant pot. You can think of the Advisor's task as if they know a specific pot (pot 1, 2 or 3) is the plant pot, and with this knowledge, must recommend you a sprinkler. The computer recommends to switch to sprinkler 3 regardless of which pot is the plant pot. There is a 1-in-3 chance that you will receive an Advisor's recommendation, and a 2-in-3 chance that you will receive the computer's recommendation.

ADVISORS ONLY

The recommendation the Chooser receives will be from one of two sources: you or the computer. The computer recommends to switch to sprinkler 3 regardless of which pot is the plant pot. There is a 1-in-3 chance that the Chooser will receive your recommendation, and a 2-in-3 chance that they will receive the computer's recommendation. When making their decisions, Choosers know their chance of receiving a recommendation from each potential source. They also know that Advisors have the same goal as them, and must choose a sprinkler to recommend for each possible plant pot.

Rounds

CHOOSERS ONLY

You will play 22 rounds of this game. At the beginning of each round, we will randomly re-match each Chooser with an Advisor. We will also determine the plant pot and source of the recommendation you will receive according to the chances mentioned earlier. After you and the Advisor you are paired with have made your decisions, the outcome of the game will be determined. At the end of each round, you will receive feedback about that outcome, as well as the decisions of the Advisor you were paired with.

ADVISORS ONLY

You will play 22 rounds of this game. At the beginning of each round, we will randomly re-match each Chooser with an Advisor. We will also determine the plant pot and source of the recommendation the Chooser will receive according to the chances mentioned earlier. After you and the Chooser you are paired with have made your decisions, the outcome of the game will be determined. At the end of each round, you will receive feedback about that outcome, as well as the decisions of the Chooser you were paired with.

Round 1 is a “practice round”. It is intended to familiarize you with your task, and is not eligible to determine your payment from the experiment. At the end of round 1, you will see an outline of the feedback you will receive at the end of the rest of the rounds.

BELIEFS EXPERIMENT ONLY

Guessing tasks

CHOOSERS ONLY

Before rounds 2, 8, 14 and 20, you will (1) make guesses about the most-common decisions Advisors have made in that round in previous sessions of the experiment, and (2) state which recommendation the computer sends when any given pot is the plant pot. Part (2) does not involve guessing; its purpose is to ensure you remember which recommendation the computer always sends.

ADVISORS ONLY

Before rounds 2, 8, 14 and 20, you will make guesses about the most-common decisions Choosers have made in that round in previous sessions of the experiment.

Your payment

CHOOSERS ONLY

One round of the game has been randomly selected to count for your payment. You will receive \$16.00 if you produce a large plant in that round, \$13.00 if you produce a small plant, and \$10.00 if you produce no plant.

Beliefs experiment only

One of the guessing tasks has also been randomly selected to count for your payment. You will receive \$5.00 if your guesses about Advisors' recommendations, as well as your statements about the recommendation the computer sends, are correct.

Advisors only

One round of the game has been randomly selected to count for your payment. If the Chooser you are paired with receives your recommendation in that round, you will receive \$16.00 if they produce a large plant, \$13.00 if they produce a small plant, and \$10.00 if they produce no plant. If the Chooser you are paired with receives the computer's recommendation in that round, you will receive \$10.00 regardless of the size of the plant they produce.

Beliefs experiment only

One of the guessing tasks has also been randomly selected to count for your payment. You will receive \$5.00 if your guesses are correct.

Quiz

On the next page, you will complete a quiz that tests your understanding of these instructions. You must answer all quiz questions correctly before proceeding to the first round. You will receive \$1.00 if you answer all quiz questions correctly on your first try. You will not be able to return to the instructions while completing the quiz, so please ensure you have read them carefully.

Quiz

Question 1

When a sprinkler is turned on...

- ☐ only the pot closest to it receives any water.
- ☐ the pot closest to it receives the most water; the pot farthest from it receives the least.

Question 2

The computer always recommends to...

- ☐ stick with sprinkler 1.
- ☐ switch to sprinkler 2.

☐ switch to sprinkler 3.

CHOOSERS ONLY

Question 3

Which of these statements is true?

- ☐ Advisors have the same goal as you.
- ☐ Advisors choose a sprinkler to recommend for each possible plant pot.
- ☐ Advisors can recommend sticking with the default or switching to one of the other sprinklers.

Question 4

In each round, you will be...

- ☐ randomly matched to an Advisor.
- ☐ matched with the same Advisor.

ADVISORS ONLY

Question 3

After receiving a recommendation to switch sprinklers, a Chooser can turn on...

- ☐ the default sprinkler or the recommended sprinkler.
- ☐ any sprinkler.

Question 4

Which of these statements is true?

- ☐ When making their decisions, Choosers know their chance of receiving a recommendation from each potential source.
- ☐ Choosers know that Advisors have the same goal as them.
- ☐ Choosers know that Advisors choose a sprinkler to recommend for each possible plant pot.
- ☐ All of the above statements are true.

Question 5

In each round, you will be...

- ☐ randomly matched to a Chooser.
- ☐ matched with the same Chooser.

Question 6

Your task consists of three choices: you must choose a sprinkler to recommend to the Chooser for each possible plant pot (pot 1, 2 or 3). You should think of this task as if...

- ☐ you know a specific pot (pot 1, 2 or 3) is the plant pot, and with this knowledge, must recommend a sprinkler.
- ☐ you are sending recommendations without knowing which pot is the plant pot.

Quiz complete

Congratulations! You have answered all of the quiz questions correctly. Once all participants have done so, we will proceed to the practice round.

Round 1 (practice round)

CHOOSERS ONLY

A new round has begun. You have been randomly matched to an Advisor, and the plant pot and source of the recommendation shown to you have been determined.

ADVISORS ONLY

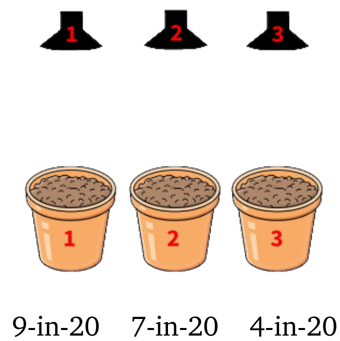
A new round has begun. You have been randomly matched to a Chooser, and the plant pot and source of the recommendation shown to that Chooser have been determined.

Round 1 (practice round)

CHOOSERS ONLY

Summary of instructions

- A seed will grow into a larger plant if it receives more water.
- Your goal (same as the Advisor's): Produce the largest possible plant.
- Your task: Decide whether, if recommended to switch to a given sprinkler, you will accept the proposed switch or stick with the default (sprinkler 1).
- There is a 1-in-3 chance you will receive the Advisor's recommendation, and a 2-in-3 chance you will receive the computer's recommendation. The Advisor chooses a sprinkler to recommend for each possible plant pot. The computer always recommends to switch to sprinkler 3.



The pots, the sprinklers and each pot's chance of being the plant pot.

Your task

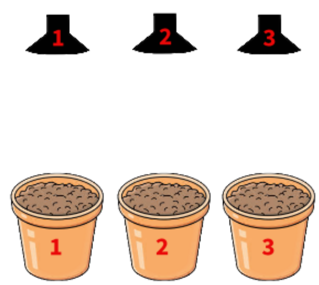
Please decide whether you will accept or reject the recommendation in each of the scenarios below.

- You receive a recommendation to switch to SPRINKLER 2. Will you accept or reject this recommendation?
  - ☐ Accept (switch to sprinkler 2)
  - ☐ Reject (stick with sprinkler 1)
- You receive a recommendation to switch to SPRINKLER 3. Will you accept or reject this recommendation?
  - ☐ Accept (switch to sprinkler 3)
  - ☐ Reject (stick with sprinkler 1)

Advisors only

Summary of instructions

- A seed will grow into a larger plant if it receives more water.
- Your goal (same as the Chooser's): Produce the largest possible plant.
- Your task: Choose a sprinkler to recommend to the Chooser for each possible plant pot. The Chooser decides whether, if recommended to switch to a given sprinkler, they will accept the proposed switch or stick with the default (sprinkler 1).
- There is a 1-in-3 chance the Chooser will receive your recommendation, and a 2-in-3 chance they will receive the computer's recommendation. The computer always recommends to switch to sprinkler 3.



9-in-20 7-in-20 4-in-20

The pots, the sprinklers and the information Choosers have (each pot's chance of being the plant pot)

Your task

Please choose a recommendation to send the Chooser in each of the scenarios below.

- You know that POT 1 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- You know that POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- You know that POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

Round 1 (practice round)

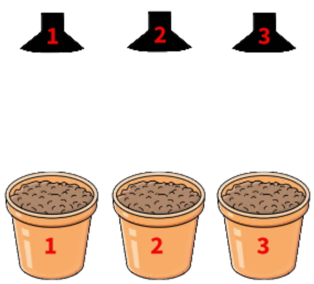
Feedback

Here is an outline of the feedback you will receive at the end of a round.

CHOOSERS ONLY

Advisor's decisions: The Advisor you were paired with made the following decisions.

- If Pot 1 is the plant pot, recommend to *[stick with sprinkler 1/switch to sprinkler 2/switch to sprinkler 3]*
- If Pot 2 is the plant pot, recommend to *[stick with sprinkler 1/switch to sprinkler 2/switch to sprinkler 3]*
- If Pot 3 is the plant pot, recommend to *[stick with sprinkler 1/switch to sprinkler 2/switch to sprinkler 3]*



This image will additionally show the sprinkler that was turned on and the plant that was produced.

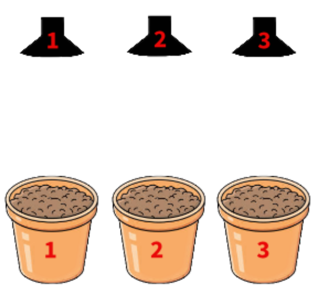
Outcome of this round: *Here, you will receive feedback on...*

- *which pot was the plant pot*
- *the source of the recommendation you received*
- *the sprinkler that was turned on*
- *the size of the plant produced*

ADVISORS ONLY

Chooser's decisions: The Chooser you were paired with made the following decisions.

- If recommended to switch to sprinkler 2, *[accept/reject]* this recommendation
- If recommended to switch to sprinkler 3, *[accept/reject]* this recommendation



This image will additionally show the sprinkler that was turned on and the plant that was produced.

Outcome of this round: *Here, you will receive feedback on...*

- *which pot was the plant pot*
- *the source of the recommendation the Chooser received*
- *the sprinkler that was turned on*
- *the size of the plant produced*

Practice round complete

Congratulations! You have completed the practice round. Once all participants have done so, we will proceed to the rest of the experiment.

MAIN EXPERIMENT ONLY

Recall that one of the rounds that follow has been randomly selected to count for your payment. You should therefore treat each round as if it is the one that counts.

BELIEFS EXPERIMENT ONLY

Recall that one of the rounds that follow, as well as one of the guessing tasks, have been randomly selected to count for your payment. You should therefore treat each round and guessing task as if it is the one that counts.

AUTHOR'S NOTE: Participants then played 21 more rounds of the Plant Pot Game. In rounds 2 through 21, the introduction and decision pages were the same as in the practice round. The structure of the feedback page was the same in rounds 2 through 22, but its content depended on the outcome of the round. Below I show an example of the feedback that players in each role may receive.

Round *n*

Feedback

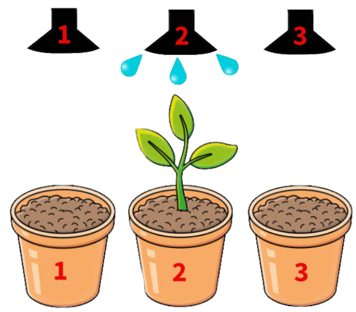
Choosers only

Advisor's decisions: The Advisor you were paired with made the following decisions.

- If Pot 1 is the plant pot, recommend to stick with sprinkler 1
- If Pot 2 is the plant pot, recommend to switch to sprinkler 2
- If Pot 3 is the plant pot, recommend to switch to sprinkler 3

Outcome of this round:

- Pot 2 was the plant pot.
- You received the Advisor's recommendation, which, as described above, was to switch to sprinkler 2.
- You accepted this recommendation, thereby switching to sprinkler 2.
- As a result, you produced a large plant.



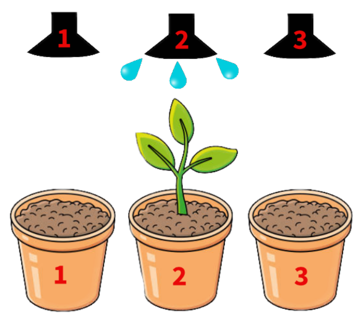
Advisors only

Chooser's decisions: The Chooser you were paired with made the following decisions.

- If recommended to switch to sprinkler 2, accept this recommendation
- If recommended to switch to sprinkler 3, reject this recommendation

Outcome of this round:

- Pot 2 was the plant pot.
- The Chooser received your recommendation, which was to switch to sprinkler 2.
- As described above, the Chooser accepted this recommendation, thereby switching to sprinkler 2.
- As a result, the Chooser produced a large plant.



Author's note: In the beliefs experiment, participants viewed and completed the page below before the introduction page in rounds 2, 8, 14 and 20.

Guessing task for round m

CHOOSERS ONLY

Instructions

- **Your task:** We want you to think about the recommendations that tend to be sent in the round you are about to play, depending on which pot is the plant pot. You must (1) guess which recommendation has most-frequently been sent by Advisors when they knew a given pot was the plant pot in round m of previous sessions of this experiment, and (2) state which recommendation the computer sends when any given pot is the plant pot. Part (2) does not involve guessing; its purpose is to ensure you remember which recommendation the computer always sends.
- **Payment:** If this guessing task counts for your payment, you will be paid $5.00 if all of your guesses about Advisors' recommendations, as well as all of your statements about the recommendation the computer sends, are correct.

Your task

(1) In ROUND m, which recommendation is most-frequently sent by **Advisors** when they know...

- ... POT 1 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... POT 2 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... POT 3 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

(2) In ALL ROUNDS, which recommendation does **the computer** send when...

- ... POT 1 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... POT 2 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... POT 3 is the plant pot?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

ADVISORS ONLY

Instructions

- **Your task:** We want you to think about the decisions Choosers tend to make in the round you are about to play. You must guess Choosers' most-common response to a recommendation to switch to a given sprinkler in round m of previous sessions of this experiment.
- **Payment:** If this guessing task counts for your payment, you will be paid $5.00 if both of your guesses are correct.

Your task

In ROUND m, what is Choosers' most-common response to a recommendation to...

- ... switch to SPRINKLER 2?
  - ☐ Accept (switch to sprinkler 2)
  - ☐ Reject (stick with sprinkler 1)
- ... switch to SPRINKLER 3?
  - ☐ Accept (switch to sprinkler 3)
  - ☐ Reject (stick with sprinkler 1)

AUTHOR'S NOTE: In round 22, Choosers' introduction and decision pages were the same as in all previous rounds. These pages were different for Advisors (see below).

Round 22

AUTHOR'S NOTE: The two introduction sentences were the same as in all previous rounds.

*In this final round, you will be able to make different recommendations depending on (1) which pot is the plant pot **AND** (2) the Chooser's decision to accept or reject recommendations to switch sprinklers.*

Round 22

AUTHOR'S NOTE: The summary of the instructions was the same as in all previous rounds.

*In this final round, you will be able to make different recommendations depending on (1) which pot is the plant pot **AND** (2) the Chooser's decision to accept or reject recommendations to switch sprinklers.*

Your task

Please choose a recommendation to send the Chooser in each of the scenarios below.

You know the Chooser would ACCEPT to switch to SPRINKLER 2 and ACCEPT to switch to SPRINKLER 3, and...

- ... you know POT 1 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2

- ☐ Switch to sprinkler 3

- ... you know POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 3 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

You know the Chooser would ACCEPT to switch to SPRINKLER 2 and REJECT to switch to SPRINKLER 3, and...

- ... you know POT 1 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 3 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

You know the Chooser would REJECT to switch to SPRINKLER 2 and ACCEPT to switch to SPRINKLER 3, and...

- ... you know POT 1 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 3 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

You know the Chooser would REJECT to switch to SPRINKLER 2 and REJECT to switch to SPRINKLER 3, and...

- ... you know POT 1 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 2 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3
- ... you know POT 3 is the plant pot. What do you recommend to the Chooser?
  - ☐ Stick with sprinkler 1
  - ☐ Switch to sprinkler 2
  - ☐ Switch to sprinkler 3

MAIN EXPERIMENT ONLY

Rounds complete

Congratulations! You have now completed all rounds of the game. Click the "Next" button to proceed to the survey.

BELIEFS EXPERIMENT ONLY

Rounds and guessing tasks complete

Congratulations! You have now completed all rounds of the game and all of the guessing tasks. Click the "Next" button to proceed to the survey.

Survey (Page 1 of 2)

CHOOSERS ONLY

We're interested in knowing more about how you decided whether to accept a recommendation.

ADVISORS ONLY

We're interested in knowing more about how you decided which recommendations to send.

Did your approach to this decision remain constant, or did it change as the rounds progressed?

- ☐ My approach remained constant.

☐ My approach changed as the rounds progressed.

Please explain why your approach did or did not change as the rounds progressed. [Text field where participant types answer]

Do you have any other thoughts or comments on the experiment? If not, leave this field blank. [Text field where participant types answer]

Survey (Page 2 of 2)

We want to know more about you! Please take a moment to answer these demographic questions.

What degree are you currently pursuing?
☐ Bachelor's degree
☐ Master's degree
☐ Doctoral degree
☐ Other degree (not listed)

In what year did you begin the degree you are currently pursuing?
☐ 2024
☐ 2023
☐ 2022
☐ 2021
☐ 2020
☐ 2019
☐ 2018 or earlier

What is your primary field of study (e.g., Anthropology, Chemistry, Mechanical Engineering, etc.)? [Text field where participant types answer]

How old are you? [Text field where participant types answer]

What is your gender?
☐ Female
☐ Male
☐ Non-binary

☐ Other (not listed)

☐ Prefer not to answer

AUTHOR'S NOTE: Below I show an example of a participant's earnings page.

Your earnings

MAIN EXPERIMENT ONLY

Round 10 was randomly selected to count for your payment. As summarized in the table below, your total earnings are **$22.00**.

Earnings breakdown	
Show-up fee	$5.00
Earnings from quiz	$1.00
Earnings from round 10	$16.00
Total earnings	*$22.00*

BELIEFS EXPERIMENT ONLY

Round 10 and the guessing task for round 2 were was randomly selected to count for your payment. As summarized in the table below, your total earnings are **$27.00**.

Earnings breakdown	
Show-up fee	$5.00
Earnings from quiz	$1.00
Earnings from round 10	$16.00
Earnings from guessing task for round 2	$5.00
Total earnings	*$27.00*

An experimenter will come by your desk shortly and pay you this amount in cash. Thanks for participating!

C Participant characteristics

	Main	Beliefs	p-value
Age	21.1	21.3	0.95
% women	60.0	67.4	0.21
% bachelor's degree program	85.9	80.6	0.25
Comprehension quiz mistakes	0.99	1.15	0.21

Notes. Mean participant characteristics by experiment. Excludes the data of one participant from each experiment who withdrew from their experimental session. The final column reports the p-value from a Wilcoxon-rank-sum test.

Table C.1: Participant characteristics