THE VALUE OF AND DEMAND FOR DIVERSE NEWS SOURCES

EVAN M. CALFORD* AND ANUJIT CHAKRABORTY**

Abstract. We study the value of and the demand for instrumentally-valuable information in a simple decision environment where signals are transparently polarized. We find that in both information aggregation and acquisition, subjects use sophisticated heuristics to counter the polarization in signals. Even though the number of precise Bayesian reports are small, most subjects (64%) generate unpolarized reports even when faced with polarized signals. Subjects placed in a market place of information rarely end up buying polarized signals and instead overwhelmingly opt for diverse information. The demand for diverse information increases as diverse information becomes more valuable and decreases as it becomes more expensive.

Keywords: Experimental economics, information acquisition, information aggregation.

JEL codes: D90, D91, D81.

---

* Research School of Economics, Australian National University; evan.calford@anu.edu.au.
** Department of Economics, University of California, Davis; chakraborty@ucdavis.edu.
Funding for subject payments was provided by the Krannert School of Management, Purdue University.
The experiments were approved by the Purdue IRB, protocol number 1706019387. Funding was provided by the Krannert School of Management, Purdue University. We thank seminar audiences at UC Davis, UC Santa Barbara, and Queensland University of Technology.

Media sources with contradictory biases can create completely different impressions of what actually happened through the selective omission of details and the choice of words. As an example of media bias, Gentzkow and Shapiro [2006] discuss the media coverage of the firefight in the Iraqi city of Samarra on December 2, 2003. Fox News, a conservative US based news channel, began its story with the following paragraph:

> In one of the deadliest reported firefights in Iraq since the fall of Saddam Hussein's regime, US forces killed at least 54 Iraqis and captured eight others while fending off simultaneous convoy ambushes Sunday in the northern city of Samarra.

And the English-language website of the Qatar-based Al Jazeera (AlJazeera.net) began its report on the same incident with:

> The US military has vowed to continue aggressive tactics after saying it killed 54 Iraqis following an ambush, but commanders admitted they had no proof to back up their claims. The only corpses at Samarra's hospital were those of civilians, including two elderly Iranian visitors and a child.

Similarly, the narrative of who is winning the war in Ukraine depends on who is doing the talking. Predictably, Russian media says that it is winning an easy war, while the United States insists that bolstered by Western support, Ukraine might be staging a surprise win. Investors following geopolitical events to predict financial markets or future fuel prices might get very different impressions of the same event from international news sources with opposing bias. Media bias also extends to financial news, which determines the quality of financial decisions a news-reader makes: Niessner and So [2018] find that a financial news story is approximately 22 percent more likely to be covered if it is negative, creating an overall negative media bias in financial news coverage.

Investors who exclusively access media sources that share identical polarization would only observe an incomplete account of the whole reality. As creatures that often make quick judgments based on heuristic-driven biases, for example hot hand bias (Marquis de Laplace, 1840), Gambler's Fallacy (Chen et al., 2016a), law of small numbers (Rabin, 2002) and base rate neglect (Kahneman and Tversky, 1973), *how do we aggregate instrumental information* *when information sources are transparently polarized?*[1] *Do polarized sources lead to severe and systematic errors?* This is our first line of inquiry.

Further, a world where investors are self-aware that polarized sources might lead to systematic errors and hence actively diversify their news sources, is vastly different to a world where they lack such awareness and hence naively end up with polarized news sources. For example, any policy that subsidizes and simplifies acquiring diverse news sources would be impactful in the former world, but have no impact in the latter. Previous experiments (e.g. Enke, 2020) that insert subjects in a world of exclusively polarized news sources help us measure polarization in the counterfactual latter world, *but don't reveal which of the two worlds is a better approximation of the world we live in.* Overcoming this limitation motivates our second line of inquiry: *How often do individuals with access to a market-place for information, end up with polarized information sources? How does their demand for diverse news sources react to its value and costs?* This paper uses experimental methods to address these questions.

To study the first question, we immerse subjects in a controlled environment of selective information-omission to compare the accuracy of their opinions under diverse versus polarized information sources. Every round, subjects are asked to guess an objective state, the average of seven i.i.d random draws from the simple uniform distribution {1,2,3,..100}. To incentivize subjects, more accurate guesses earned a higher expected payment. In our first two treatments, subjects were *randomly* shown only three of the seven random draws as *signals* beforehand. The polarization in observed signals is *extreme* when all three observed signals are high (between 51-100), thus creating a high-polarity, or when all three observed signals are low (1-50), creating a low-polarity. This is similar to being exclusively exposed to media sources that share the same reporting polarization on an issue. Instead, if the three observed signals are a mix of high and low, then we will describe the signals as *diverse.*

[1]Psychology studies have long proposed a dual-process model for two modes of information-processing: a "fast, associative" one "based on low-effort heuristics" (System 1), and a "slow, rule based" one (System 2) that relies on "high-effort systematic reasoning" (Kunda, 1990).

In the Baseline treatment (called the **No Colors** treatment), the subjects do not have any information about the unobserved signals, other than knowing that they are equally likely to be any number between {1,2,3,..100}. Given the observed three signals were randomly chosen from the seven draws, subjects might assume that their observed sample is "representative" of the seven draws. An intuitive way to estimate of the average of all seven signals would be to calculate the average of the three observed signals. Unfortunately, despite the intuitive appeal, such a sample-average rule ignores the information available about the unobserved signals, that they are equally likely to be any number in {1,2,...100}. Consequently, the sample-average would be high-polarized when all observed signals were high, and low-polarized when all observed signals were low. Another intuitive but incorrect estimation rule, one that falls prey to the Gambler's fallacy, is to that mistakenly infer that observing three low (or high) signals increases the likelihood of the unobserved signals being of the opposite polarity, causing reports to be biased in the opposite direction of the extremity of observed signals. Finally, the correct Bayesian rule would add the three observed signals with four times 50.5 (as the four unobserved signals are uniformly distributed between 1 and 100), and then divide the sum by 7.

Contrary to our prior expectations, subjects are remarkably unbiased in adjusting their opinions for extreme signal bias: only 16% of subjects exhibited a consistent bias towards the sample-average heuristic and only 20% exhibited a consistent bias towards the Gambler's fallacy. The majority (64%) of subjects are able to counteract the effect of polarized information. Even though the number of precise Bayesian reports are quite small, we find no evidence for subjects over-weighting the information of the observed signals or under-weighting/ignoring the prior information available about the unobserved signals.

Next, as a benchmark for an easier decision environment, we designed the **Colors** treatment, where the subjects are also informed how many of the four unobserved signals lie between 1-50 (called Blue signals) or 51-100 (called Red signals). The correct Bayesian rule here would add the three observed signals with 25.5 times the number of Blue signals and 75.5 times the number of Red signals. At a minimum, the Colors treatment alerts the subjects that the observed signals might not be informative about the unobserved signals (thus moving them away from Gambler's fallacy), and at a maximum, helps subjects account for the unobserved signals. Reduced form estimates suggest that this extra information improves report quality for inexperienced subjects who observe extreme signals, but otherwise has a limited impact on report quality. In addition, we find that playing the more informative Colors treatment before the Baseline (No Colors) treatment does not improve the report-accuracy in the Baseline treatment.

We perform a structural estimation exercise, in part motivated by the dual-process cognitive model (see footnote 1), to further unpack how subjects aggregated information. We assume that reports were generated either from a precise mental model that uses the Bayesian rule (System 2 thinking) or from an imprecise mental model that heuristically aggregates all the available information (System 1 thinking). For each treatment (Colors and No Colors) and each signal type (diverse and extreme), we separately estimate the likelihood that the reports are generated through each of the two mental models, and the parameters of the imprecise heuristic rule. We find that more than 70% of all reports, under all conditions, are estimated to have originated from the imprecise heuristic. In rounds where subjects observe diverse signals, there are only minor differences in the estimated relative frequency of each rule and the heuristic used, across the No Colors and Colors treatments. For extreme signals, however, we observe a marked difference in behavior across the Colors and No Colors treatments: reports are more likely to have come from the precise Bayesian model in Color treatments. Further, the estimated heuristic-models produce *noisy but unbiased reports* under all treatment-signal combinations, except when subjects receive extreme signals in the Colors treatment.

Next, we study if subjects in a *market-place for instrumental information* are unlikely to end up with extreme signals, as they realize that extreme signals might lead to worse decisions. In our **Active Choice** and **Average** treatments, subjects are only assigned their first signal, and then can buy two more signals. They can choose to buy two, one or zero signals with the same polarity of their first signal. Signals of the same or opposite polarity are priced differently. Buying two signals of the same polarity would result in an extreme information source. Buying one or zero signals with the same polarity would result in a diverse information source, but would come at different costs. Observing what information subjects buy, allows us to investigate how subjects facing polarized information, opt into or opt out of particular informational environments. When diverse and extreme signals are equally costly, we find that approximately 95% of subjects choose a diverse portfolio of signals. The demand for diverse signals persists as the cost of diverse information increases. As the price for diverse signals increases, subjects become less likely to buy them, but even at the most extreme price differentials 20% of subjects continue to purchase diverse signals. Thus, subjects who are not constrained by experimental rules, seldom face extreme signals.[2]

In our **Average** treatment, we impose the rule that after subjects choose a signal portfolio, a sample-average heuristic would calculate the report on their behalf. This intervention significantly and exogenously increases the value of diverse information as the sample-average heuristic is severely inaccurate under extreme information. As accurate reports pay more, subjects should select signals that preserve accuracy under the imposed sample-average rule. In our experiments, most of the subjects realize that the accuracy premium from diverse signals is even higher under the sample-average rule and it is reflected in their signal choices. The demand for diverse signals is higher in the Average treatment than the Active Choice treatment, increasing by more than 50% at intermediate information costs.

On aggregate, we find that in both information aggregation and acquisition, subjects use sophisticated heuristics to counter the polarization in signals. Our results advocate for greater transparency in media bias, as we show that consumers who wish to make an informed choice, actively seek diverse news sources in such a world. For example, websites like Allside.com or Adfontes-media.com that explain and measure various dimensions of media polarization can be extremely useful in the quest for transparency. In particular, All Sides

[2]Notably, we also find evidence of metacognition in signal choice: subjects who were worse at aggregating information were also more likely to purchase diverse signals, perhaps in an attempt to make their inference problem easier.

is a news website that presents multiple sources side by side in order to provide the full scope of news reporting. It also provides a Bias Ratings page that allows a visitor to filter a list of news sources by the polarization on the political spectrum (left, center, right).[3] Even though we do find a strong preference for diversity of information, our structural model does suggest that increased availability of information could increase ex-post polarization among some consumers who continue to consume polarized media sources.

Our paper is related to information acquisition and to a large literature, pioneered by Tversky and Kahneman [1971, 1974], that studies the prevalence of biases and heuristics in probabilistic decisions. For clean identification, our experimental environment excludes the scope for motivated reasoning (Kunda [1990]), the tendency of people to conform assessments of information to some goal or end extrinsic to accuracy. Similarly, it also excludes mistakes originating in the failure of hypothetical or contingent reasoning; any event that subjects should condition on is clearly and explicitly displayed. We discuss the related literature in detail in Section 1.

1. Related Literature

The sample-average rule, that subjects might find naturally attractive in our decision environment, belongs to the class of simplistic heuristics that overweight one type of information over other available information.[4] For example, experimental subjects frequently discard or under-weight base-rate information because it is not relevant to judgements of representativeness (Kahneman and Tversky, 1973), or because likelihood information is more "vivid, salient, and concrete" (Nisbett and Ross, 1980). Subjects in our treatments might similarly overweight the information from the three signals that they observe, thinking it is representative of the seven signals, or because the information it provides is more salient and concrete. A recent literature in cognitive psychology connects such behavior to people behaving like "naive intuitive statisticians" who despite being skilled in making judgements based on memory-stored frequencies, often naively assume that their information samples are representative and

[3]See their media bias chart here.

[4]Benjamin [2019] provides a detailed literature review.

that sample properties can be directly used to estimate population analogs (Fiedler and Juslin, 2006).

The Gambler's Fallacy (GF in short) is one of the oldest documented probabilistic biases. Marquis de Laplace [1840] described people's belief that the fraction of boys and girls born each month must be roughly balanced, so that if more of one sex has been born, the other sex becomes more likely. Rabin [2002] and Oskarsson et al. [2009] review the extensive literature that documents the GF in surveys and experiments. Chen et al. [2016a] finds consistent evidence of negative autocorrelation in decision making that is unrelated to the merits of the cases considered in three separate high-stakes field settings: refugee asylum court decisions, loan application reviews, and Major League Baseball umpire pitch calls. They link it to people underestimating the likelihood of sequential streaks occurring by chance—leading to negatively autocorrelated decisions that result in errors. Similarly, experimental participants playing a game with a unique mixed-strategy Nash Equilibrium or tennis players making a serve switch their actions too often (Rapoport and Budescu, 1997, Gauriot et al., 2016), and this excessive switching could reflect the mistaken GF intuition for what random sequences look like.

Studies on information extraction are also closely related to the growing literature on contingent reasoning or hypothetical thinking. For example, to avoid the winner's curse, bidders in a common value auction should extract information from their private signal, while conditioning on the hypothetical event of winning the auction. Similarly, voters should extract information from their private signal, while conditioning on the hypothetical event of being pivotal to the outcome. Esponda and Vespa [2019], Araujo et al. [2021] find that experimental subjects routinely fail to perform such contingent reasoning while processing their private information. Enke [2020] finds that when subjects are exclusively shown information consistent with their initially reported prior, they often behave as if the sample selection does not even come to their mind.[5] Enke [2020] provides further causal evidence that the frequency of

[5]Enke's Random treatment is similar to our NoColors treatment, with one important difference in the user interface that critically reflects the different behaviors we seek to capture. In our user interface, subjects are visually prompted about the missing signals (although they get no further information about these in the NoColors treatment), while Enke deliberately

such incorrect mental models is a function of the computational complexity of the decision problem. To disentangle the value of and demand for diverse information from these other behavioral forces, we offer subjects a simple decision problem where the difference between extreme and diverse signals is transparent, even to subjects who cannot reason through hypothetical events.

Our Active Choice and Sample Average treatments are novel and connect to the nascent literature on the value of instrumental information. Reshidi et al. [2021] contrast information acquisition in groups versus individual treatments, and in static versus dynamic contexts. Duffy et al. [2019, 2021] study how subjects choose between social and private information sources that vary in relative quality. Charness et al. [2021] and Montanari and Nunnari [2019] study how subjects update their beliefs about a payoff-relevant state of the world while choosing *exactly one of two information sources* (signals) which have mutually opposite biases, and thus are the closely related. In both studies, subjects had to guess the probability that a single ball drawn randomly from an urn would be of a particular color, and also guess the color of the ball. To inform their guesses, subjects first chose one of a pair of computerized advisors, from which to receive an informative signal about the ball drawn. Subjects were fully informed of the probabilities with which each advisor would provide each signal as a function of the true color of the ball drawn from the urn. Charness et al. [2021] find that the fraction who choose information optimally and the fraction who use a mistaken confirmation-seeking rule are roughly equal. In Montanari and Nunnari [2019], when the two information sources are equally reliable, subjects select information optimally. But, when the source less supportive of the prior belief is more informative, subjects display a dis-confirmatory pattern of information acquisition that is not always consistent with the theoretical predictions. Even in cases where information is not instrumentally valuable, subjects might have preferences over how information is disclosed (Zimmermann, 2015, Ganguly and Tasoff, 2017,

does not prompt about the missing signals on the decision page. Enke [2020]'s design asks the eponymous question: is what you see all there is? While Enke studies whether people ignore signals that are not visually presented, we study how people process signals where the realization of (rather than the existence of) the signal is shrouded. The difference also extends to our findings: while Enke [2020] finds systematic evidence of sample-average bias, we find more subjects with the opposite (Gambler's fallacy) bias.

Masatlioglu et al., 2017, Nielsen, 2020). An emerging literature on motivated reasoning addresses how we often seek particular information and stay willfully ignorant of other information, because we wish to arrive at our desired conclusion (Festinger, 1962). Here, we deliberately frame our experimental tasks to remove the scope for motivated reasoning, as we want to to study to the demand for information with *purely* instrumental value.

Acquisition of instrumental information has also been studied in applied settings. Fuster et al. [Forthcoming] use an experimental survey instrument to determine which pieces of economic data subjects prefer to consult when predicting house price movements. Burke and Manz [2014] asked subjects to forecast inflation in a simulated laboratory economy, and provided subjects with a choice of viewing historical information on inflation, interest rates, unemployment, population growth, or price changes of specific commodities, before making their forecast. In both environments choices of more informative sources were correlated with measures of economic sophistication. Mikosch et al. [2021] study how information acquisition about the future development of the exchange rate is related to the exposure to and uncertainty about exchange rate risk, and the perceived information acquisition costs. Roth et al. [2022] find that a higher personal exposure to unemployment risk during recessions increase the demand for an expert forecast about the likelihood of a recession. Capozza et al. [2021] provide a detailed review the emerging literature on information acquisition in field settings.

Most of the lab-experimental literature on information-acquisition compares empirical information choice to a theoretical optimal calculated for the Bayesian subjects. This allows for sharp theoretical predictions, but the test for optimal signal choice becomes a joint test of optimal choice and subjects updating using Bayes Rule. The information-choice data can only reject optimal choice when information-updating behavior is consistent with Bayes Rule. There are, however, some exceptions in the literature. Charness et al. [2021] use a treatment with exogenously assigned signals to show that the Bayesian optimal information choice was also optimal for the behavioral non-Bayesian types in their subject pool (i.e. the optimality of information choices is not dependent on an assumption of Bayesian updating). Ambuehl and Li [2018] measure the value of information against both a Bayesian benchmark and an empirical updating benchmark and find that subject's value for information is higher, but still lower than optimal, when measured against the empirical benchmark. We instead test if subjects react optimally to an increase in the value of diverse signal choice and we do not require subjects to be Bayesian. All we require is that diversification is more valuable under the sample-average rule than whatever updating rule subjects actually use. As we show later in Figure A.2, this is indeed true for all individual participants in our study. Further, in our Average treatment we exogenously impose a sub-optimal updating rule on subjects which allows us to measure, unconfounded by updating ability, subject propensity to select information that offsets updating biases.

2. Experimental Design

To study the acquisition and aggregation of extreme information, we implement 4 treatments: No Colors (NC), Colors (C), Active Choice (AC) and Average (Avg). The first two treatments randomly allocate signals (information) to study information aggregation in isolation. The last treatment fixes the information aggregation procedure exogenously to study information acquisition in isolation, and the AC treatment combines both information aggregation and information acquisition tasks.

2.1. **Baseline/ No Colors treatment (NC):.** The first treatment is No Colors (NC henceforth). Each round in the NC treatment has the following structure:

(1) **Signal realization stage:** A set of seven signals are independently drawn from the numbers 1 to 100. Each signal, s_i for $i \in \{1, \dots, 7\}$ is assigned a color based on its realization: Blue if $s_i \leq 50$ and Red if $s_i \geq 51$.
(2) **Information stage:** In the experiment, we framed s_1 as the subject's own signal and each other signal $\{s_2, .., s_7\}$ as belonging to a passive computer "player". Subjects see the realization of s_1, and two other randomly chosen signals (say s_2 and s_3).[6] We use this framing to imitate information flow in the subject's information network.

[6]Given the signals are ordered randomly, this is equivalent to showing subjects any three signals randomly.

Subjects are not informed about the colors of the remaining 4 signals. For example, in a particular round, if the signal realization were $\{\underbrace{10}_{s_1}, \underbrace{20}_{s_2}, \underbrace{30}_{s_3}, 40, 50, 60, 70\}$, and s_2 and s_3 were shown to the subject, then the subject would observe $\{10, 20, 30\}$ and they would *not know* the color composition of the unobserved 4 signals. Thus the name No Colors treatment.

(3) **Updating/ aggregation stage:** Subjects state their best estimate of the average value $\overline{s} = \frac{\sum_{i=1}^{7} s_i}{7}$ given the information above.

One crucial feature of our design is that we did not provide any feedback to the subjects in between the rounds, in any of our treatments. Subjects were provided with a hand-held calculator, although few subjects elected to use the calculators; we did not want to test the ability of subjects to do basic arithmetic.

At the end of the experiment, one of the rounds was randomly chosen as the round for which subjects were paid. For the chosen round, the guessing error was calculated as the absolute difference between the reported guess and the realized $\overline{s} = \frac{\sum_{i=1}^{7} s_i}{7}$. Each subject won a large prize worth 200 points with $\max(100 - 6 \times error, 0)\%$ chance, and a small prize worth 50 points with the complementary probability. This payment function ensures that reports are unaffected by the curvature of the subject's utility function over money. The linear loss function was chosen as it is simple and it incentives reporting the expected value of $\bar{s}$ truthfully under an assumption that the belief about $\bar{s}$ is symmetric and single peaked.[7]

2.2. **Colors treatment (C):.** The Colors treatment is identical to the No Colors treatment, except, in the **Information stage**, subjects are also shown the colors of the remaining 4 remaining signals. For example, in a particular round, suppose the signal realizations were $\{\underbrace{10}_{s_1}, \underbrace{20}_{s_2}, \underbrace{30}_{s_3}, \underbrace{40}_{\leq 50}, \underbrace{50}_{\leq 50}, \underbrace{60}_{>50}, \underbrace{70}_{>50}\}$. If the realizations s_2 and s_3 were shown to the subject, then the subject observed $\{10, 20, 30\}$ and additionally they would know that among the unobserved 4,

[7]This assumption obviously holds for Bayesian subjects, but can hold more generally.

there are 2 Blue ($<= 50$) and 2 Red (> 50) signals. The Colors treatment makes the importance of unobserved signals salient.

2.3. **Active choice treatment (AC):.** The AC treatment is identical to the Colors treatment, except, in the **Information stage**, subjects make an active choice about the signals they observe. Subjects observe s_1 by default, and they can choose to observe the realizations of two Blue signals or two Red signals or one signal of each color. For example, in a particular round, if the signal realization were $\{\underbrace{10}_{s_1}, \underbrace{20, 30, 40, 50}_{Blue}, \underbrace{60, 70}_{Red}\}$, then the subject only sees $\{10\}$ by default. If she chooses to see 2 Blue signals, then she is randomly shown exactly two signals from $\{\underbrace{20, 30, 40, 50}_{Blue}\}$ and she is told that among the unobserved 4, there are 2 Blue and 2 Red signals. If she instead chooses to see 2 Red signals, then she is shown both $\{60, 70\}$ and she is told that all the unobserved 4 signals are Blue. If she instead chooses to see 1 signal of each color, then she is shown one of $\{20, 30, 40, 50\}$, one of $\{60, 70\}$ and she is told that among the unobserved 4, there are 3 Blue and 1 Red signals.

The subjects made their signal selection without knowing if their choice were available. In the improbable event that the selected signals are not available (for example, all but one of $\{s_2, s_3, .., s_7\}$ are Blue but the subject requests two Red signals), then the subject is shown a combination of signals that is as close as possible to matching the subjects requested combination of colors (in this example, the subject would be shown one each of Red and Blue signals). The subject is always informed of the colors of any unobserved signals (in this example, there are 4 unobserved Blue signals).

Further, each round, subjects are paid p_1 for each signal they observe that is the same color (Blue/ Red) as s_1 and p_2 for each observed signal that is of the opposite color from s_1. We vary p_1 and p_2 to generate the price-differences

$$\Delta p = (p_1 - p_2) \in \{-6, -2, 0, 2, 4, 6, 8, 10, 14, 20\}$$

When $\Delta p > 0$, subjects have an explicit monetary incentive for choosing more signals that are of the same color as s_1. When $\Delta p < 0$, they have the opposite incentive, and when $\Delta p = 0$, there is no explicit incentive and subjects should choose the set of signals that they subjectively view as the most informative.

Features	Treatments			
	Color (C)	No Color (NC)	Active Choice (AC)	Average (Avg)
Signal acquisition	Random	Random	Active choice	Active choice
Information Aggregation	Active choice	Active choice	Active choice	Sample-Avg
Colors of unobserved balls	Available	Unavailable	Available	Not Applicable

TABLE 1. Features of different treatments.

2.4. **Average treatment (Avg):** The Average treatment is identical to the Active Choice treatment in the Information stage: subjects choose the colors of the signals, subject to the available price-difference Δp. But, subjects know that their report in the updating/aggregation stage is *exogenously constrained to be the sample average* of their three observed signals. Thus, if a subject observes $\{10, 20, 30\}$, then the software would report 20 on their behalf. Similarly, if they observe $\{10, 20, 60\}$, then the software would report 30 on their behalf. Thus the name Average treatment. This treatment manipulation increases the value of diversified signals for all subjects, given that the sample-average updating rule suffers a large bias when all observed signals are extreme. As before, we vary p_1 and p_2 to generate the price-differences

$$\Delta p = (p_1 - p_2) \in \{-6, -2, 0, 2, 4, 6, 8, 10, 14, 20\}$$

In the Appendix, we report on a fifth treatment. We do not include this treatment in the main text because, as discussed in the Appendix, there is evidence of substantial subject confusion in the novel part of the extra treatment. The sessions that included the fifth treatment also included some rounds of the Colors and Avg treatments as well. We exclude all sessions that included the fifth treatment from the main text, but include the data from those sessions in the Appendix. All our results are robust to including the fifth treatment sessions.

2.5. **Sessions.** Each session was run with around 15 subjects. Sessions lasted 40 rounds, grouped into 2 treatment-blocks of 20 rounds each. In the table below we summarize the treatment composition of the sessions

Session name	Treatment(s) run		#Sessions
	Rounds 1-20	Rounds 21-40	
C1+NC2	Colors	No Colors	2
NC1+C2	No Colors	Colors	2
AC+Avg	Active Choice	Average	2

TABLE 2. Number of sessions conducted with each pairwise combination of treatments. C1 and C2 mean Colors treatment were run in the first and second half of the session respectively. Similarly, NC1 and NC2.

Subjects were paid for the sum of points earned during one randomly selected round. At the end of the experiment, points were converted to US Dollars at an exchange rate of $0.07 per point. Thus, the larger prize (200 points) was worth $14 and the smaller prize (50 points) was worth $3.50. They were also paid a $5.00 show up fee in addition to any money they earned during the experiment. We conducted 6 sessions in total, with a total of 89 subjects. Our experiments were conducted in-person at the Purdue University, using student subjects drawn from Purdue's implementation of the ORSEE subject database [Greiner, 2015], during 2018. The experiments were programmed using oTree [Chen et al., 2016b].

3. HYPOTHESES

3.1. **Information aggregation.**

Given the information subjects have while making a guess, the Bayesian report for both Color and No Color rounds can be calculated as follows. For the Color treatment the Bayesian report is given by

$$R_C^* = \frac{s_1 + s_2 + s_3 + 25.5 \times (\#Blue) + 75.5 \times (\#Red)}{7}$$

where $\#Blue$ and $\#Red$ are the number of unobserved Blue and Red signals, respectively, the subscript C denotes the Colors treatment, and the superscript $*$ denotes the optimal Bayesian report. In the No Colors treatment the Bayesian report is

$$R^{*}_{NC} = \frac{s_1 + s_2 + s_3 + 50.5 \times 4}{7}$$

where the subscript NC denotes the No Colors treatment. The Color treatment provides subjects with more information about the unobserved signals and will, on average, lead to better performing reports in the Color treatment than the No Colors treatment. Our first set of hypotheses concern the effects of the distribution of the three observed signals on the quality of reports. We call a set of three signals *extreme* if they are all Red (51 or above) or all Blue (50 or below). If a set of signals is not *extreme* then it is *diversified*. For example, $\{10, 20, 30\}$ would be a set of extreme signals, and so would $\{95, 60, 70\}$. We initially focus on two behavioral biases that are ex-ante plausible in the NoColor rounds.

1) **Sample average** bias: An intuitive but incorrect decision rule would be to completely disregard any information present in the unobserved signals and simply report the sample average of the three observable signals. That is, a naive, sample-average report would be to report

$$\underbrace{R^{SA}_{NC} = \frac{s_1 + s_2 + s_3}{3}}_{\text{sample average}}$$

where the subscript SA stands for sample-average. Given the observed three signals were randomly chosen from the seven draws in the NC treatment, subjects might assume that their observed sample is "representative" of the seven draws and be attracted to such a heuristic. In this case, when the signals are extreme, the guess would be more inaccurate on average and would be *polarized towards the signal-extremity*. For example, when all the three observed signals are Blue, subjects would, on average, under-report by a large margin

$$R^{SA}_{NC}|(0R, 3B) = 25.5 < R^{*}_{NC}|(0R, 3B) = \frac{25.5 \times 3 + 50.5 \times 4}{7} = 39.79$$

Similarly, when all the signals are Red, subjects would over-report by a large margin. Instead when the observed signals are mixed, for example two Blue and one Red, the margin of error is much smaller:

$$R_{NC}^{SA}|(1R,2B) = 42.2 < R_{NC}^{*}|(1R,2B) = \frac{51 + 75.5 + 50.5 \times 4}{7} = 46.92$$

2) **Gambler's Fallacy**: The gambler's fallacy (GF) is the common, but mistaken, belief that that i.i.d. random variables are "self-correcting towards the mean" and hence exhibit negative serial correlation.[8] In our setting, the gambler's fallacy implies that subjects mistakenly believe that each observed Red (or Blue) signal reduces the likelihood of the unobserved signals being that color. For simplicity, one can think of this as the rule

$$R_{NC}^{GF} = \frac{s_1 + s_2 + s_3 + 4 \times (75.5p + 25.5(1-p))}{7}$$

where p depends on the colors of the signals s_1, s_2, s_3. Thus, if k is the number of Red balls among s_1, s_2, s_3, then, according to the GF,

$$p(k=0) > p(k=1) > \underbrace{.5}_{\text{under Bayes}} > p(k=2) > p(k=3)$$

Thus, when the signals are extreme, the guess would be further from the Bayesian estimate on average and would be *polarized in the opposite direction of the signal-extremity.* One could construct hybrid rules by taking the following convex combinations of the non-Bayesian heuristics with the Bayesian rule, that is, $\alpha R_{NC}^{*} + (1-\alpha)R_{NC}^{SA}$ and $\alpha R_{NC}^{*} + (1-\alpha)R_{NC}^{GF}$. For any value of $\alpha \in [0,1)$ these hybrid rules would inherit the directional bias of their parent non-Bayesian rule. In particular, the former rule would *over-weight* the information in the observed signals, and the latter would *under-weight* the same. We construct our first hypothesis under the generalized sample-average rule $\alpha R_{NC}^{*} + (1-\alpha)R_{NC}^{SA}$ for $\alpha \in [0,1)$.

Hypothesis 1 (Error and Bias, Sample Average)**.** *In the No Colors and Colors treatments, extreme signals create reports that are further away from the Bayesian report and polarized towards the corresponding extremity.*

[8] The fallacy earns its name from the story of a gambler who, after observing a run of black numbers at a roulette table, exclaims "We are due for a red number next!"

An alternative hypothesis, using the Gambler's Fallacy and its generalized aggregation rule, would suggest a bias towards the opposite direction under extreme signals.

The Gambler's Fallacy, as described above, is ruled out in the Colors treatment. A heuristic like sample-average is also unlikely when subjects are explicitly informed about the colors of the unobserved realizations. Thus, subjects are more likely to make mistakes in the No Colors rounds, which is our next hypothesis.

Hypothesis 2 (Colors vs No Colors)**.** *Polarization is higher when the color information of unobserved balls is unavailable (No Colors treatment).*

3.2. **Learning:** Subjects get no feedback between rounds. But, consider a subpopulation of subjects who use the Sample average rule R^{SA}_{NC} in the No Color treatment. They completely disregard any information that might be present in the unobserved signals. Their play might be influenced by the order in which they play both the Color and No Color treatments. In particular, any prior experience in the Color treatment might make it salient for them that the unobserved signals play a significant role in determining the target $\overline{s} = \frac{\sum_{i=1}^{7} s_i}{7}$.[9] This experience might move them away from a naive sample-average rule to some rule that accounts for the unobserved signals (e.g, $\alpha R^*_{NC} + (1-\alpha)R^{SA}_{NC}$) and thus improve their quality of information aggregation when they get to the No Color treatment. To test this, we could compare the data from subjects who experienced Colors before the NoColors treatment (i.e. using the data from the C1+NC2 sessions) to those who did not (from the NC1+C2 sessions).

Hypothesis 3 (Learning)**.** *Compared to the NC1 condition, subjects in the NC2 condition aggregate information more accurately.*

3.3. **Information choice.** For any subject i who uses an aggregation strategy $\hat{R}^i_C$ in the Active Choice rounds, the optimal choice of signal depends jointly on the signal-prices and i's beliefs about how $\hat{R}^i_C$ interacts with signal choice. The variation in signal prices Δp helps us measure the demand for signals without assuming any structure on $\hat{R}^i_C$. When $\Delta p = 0$, all signal compositions

[9] Recall that subjects were never provided feedback, so this learning can only occur if subjects realize the connection between colors and the sample average introspectively.

are equally expensive. Thus, any subject i from the Active Choice treatment should choose a signal combination that she believes would deliver the most accurate report, given $\hat{R}^i_C$. If they believe that extreme signals might result in more a more inaccurate report, then they would prefer diverse signals. It is important to note that, i's choices are guided by her *beliefs* about which signal choices lead to more accuracy, i.e, the perceived value of information, which might not be equal to the actual value of information. In the range $\Delta p > 0$, as Δp increases, diversification gets increasingly expensive, while its perceived value stays the same. Thus, they should choose diverse signals less often.

Hypothesis 4 (Demand for diversity (Price response))**.** *In the AC treatment, subjects prefer diverse signals over extreme signals for* $\Delta p = 0$*. As* Δp *increases, subjects choose diverse signals less often.*

The sample-average rule imposed in the Average treatment, disregards all information about the unobserved signals. Thus, unless the subjects themselves are using the sample-average rule in the Colors treatment, which is unlikely, the actual value of diversification increases significantly when the sample average rule is imposed.[10] Does the perceived value of diversification react to this change? Our next hypothesis is about how the perceived value of diversification, as measured through the demand for diverse signals, changes when the sample-average rule is imposed:

Hypothesis 5 (Demand for diversity (Avg vs AC))**.** *Compared to the AC treatment, subjects in the Avg treatment are more likely to choose diverse signals over extreme signals for all* $\Delta p > 0$*.*

To summarize, hypotheses 1, 2, and 3 presume that subjects use a naive aggregation rule (for e.g, sample-average or Gambler's fallacy) to predict comparative statics over signal diversity, experience, or informativeness of treatments (C versus NC treatments). Sophisticated aggregation behavior would result in their rejection. Our last two hypotheses (4 and 5) posit sophisticated signal choice favoring diverse signals, and would be rejected if subjects respond only to the signal prices, p_1 and p_2.

[10]Recall that we expect the sample-average rule to be more prevalent in the NoColors treatment, rather than the Colors treatment.

4. Results

In this section, we use all the data from our four main treatments to test our hypotheses. Most of the effects are identified through the variation of the treatments within the same subject. We begin with a reduced-form analysis of the deviation from Bayesian estimates across treatments. Following this, we confirm the basic insights of the reduced form regressions by estimating a structural model of the updating process. The third subsection provides a subject-level analysis and broadly confirms the underlying pattern of the data: there is only limited evidence of systemic polarization in subject reports. The final subsection presents the results on subject willingness to pay for diverse information.

4.1. **Reduced form analysis.**

We address Hypotheses 1, 2 and 3 using the data from all rounds of C1+NC2 and NC1+C2 sessions.[11] For ease of exposition, we define a subject to be inexperienced during rounds 1-20, and experienced when they are in rounds 21-40, having played a different treatment previously in rounds 1-20. To isolate the effects of diverse signals, the observation of colors, and experience, we regress the absolute deviation from the Bayesian report (y) on indicator variables for extreme signals, the C treatment, and Experience, plus all interaction terms, clustering standard errors at the subject level. In equation 4.1, the baseline observations are the NC1 rounds where inexperienced subjects observed diverse signals.

$$\begin{aligned} y =& \beta_0 + \beta_1 \mathbb{1}_{Extreme} + \beta_2 \mathbb{1}_{Color} + \beta_3 \mathbb{1}_{Extreme} \times \mathbb{1}_{Color} + \beta_4 \mathbb{1}_{Experienced} + \beta_5 \mathbb{1}_{Extreme} \times \mathbb{1}_{Experienced} + \\ & \beta_6 \mathbb{1}_{Color} \times \mathbb{1}_{Experienced} + \beta_7 \mathbb{1}_{Extreme} \times \mathbb{1}_{Color} \times \mathbb{1}_{Experienced} + \epsilon \end{aligned} \tag{4.1}$$

[11] In the Appendix, we repeat the analysis while including the Colors and Avg rounds from the fifth treatment, increasing our data set further. We do not include this treatment in the main text because, as discussed in the Appendix, there is evidence of substantial subject confusion in the novel part of the fifth treatment. All our results are robust to including the fifth treatment sessions.

In Table 3, we present the the estimated marginal effects (ME) of each of the three factors of interest: Extreme signals (top left), Experience (top right) and Colors (Bottom), for the 2×2 values taken by the other two factors.[12] For example, the top left table measures the marginal effect of moving from diverse signals to extreme signals, at each level of Colors (along the rows) and Experience (along the columns). It shows that observing extreme signals significantly increases the average absolute reporting error, for both experienced and inexperienced subjects, but *only* in the NC rounds. The size of the effect is 2.08 units ($p < 0.001$) for inexperienced NC subjects, and 1.71 units ($p < 0.01$) for experienced NC subjects. Given the incentive structure of the experiment, this increases expected earnings by approximately \$1.26 or \$1.07 for experienced or inexperienced subjects, respectively. Further, the standard errors indicate that the effects of extreme signals are estimated with high precision. The standard errors, of around half a unit on the 100 point scale used in the experiment, are approximately 28 times smaller than the expected absolute difference between the Bayesian estimate and the Sample Average heuristic when extreme signals are observed.

The bottom left panel of Table 3 shows the marginal effect of observing Colors on the average absolute deviation from Bayesian reports. A statistically significant effect is only found for inexperienced subjects who observe extreme signals: the improvement in reports, when observing extreme signals, for subjects who are participating in the C treatment (relative to those participating in the NC treatment) and have not yet experienced the other treatment is 2.40 units ($p < 0.05$).[13] The effects of Experience are shown in the top right panel. There is an improvement in reports in the NC treatment for subjects who have already had experience in the C treatment (relative to those who play the NC treatment first), but the effects are not significant at the 5% level.

Result 1. (a): Diverse signals improve reports, relative to the Bayesian benchmark, only when color information is not available. (Qualified support for Hypothesis 1.)

[12] The main regression is presented in Table A.1 in the appendix.

[13] That is, this statistic is a between subject measure of the effect of observing the colors of unobserved signals among inexperienced subjects.

ME of Extreme signals	Rounds 1-20	Rounds 21-40
Under Colors	0.75	0.54
	(0.59)	(0.56)
Under NoColors	2.08	1.71
	(0.51)	(0.49)

ME of Experience	Diverse signals	Extreme signals
Under Colors	0.61	0.41
	(0.82)	(1.20)
Under NoColors	-1.48	-1.85
	(0.77)	(1.06)

ME of Colors	Rounds 1-20	Rounds 21-40
Under Diverse signals	-1.06	1.04
	(0.74)	(0.84)
Under Extreme signals	-2.40	-0.14
	(1.11)	(1.14)

Table 3. We report the marginal effects (ME) of three factors: Extreme signals (top left), Experience (top right) and Colors (Bottom) from regression-equation (4.1). The marginal effect of each factor is reported for all 2×2 values taken by the other two factors. For example, the top left table measures the effect of moving from diverse signals to extreme signals, at each level of Colors (along the rows) and Experience (along the columns). All values are calculated from a regression of the absolute deviation from the Bayesian report on indicators for Extreme signals, Colors, and Experience, plus all interaction terms, with standard errors clustered at the subject level. Standard errors are in parenthesis.

(b): Observing colors improves reports, relative to the Bayesian benchmark, only when signals are extreme and subjects are inexperienced. (Qualified support for Hypothesis 2.)
(c): Prior experience with the Colors treatment does not cause a statistically significant improvement in reports in the NoColors treatment. (Fails to support Hypothesis 3.)

Result 1 documented the effects of extreme signals and observing colors on the absolute error of subject reports, but is silent on whether errors are generated by biased reports or are generated by unbiased variance in reports.

We define *polarization towards the extremity* as instances where the report was lower than (Bayesian estimate $-$ 1) when all signals were low, or the report was higher than (Bayesian estimate + 1) when all signals were high.

Conversely, we define *polarization against the extremity* as in the report was higher than (Bayesian estimate + 1) when all signals were low, and lower than (Bayesian estimate − 1) when all signals were high. We allow the ± 1 tolerance band around the Bayesian estimate to allow for inconsistencies between how the computer and subjects rounded fractions, and our results are robust to alternative tolerance bands, for example, ± 1.5 or ± 2. Misreporting towards the extremity is consistent with subjects employing the generalized sample-average ($\alpha R^*_{NC} + (1-\alpha) R^{SA}_{NC}$), and misreporting against the extremity is consistent with the generalized Gambler's fallacy ($\alpha R^*_{NC} + (1-\alpha) R^{GF}_{NC}$).

Under systematic misreporting towards the extremity, subjects would be more likely to over-report under all-high signals. Similarly, subjects would be more likely to under-report under all-low signals. In columns [1] and [2] of Table 4, we report a multinomial probit regression of whether the subjects under or over-reported, on whether the signals were all-high or all-low. We find that extreme signals increase the probability of both under and over reporting, implying noisier reports rather than systematic under/over reporting, although the effect is stronger and significant for high signals causing under reporting (and vice versa). Importantly, there is no significant difference in the proportion of under (or over) reporting when signals are all-high as compared to all-low. If extreme signals caused polarized reports, we would expect the rate of under reporting to be substantially different, and also to differ in sign, when facing all-high as compared to all-low signals.

In the Appendix we include an alternative analysis that focuses on the magnitude, rather than the probability, of misreporting. The conclusions are the same: there is no evidence of systemic *polarization* towards, or away from, the extremity in our sample.

Result 2: In the baseline (non-color rounds), aggregate behavior is inconsistent with systematic misreporting polarized towards the extremity. (Fails to support Hypothesis 1.)

	$\mathbb{1}_{Underreport}$	$\mathbb{1}_{Overreport}$
	[1]	[2]
	NC1,NC2	NC1,NC2
$\mathbb{1}_{Red}$	0.43	0.36
	(.21)	(.20)
$\mathbb{1}_{Blue}$	0.25	0.52
	(.20)	(.22)
Constant	-0.06	-0.07
	(0.16)	(0.16)
N	1180	1180

TABLE 4. Multinomial probit regression of the probability of underreporting (column [1]), correctly reporting (base group, not shown), and overreporting (column [2]) on indicator variables for observing all Red signals or all Blue signals (with diverse signals as the base group). Standard errors clustered at subject level are reported in parentheses. There were 59 clusters.

4.2. **Structural model.** Figure 4.1 plots the histogram of deviations from the Bayesian response across the four choice environments: Colors or No Colors treatment crossed with Diverse or Extreme signals. In all four cases, we find that the deviations are approximately unimodal and symmetrically distributed with a substantial mode at the Bayesian response. As previously reported in Table 3, Color rounds have a lower absolute deviation from Bayes reports, but Figure 4.1 shows that the Colors rounds also show evidence of *polarization* towards the extremity (or sample average behavior) under extreme signals: the distribution of deviations under all-high signals line-up slightly to the left of that of all-low signals.[14] Despite the familiar shape of the distribution, the deviations are not well-approximated by a normal distribution: any normal distribution that fits the mode of the distribution substantially underestimates the tail mass of the guess distribution, and any normal distribution that fits the tails substantially undershoots the mode.

Instead, the data appears to be approximated by the convex combination of two normal distributions, one with very small variance and one with large variance. Such a convolution is not entirely arbitrary, and can be motivated

[14]This and the results in Table 4 reject Hypothesis 2: Contrary to the hypothesis, polarization is only present in the Colors rounds.

from the psychology literature on Type 1 and Type 2 thinking processes. Responses that are tightly clustered around the Bayesian response are likely to be the outcome of careful consideration (Type 2 deliberation) by people who know how to calculate the exact Bayesian formula.[15] Responses that are more widely dispersed are likely to be the outcome of decisions made by people who were unaware of the Bayesian formula or were unable to use it. Such responses are often attributed to quick Type 1 thinking or heuristic-based decision making.[16] For illustrative purposes, for the Color rounds and diverse signals, we plot the $N(0,2)$ and $N(0,7)$ distributions on top of the histogram. The plot is suggestive that the histogram of subject decisions could be presented as a mixture of those two distributions: $N(0,2)$ and $N(0,7)$.

Motivated by the observations above, we model subject reports with the following simple modeling assumptions:

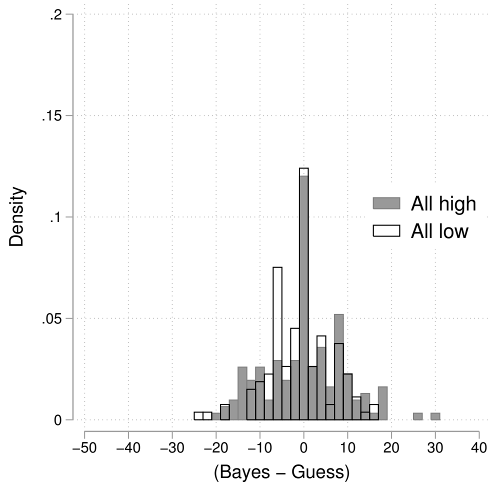
i) In treatment $t \in \{C, NC\}$ and with signal-type $s \in \{e, d\}$, where e stand for extreme and d for diverse, we assume that reports are approximately Bayesian with probability $p_{s,t}$. In this case, the report is generated from the structural equation

$$R_{s,t} = Bayes + \epsilon \tag{4.2}$$

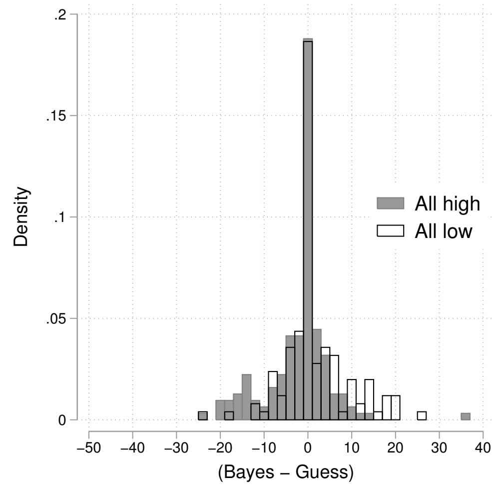
where $\epsilon \in N(0,2)$ is i.i.d noise irrespective of the treatment and signals. Note that we have fixed the variance exogenously at 2 and this implies that, with 95% probability, the reports are within $\pm 2\sqrt{2}$ of the Bayes report. Reducing this exogenous variance, to say 1.5 or 1, does not substantively alter the conclusion of the model. We fix the variance of the error term exogenously, because the model is not identified otherwise. We allow $p_{s,t}$ to depend on the

[15]Relatively few of these responses correspond to the exact Bayesian response, for at least two reasons. First, there appears to be a substantial number of subjects who calculate the Bayesian update while assigning unobserved Blue signals a value of 25, and unobserved Red signals a value of 75. The correct calculations would assign a value of 25.5 and 75.5 to Blue and Red signals, respectively. Second, some subject responses appear to be rounded, and some are reasoned approximations rather than calculations of the Bayesian update.

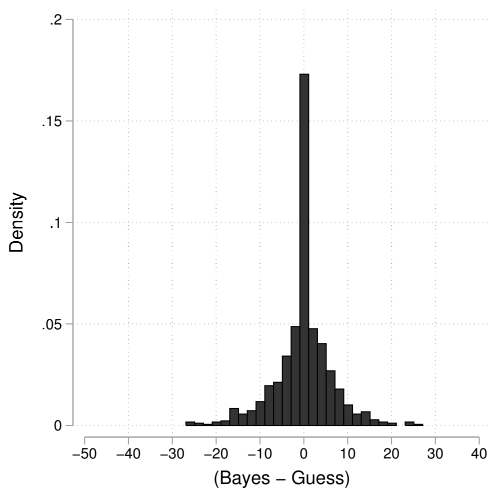
[16]Instead of calculating (or perhaps carefully approximating) the Bayesian update, such Type 1 subjects might make a quick judgment (i.e. estimate the average of the three observed signals, and then arbitrarily adjust up or down depending on the color of unobserved signals) and move on.



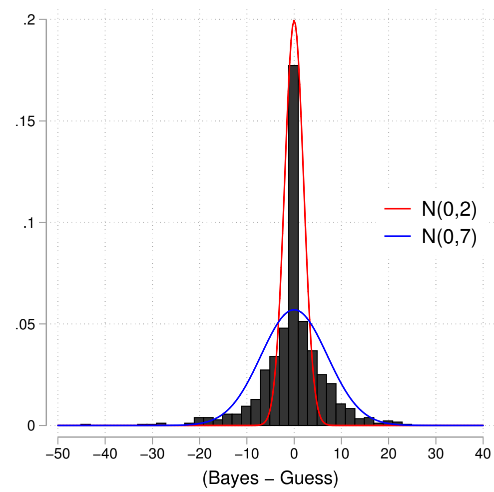
(A) No Color rounds Extreme.



(B) Color rounds Extreme.



(C) No Color rounds Diverse.



(D) Color rounds Diverse.

FIGURE 4.1. Distribution of (Bayes Update-Guess) from all guesses when all three observed signals were high or were low. Bin width =2.

treatment t and signal-type s to accommodate the variation in the frequency of accurate reports across response conditions.

ii) With the complementary probability, $1 - p_{s,t}$, reports are provided by a heuristic decision pattern which is, potentially, biased. Bias could originate from over or under-weighting of the observed signals $\{s_1, s_2, s_3\}$ or from mis-weighting the "base rate" of unobserved signals. In this case, the report is generated from the structural equations

$$R_{s,\ t=C} = \alpha_{s,C}(s_1 + s_2 + s_3) + \beta_{s,C}\left(25.5 \times (\#Blue) + 75.5 \times (\#Red)\right) + \epsilon_{s,C} \tag{4.3}$$

and

$$R_{s,\ t=NC} = \alpha_{s,NC}(s_1 + s_2 + s_3) + \beta_{s,NC}\left(50.5 \times 4\right) + \epsilon_{s,NC} \tag{4.4}$$

for the Color and No Color treatments, respectively, where $\#Blue$ and $\#Red$ are the number of unobserved Blue and Red signals. Bayesian decision making, with noise, corresponds to $\alpha = \beta = 1/7 \simeq .14$ and lower or higher estimated values indicate under or overweighting of either the observed, or unobserved, signals, respectively. Finally, we assume that the noise in reporting is Gaussian $\epsilon_{s,t} \in \mathcal{N}\left(0, {\sigma_{s,t}}^2\right)$ with variance that can depend on the treatment and the diversity or extremity of signals.

We estimate $(p_{s,t}, \sigma_{s,t}, \alpha_{s,t}, \beta_{s,t})_{t\in\{C,NC\},s\in\{e,d\}}$ using Maximum Likelihood Estimation over the observed reports from all the C and NC rounds, and report the results in Table 5.

The estimates help us explicitly model the mental models used by subjects in forming reports. They also clarify and explain our reduced-form results. For example, consider the marginal effect of extreme signals. In Color rounds, extreme signals don't increase absolute deviations. As one explanation, note that the proportion of observations that are consistent with the more accurate Type 2 decision making is slightly, but not statistically significantly, higher in the case of extreme signals than diverse signals (.27 vs .24). On the other hand, the variance of the heuristic Type 1 decision making system is slightly, but not statistically significantly, higher in the case of diverse signals (8.49 vs 7.41). These two small effects are essentially offsetting each other in aggregate.

Next, consider the No Colors rounds. In this case, the proportion of signals that are consistent with the more accurate Type 2 system is substantially larger ($0.19 > 0.05$, p-value .003) in the case of diverse signals. In addition,

		t = No Colors	t = Colors
Extreme signals	$p_{e,t}$	0.05	0.27
		(0.04)	(0.04)
	$\sigma_{e,t}$	8.33	8.41
		(0.39)	(0.45)
	$\alpha_{e,t}$	0.13	0.17
		(0.01)	(0.01)
	$\beta_{e,t}$	0.15	0.12
		(0.01)	(0.01)
Diverse signals	$p_{d,t}$	0.19	0.24
		(.03)	(.03)
	$\sigma_{d,t}$	7.12	7.49
		(.21)	(.23)
	$\alpha_{d,t}$	0.15	0.16
		(0.01)	(0.01)
	$\beta_{d,t}$	0.14	0.13
		(0.01)	(0.01)

TABLE 5. Maximum likelihood estimate of the model parameters from Equations 4.3 and 4.4.

the variance of the Type 1 heuristic system is larger ($8.33 > 7.12$, p-value .007) in the case of extreme signals. These two effects reinforce each other, and generate substantially worse reports in the case of extreme signals, thus confirming the reduced form regression presented in Table 3.

Deviations from the Bayesian update are also explained by subjects marginally over-weighting their signals in the Colors rounds: We estimate α coefficients of .17 from extreme and .16 from diverse signal rounds, which are both marginally higher than the benchmark value of .14 (the differences are statistically significant at conventional levels). Similarly, the estimated β coefficients show that they marginally underweight the base-rate information contained in the unobserved signals. In contrast, in the No Colors rounds, these coefficients are not significantly different from .14, indicating that there is no evidence for over-weighting of signals in the NC rounds.

Result 3: Under both extreme and diverse signals, subjects update with unbiased noise in the NC rounds. In C rounds, reports are slightly polarized under extreme signals, but are less noisy than those from the NC rounds.

4.3. **Subject level bias.** Could the lack of substantial polarization at the aggregate level in the NC rounds be attributed to two different and opposite subject-level biases canceling each other out in aggregate? To dig deeper, we conduct a subject-level analysis. For each subject who faced extreme signals in the No Colors rounds 3 or more times, we calculate how frequently they misreported towards and against the extremity of their observed signals. To allow for rounding errors, reports that are within ± 1 of the true Bayesian report are not accounted as a misreport.[17] In Figure 4.2, we bubble-plot these fractions of misreports that were towards or against the extremity, for these same subjects. There are 9 (out of 54[18]) subjects who misreport in the direction of the sample-average heuristic at least once but never misreport in the direction of the Gambler's fallacy, and 11 subjects who do the opposite. Thus, 9 and 11 subjects are, respectively, consistent with sample-average and Gambler's fallacy, and the remaining 34/54 (64%) subjects do not show a systematic polarization under extreme signals.[19] The data overall is slightly biased towards the lower right of the figure, suggesting that mistakes a la Gambler's fallacy were marginally more prevalent than Sample average rule. 10 subjects showed no bias in either direction whatsoever.

When we recreate the same figure for the Colors rounds, the data shows a distinct shift, which is consistent with the above result that $\alpha_{e,t} > 0.14$, towards the top left of the figure. There are 11 (out of 54) subjects who misreport in the direction of the sample-average heuristic at least once but never misreport in the direction of the Gambler's fallacy, and 8 subjects who do the opposite. 10 subjects showed no bias in either direction whatsoever.

Result 4: We find little evidence of systemic polarization at the subject level.

[17] Our analysis is robust to using an alternative tolerance level of $\pm .5$.

[18] There were 54 subjects who faced extreme signals in the No Colors rounds 3 or more times.

[19] Alternative classification procedures lead to similar conclusions. For example, we could classify a subject as exhibiting the Gambler's fallacy if 80% of their choices are polarized in the direction of the fallacy (and similarly for the sample-average heuristic). In this case, we would classify 8 subjects as exhibiting the Gambler's fallacy, 6 subjects as exhibiting the sample-average heuristic, and 40 subjects as being unpolarized.

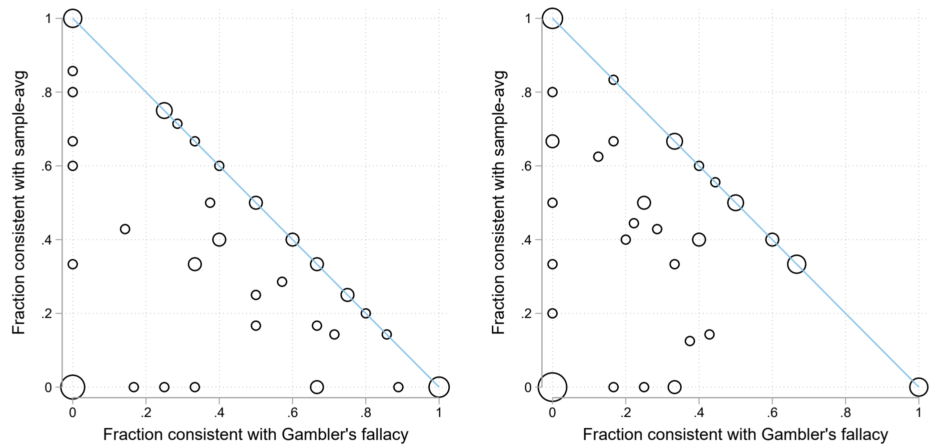


FIGURE 4.2. Proportion of choices, by subject, that are consistent with either the gambler's fallacy or the sample average bias for extreme signals in the NC (left) and C (right) treatments. Restricted to subjects who observed extreme signals at least three times. The size of each bubble represents the number of subjects at each point.

4.4. **Demand for signals.** In the Active Choice and Average treatments, signals not only influence the precision of reports but also pay subjects directly: subjects receive p_1 and p_2 for each signal that is of same or opposite polarity to s_1, respectively. In this subsection, we analyze whether subjects exhibit a preference for diverse signals (or, perhaps, a preference for extreme signals) or whether they prefer to only maximize the signal acquisition payoffs.

An unfortunate programming constraint, which was not noticed until after the experiments were run, allows us to only observe the signals finally received by the subjects and not the chosen or requested signals. The received signals can differ from the requested signals when, for example, the subject requests two Blue signals but only one Blue signal is available in $\{s_2, s_3, .., s_7\}$. In the rounds where $\{s_2, s_3, .., s_7\}$ contained at least two signals of either color, the signals requested and received are guaranteed to be identical. We use only data from these rounds for our following results on signal choice. Given subjects made their signal choice without knowing the composition of $\{s_2, s_3, ..s_7\}$, conditioning on the composition does not bias the analysis of the signal choice.[20] For comparison, we repeat the analysis with the full data set, containing all rounds, in Appendix A. The results are similar.

In Figure 4.3, we plot the proportion of extreme signal choices against Δp, the difference between the price of own signal and the price of the other signal. We do this separately for when subjects aggregate on their own (Active Choice rounds) versus when they aggregate under sample-average rule (Average rounds), pooling similar price differences to simplify the figure.

When $\Delta p = 0$, approximately 10% of signal choices are extreme as most subjects prefer a diverse portfolio of signals. As the price-difference increases, subjects become more likely to choose extreme signals. In the Active Choice treatment, even at the highest price-difference group ($\Delta p \in \{14, 20\}$), approximately 20% of decisions are in favor of diverse signals and this figure is higher still in the Average treatment (approximately 35%).

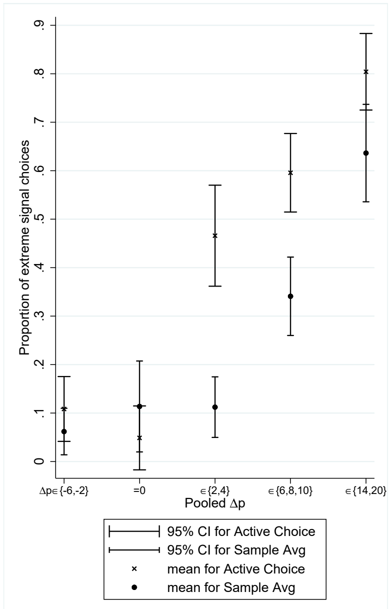
Result 5: Subjects rarely choose extreme signals when $\Delta p = 0$. (Support for Hypothesis 4.)

Next we ask if subjects were aware that diverse signals are more "valuable" in the Average treatment, and if it was reflected in their demand for signals (Hypothesis 5). Note that this hypothesis relies on the implicit assumption that diverse signals are actually more valuable in the Avg treatment than in the AC treatment: an assumption that is testable in our data. As we establish in Appendix A.3,not only is the assumption supported on average, across the subject population, we find that that it holds individually for all subjects.

As already seen in Figure 4.3, at every $\Delta p > 0$, subjects are less likely to choose extreme signals under Average than under Active Choice. Table 4.3 presents two regressions designed to study the determinants of extreme signal choices.[21] In column [1], which controls for Δp and the treatment (Average or Active Choice) we observe a negative coefficient on the dummy for Average.

[20] $\{s_2, s_3, ..s_7\}$ are drawn independently of s_1, and hence s_1 is uninformative about the other 6 signals.

[21] In the Appendix, we report the same regressions while including the Colors and Avg rounds from the fifth treatment, increasing our data set further. We do not include this treatment in the main text because, as discussed in the Appendix, there is evidence of



Probit (Active choice + Avg)		
	Dependent variable	
	1 if signal choice was Extreme, 0 otherwise	
	[1]	[2]
Δp	0.10	0.10
	(0.02)	(0.02)
$\Delta p \times \mathbb{1}_{SmplAvg}$	-0.00	-0.01
	(0.01)	(0.02)
$\mathbb{1}_{SmplAvg}$	-0.61	-0.92
	(0.20)	(0.28)
Dev		-0.11
		(0.03)
Dev$\times \mathbb{1}_{SmplAvg}$		0.08
		(0.04)
Constant	-0.65	-0.19
	(0.14)	(0.19)
N	909	909

FIGURE 4.3. The left hand panel plots the proportion of extreme signal choices against Δp, the difference between the price of an own colored signal and an other colored signal, separately for when subjects update on their own (Active Choice rounds) versus when they update under sample-average rule (Average rounds) with 95% confidence intervals. The right hand panel presents a probit regression of extreme choice on Δp, with a dummy variable for the Average rounds with standard errors clustered at the subject level (30 subjects). In regression [2], we additionally control for "Dev", which is calculated at the subject level as the average deviation from the Bayesian update across the first 20 Active Choice rounds and restricted to rounds with diverse signals. Both panels restrict the data to only include rounds where there were at least two red signals and at least two blue signals in $\{s_2, s_3, s_4, s_5, s_6, s_7\}$

In column [2] we add a control, Dev, which captures subject-level guess accuracy in the Active Choice rounds when the subject observed diverse signals.

substantial subject confusion in the novel part of the extra treatment. All our results are robust to including the fifth treatment sessions.

Thus, the higher the value of Dev, the worse was the quality of information aggregation by the subject.[22]

From column [2] of Table 4.3 we conclude, given the negative coefficient on Dev, that subjects who are worse at aggregating information are more likely to select diverse signals in the AC treatment. To provide some context for the estimated value of -0.11, in the Active Choice treatment, at the sample average $\Delta(p)$ and average value of Dev, a one unit improvement in signal aggregation ability leads to a 4 percentage point decrease in the likelihood of choosing extreme signals. That is, subjects who are worse at aggregating exhibit some self-awareness of this and respond by giving themselves an easier updating problem. For the Avg treatment, the effect of Dev is $-0.11 + 0.08 = 0.03$ and statistically insignificant, suggesting that the choice of signals is independent of aggregating ability in the Avg treatment. This forms our final result, and suggests that subjects are able to separate information acquisition from information processing. The estimates of $\Delta(p)$ and Dev in Table 4.3 are also rather precise, with standard errors of only 0.02 and 0.03, respectively.

Result 6: Subjects choose extreme signals less frequently (i) in the Average treatment and (ii) when they are poor at aggregating information in the Active Choice treatment. (Support for Hypothesis 5.)

Result 7: Signal choices in the Average treatment are independent of guess accuracy in the Active Choice treatment.

Our results suggest, overall, that subjects were aware that extreme signals might reduce the quality of reports, and they were therefore willing to pay a cost to buy diverse signals. The results also suggest a mechanism that led to this preference for diversity. In the AC rounds, subjects observed s_1 by default, and then chose to observe the realizations of two Blue signals, or two Red signals, or one signal of each color. Let n be the number of signals chosen that are of the same color as s_1. Thus, $n = 2$ is the only choice that results in extreme signals. When $\Delta p = 0$, the two ways of forming diverse signals, $n = 0$

[22]If the Dev variable was calculated using all rounds, then there would be a potential endogeneity problem: some subjects might choose diverse signals more often and if diverse signals have lower deviations, that would systematically effect the total deviation across all observations. We checked the robustness of column [2] by recalculating Dev using only rounds with extreme signals, finding that the results are qualitatively unchanged.

and $n = 1$ were equally affordable, and empirically resulted in approximately similar precision in the Colors treatment when signals were randomly assigned. In the Active Choice (AC) rounds, subjects chose the former in 84% of all $\Delta p = 0$ rounds, and the latter in only 9% of all $\Delta p = 0$ rounds. This feature of the data hints that, perhaps, not all signal-combinations were perceived to be equally valuable by subjects, and hints at their signal-choices being guided by a diversity-seeking heuristic.

5. Conclusion

In this paper we study the value of and the demand for diverse information sources in a simple decision environment where information-processing does not require contingent reasoning. We find that most subjects are unable to make perfect Bayesian updates but instead utilize heuristics that are remarkably resistant to making mistakes when receiving news with transparent bias. Heuristics used for aggregation are not systematically polarized even when signals are extreme. We find little evidence for subjects following a naive sample-average rule or committing the Gambler's fallacy, or over/ under-weighting any of the available information. Previous research (e.g. Enke, 2020) has identified conditions under which information polarization can lead to ex-post polarization. Our results, conversely, demonstrate that when information polarization is transparent, and motivated reasoning is not present, subjects are surprisingly good at constructing a balanced portfolio of signals and then constructing unpolarized estimates of the true state of the world.

Further, our results show that when it comes to instrumental information, few fail to appreciate the value of diverse information sources. The resultant demand for diverse information reacts "rationally" to the value and cost of diverse information, increasing in the Avg rounds when extreme signals lead to lower quality reports. Remarkably, subjects who perform poorly when aggregating information appear to be cognizant of their limitations and exhibit a stronger demand for diversified information in the AC rounds. Finally, when we exogenously impose a naive sample-average aggregation rule the subject level demand for diversified information sources is, rationally, not dependent on subject level aggregation ability. Our results advocate for subsidizing and simplifying the acquisition of diverse news sources, for example through ensuring greater transparency in media bias, so that individuals can choose the right portfolio of information and make better choices.

References

Sandro Ambuehl and Shengwu Li. Belief updating and the demand for information. *Games and Economic Behavior*, 109:21–39, 2018.

Felipe A Araujo, Stephanie W Wang, and Alistair J Wilson. The times they are a-changing: experimenting with dynamic adverse selection. *American Economic Journal: Microeconomics*, 13(4):1–22, 2021.

Daniel J Benjamin. Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:69–186, 2019.

Mary A. Burke and Michael Manz. Economic literacy and inflation expectations: Evidence from a laboratory experiment. *Journal of Money, Credit and Banking*, 46(7):1421–1456, 2014.

Francesco Capozza, Ingar Haaland, Christopher Roth, and Johannes Wohlfart. Studying information acquisition in the field: A practical guide and review. *Available at SSRN 3952561*, 2021.

Gary Charness, Ryan Oprea, and Sevgi Yuksel. How do people choose between biased information sources? evidence from a laboratory experiment. *Journal of the European Economic Association*, 19(3):1656–1691, 2021.

Daniel L Chen, Tobias J Moskowitz, and Kelly Shue. Decision making under the gambler's fallacy: Evidence from asylum judges, loan officers, and baseball umpires. *The Quarterly Journal of Economics*, 131(3):1181–1242, 2016a.

Daniel L. Chen, Martin Schonger, and Chris Wickens. oTree – an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97, 2016b.

John Duffy, Ed Hopkins, Tatiana Kornienko, and Mingye Ma. Information choice in a social learning experiment. *Games and Economic Behavior*, 118: 295–315, 2019.

John Duffy, Ed Hopkins, and Tatiana Kornienko. Lone wolf or herd animal? information choice and learning from others. *European Economic Review*, 134:103690, 2021.

Benjamin Enke. What you see is all there is. *The Quarterly Journal of Economics*, 135(3):1363–1398, 2020.

Ignacio Esponda and Emanuel Vespa. Contingent thinking and the sure-thing principle: Revisiting classic anomalies in the laboratory. Technical report, Working paper, 2019.

Leon Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1962.

Klaus Fiedler and Peter Juslin. Taking the interface between mind and environment seriously. *Information sampling and adaptive cognition*, pages 3–29, 2006.

Andreas Fuster, Ricardo Perez-Truglia, Mirko Wiederholt, and Basit Zafar. Expectations with endogenous information acquisition: An experimental investigation. *Review of Eocnomics and Statistics*, Forthcoming.

Ananda Ganguly and Joshua Tasoff. Fantasy and dread: The demand for information and the consumption utility of the future. *Management Science*, 63(12):4037–4060, 2017.

Romain Gauriot, Lionel Page, and John Wooders. Nash at wimbledon: evidence from half a million serves. *Available at SSRN 2850919*, 2016.

Matthew Gentzkow and Jesse M Shapiro. Media bias and reputation. *Journal of political Economy*, 114(2):280–316, 2006.

Ben Greiner. Subject pool recruitment procedures: Organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125, July 2015.

Daniel Kahneman and Amos Tversky. On the psychology of prediction. *Psychological review*, 80(4):237, 1973.

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

Pierre Simon Marquis de Laplace. *Essai philosophique sur les probabilités*. Bachelier, 1840.

Yusufcan Masatlioglu, A Yesim Orhun, and Collin Raymond. Intrinsic information preferences and skewness. *Ross School of Business Paper*, 2017.

Heiner Mikosch, Christopher Roth, Samad Sarferaz, and Johannes Wohlfart. Uncertainty and information acquisition: Evidence from firms and households. *Available at SSRN*, 2021.

Giovanni Montanari and Salvatore Nunnari. Audi alteram partem: An experiment on selective exposure to information. *Unpublished manuscript*, 2019.

Kirby Nielsen. Preferences for the resolution of uncertainty and the timing of information. *Journal of Economic Theory*, 189:105090, 2020.

Marina Niessner and Eric C So. Bad news bearers: The negative tilt of the financial press. *Available at SSRN 3219831*, 2018.

Richard E Nisbett and Lee Ross. Human inference: Strategies and shortcomings of social judgment. 1980.

An T Oskarsson, Leaf Van Boven, Gary H McClelland, and Reid Hastie. What's next? judging sequences of binary events. *Psychological bulletin*, 135(2):262, 2009.

Matthew Rabin. Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3):775–816, 2002.

Amnon Rapoport and David V Budescu. Randomization in individual choice behavior. *Psychological Review*, 104(3):603, 1997.

Pëllumb Reshidi, Alessandro Lizzeri, Leeat Yariv, Jimmy H Chan, and Wing Suen. Individual and collective information acquisition: An experimental study. Technical report, National Bureau of Economic Research, 2021.

Christopher Roth, Sonja Settele, and Johannes Wohlfart. Risk exposure and acquisition of macroeconomic information. *American Economic Review: Insights*, 4(1):34–53, 2022.

Amos Tversky and Daniel Kahneman. Belief in the law of small numbers. *Psychological bulletin*, 76(2):105, 1971.

Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, 185(4157):1124–1131, 1974.

Florian Zimmermann. Clumped or piecewise? evidence on preferences for information. *Management Science*, 61(4):740–753, 2015.

Appendix A. Appendix [For Online Publication Only]

A.1. **Regression table of Equation 4.1.**

	Absolute deviation from Bayesian report
$\mathbb{1}_{Extreme}$	2.08
	(0.51)
$\mathbb{1}_{Colors}$	-1.06
	(0.74)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Color}$	-1.33
	(0.78)
$\mathbb{1}_{Experienced}$	-1.48
	(0.77)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Experienced}$	-0.37
	(0.71)
$\mathbb{1}_{Color} \times \mathbb{1}_{Experienced}$	2.10
	(1.47)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Colors} \times \mathbb{1}_{Experienced}$	0.16
	(1.19)
Constant	4.87
	(0.50)
N	2360

Table A.1. The effects of Extreme signals, the observability of Colors, and subject experience, on the absolute deviation of subject reports from the Bayesian report. The omitted category is observations from the No Colors rounds 1-20 where subjects observed diverse signals. Standard errors clustered at subject level are reported in parentheses (59 clusters). Data includes all rounds of C1+NC2 and NC1+C2 sessions.

A.2. **Magnitude of bias.** In the main text, we evaluate polarization in reports by evaluating the probability that a subject over or under reports as a function of observing all high or all low signals. Here, we provide a robustness check by examining the magnitude of polarization as a function of observing all high or all low signal in Table A.2. The first column of Table A.2 presents a restricted version of the regression contained in Table A.1, while the second and third tables estimate the polarization of reports. The second column uses only rounds where the sample average heuristic lies above the Bayesian estimate, and the third column uses rounds where the sample average heuristic lies below the Bayesian estimate. If reports are polarized, for either extreme or diverse signal observations, then either the constant or the coefficient on Extreme must be different from zero. As the Table shows, all coefficients in both regressions are close to zero and not statistically significant, indicating that there is no evidence of polarization in our sample. Note that the second and third column use only data from the NC treatment, given that the sample average heuristic is unnatural in the C treatment.

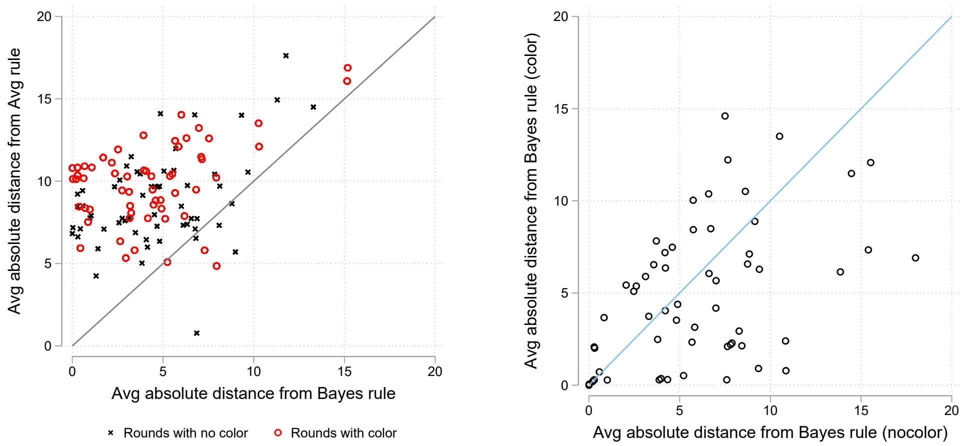
Sample	$\lvert deviation \rvert$ Full C+NC	$deviation$ avg>Bayes NC only	$deviation$ avg<Bayes NC only
$\mathbb{1}_{Extreme}$	1.83	0.53	-0.78
	(0.37)	(1.00)	(0.86)
$\mathbb{1}_{Colors}$	-0.07		
	(0.30)		
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Colors}$	-1.19		
	(0.48)		
Constant	4.21	-0.24	0.06
	(0.39)	(0.46)	(0.37)
N	2360	615	565

TABLE A.2. The first column regresses the absolute deviation of subject reports from the Bayesian benchmark on an indicator for Extreme signals and a Color treatment indicator, using the full sample (all rounds of C1+NC2 and NC1+C2 sessions). The second and third columns regress the deviation of subject reports from the Bayesian benchmark on an indicator for Extreme signals, using only data from the NC treatment, using samples restricted to upwards and downwards polarized signals. Standard errors clustered at subject level are reported in parentheses (59 clusters).

We also calculate, at the subject level, the average absolute deviation from the Bayesian report and the naive sample-average report across each of the Colors and No Colors rounds, and then plot this data in Figure A.1a. We plot the average absolute distance from the Bayesian report on the x-axis, and the distance from the sample-average report on the y-axis. Subjects who are positioned above the 45-degree line are closer to the Bayesian average, and subjects below the 45-degree are closer to the sample-average report, with distance from the 45-degree line giving an indication of the size of the advantage of one rule over the other. It is immediately visually apparent that (i) most subjects are above the 45-degree line and, therefore, on average, closer to the Bayesian report than the sample-average report (ii) there is no strong relationship between the observability of Colors and average deviations from either rule. In fact, more than 60% of subjects were, on average, within 5 points of the Bayesian update for both treatments.

In figure A.1b we drill further down into the distinction between the C and NC treatments. In this panel, we plot the subject-level average absolute deviation from Bayes rule in the No Colors and Colors rounds on the x and y axis, respectively. Here, subjects above the 45-degree line provide better reports in the NC treatment, and subjects below the 45-degree line provide better reports in the C treatment, relative to the Bayesian optimal report.
Reports are substantially closer to the Bayesian paradigm than to the naive sample-average paradigm, for both the C and NC treatments.

A.3. **Relative value of diverse signals across Active Choice and Average treatments.** In this subsection, we establish that, for every subject in our sample, the improvement in guess accuracy from diverse signals, relative to extreme signals, is larger in the Active Choice treatment than the Average treatment. For each subject we calculate the average absolute difference of their guesses from the true $\bar{s}$ for all Active choice rounds, separately for diverse and extreme signals, and interpret the difference between these two measures as the loss from choosing extreme signals. In Figure A.2 we plot the CDF of the subject specific losses from extreme signals. It is clear from the figure that, while the aggregation rule used by the median subject experiences essentially no gain from diversity, there is substantial heterogeneity across subjects. We also calculate and plot the average loss from extreme signals in the Average rounds (under the sample-average aggregation) as the vertical line at, approximately, 8.4. Despite the heterogeneity, the gain from diverse signals in the AC treatment is less than 8.4 units for all subjects.
Observation: Diverse signals improve reports (relative to $\bar{s}$) in the Average rounds more than they do in the Active Choice rounds.



(A) Average absolute deviation from the Sample Average rule plotted against the average absolute deviation from the Bayesian optimal report, at the subject level, for the C treatment (red circles) and the NC treatment (black crosses).

(B) Average absolute deviation from the Bayesian optimal report in the C treatment plotted against the average absolute deviation from the Bayesian optimal report in the NC treatment, at the subject level, when signals were extreme.

FIGURE A.1. Heterogeneity

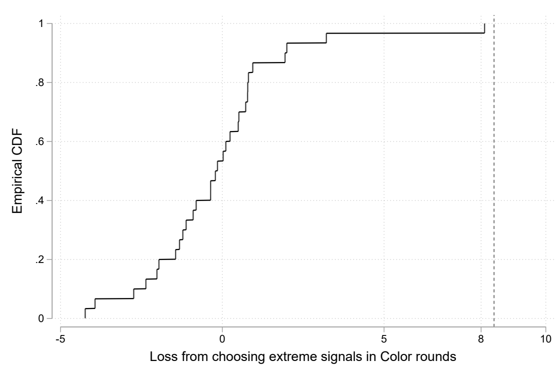
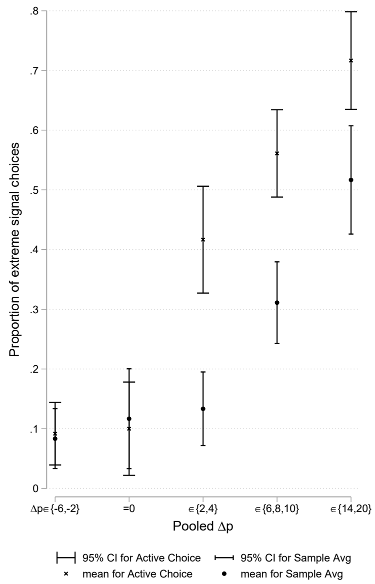


FIGURE A.2. CDF of the average gain (at the subject level) from diversification over all Color rounds. The vertical dashed line denotes the average gain from diversification observed in the Average rounds.



Probit (Active choice + Avg)		
	Dependent variable	
	1 if signal choice was Extreme, 0 otherwise	
	[1]	[2]
Δp	0.09	0.09
	(0.01)	(0.01)
$\Delta p \times \mathbb{1}_{SmplAvg}$	-0.01	-0.01
	(0.01)	(0.01)
$\mathbb{1}_{SmplAvg}$	-0.45	-0.79
	(0.16)	(0.24)
Dev		-0.08
		(.03)
Dev$\times \mathbb{1}_{SmplAvg}$		0.08
		(0.04)
Constant	-0.71	-0.32
	(0.12)	(0.17)
N	1200	1200

FIGURE A.3. Robustness check for the information selection results from Figure 4.3, including data from all rounds.

A.4. **Robustness of the information selection results.** As described in the main text, our data only allows us to observe the signals received by the subjects (and not the signals requested by the subjects). In the main text we restrict attention to a subset of rounds for which we know that requested and received signals must be the same. Here, we provide a robustness test by including data from all rounds. Figure A.3 is a robustness check on Figure 4.3.

A.5. **The fifth treatment.** In a fifth treatment, we gave subjects the opportunity to construct an algorithm that would calculate the subject's report of $\bar{s}$ automatically given the signals and colors that the subject observed. Despite our attempts to design an interface that would be intuitive and easy for subjects to understand, the algorithms that subjects constructed demonstrated that subjects did not understand the algorithm construction process sufficiently. Sessions that contained the algorithm treatment consisted of 20 rounds of the Colors treatment, followed by 10 rounds of the algorithm treatment, followed

	Absolute deviation from Bayesian report
$\mathbb{1}_{Extreme}$	2.08
	(0.51)
$\mathbb{1}_{Color}$	-0.53
	(0.79)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Colors}$	-1.49
	(0.70)
$\mathbb{1}_{Experienced}$	-1.48
	(0.77)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Experienced}$	-0.37
	(0.71)
$\mathbb{1}_{Colors} \times \mathbb{1}_{Experienced}$	1.57
	(1.43)
$\mathbb{1}_{Extreme} \times \mathbb{1}_{Colors} \times \mathbb{1}_{Experienced}$	0.32
	(1.12)
Constant	4.87
	(0.50)
N	2900

Table A.3. (Robustness check on Table A.1 when we add the Colors rounds from the fifth treatment.) The effects of Extreme signals, the observability of colors, and subject experience, on the absolute deviation of subject reports from the Bayesian report. The omitted category is observations from the No Colors rounds 1-20 where subjects observed diverse signals. Standard errors clustered at subject level, including subjects from the Algorithm sessions, are reported in parentheses (86 clusters).

by 10 rounds of the Sample Average treatment. In the remainder of this subsection, we repeat some of the analysis from the main text with the inclusion of data from Colors and Sample Average treatments in these sessions. The results are similar to those presented in the main text.

Table A.3 is a robustness check on Table A.1, and Table A.4 is a robustness check on the right hand panel of Figure 4.3.

Probit (Active choice + Avg)		
	Dependent variable	
	1 if signal choice was Extreme, 0 otherwise	
	[1]	[2]
Δp	0.10	0.10
	(0.02)	(0.02)
$\Delta p \times \mathbb{1}_{SmplAvg}$	-0.00	-0.01
	(0.01)	(0.02)
$\mathbb{1}_{SmplAvg}$	-0.55	-1.03
	(0.17)	(0.23)
Dev		-0.12
		(0.03)
Dev$\times \mathbb{1}_{SmplAvg}$		0.13
		(0.04)
Constant	-0.65	-0.13
	(0.14)	(0.19)
N	1116	1116

TABLE A.4. (Robustness check on the right hand panel of Figure 4.3 when we add the AC rounds from the fifth treatment.) Probit regression of extreme choice on Δp, with a dummy variable for the Average rounds with standard errors clustered at the subject level, including subjects from the Algorithm sessions (57 subjects). In regression [2], we additionally control for "Dev", which is calculated at the subject level as the average deviation from the Bayes rule across the first 20 Active Choice rounds.

Instructions Appendix for The Value of and Demand for Diverse News Sources

Evan M. Calford*and Anujit Chakraborty†

November 1, 2022

Abstract

This document contains the instructions for the paper "The Value of and Demand for Diverse News Sources."

*Research School of Economics, Australian National University. email: Evan.Calford@anu.edu.au
†Department of Economics, University of California, Davis. email: chakraborty@ucdavis.edu

A Instructions for the Colors and NoColors treatment

Immediately following this page are the Instructions for the Colors and NoColors treatment. There were two types of sessions that contained these two treatments. In the C1NC2 sessions the Colors treatment was followed by the NoColors treatment. In the NC1C2 sessions the NoColors treatment was followed by the Colors treatment.

Text common to both treatments is shown in black. Text specific to the NC1N2 sessions is shown in Red. Text specific to the C1NC2 sessions is shown in Blue. Figures 1 and 2 are shown as displayed in the NC1C2 sessions. For the C1NC2 sessions the order of the figures was reversed.

Instructions

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make considered decisions you can earn a considerable amount of money, which will be paid to you in cash at the end of the experiment.

Your computer screen will display useful information. Remember that the information on your computer screen is private. Please do not communicate with the other participants at any point during the experiment. If you have any questions, or need assistance of any kind, raise your hand and the experimenter will come and help you.

Please switch your phones off and place them away. The only materials you will need for this experiment are the computer and the calculator in front of you. We will also provide you with some paper if you wish to take notes.

In the experiment you will make many sets of decisions; each set of decisions will be called a *round*. At the end of the experiment you will be paid for one, randomly selected, round. Each round has an equal chance of being chosen as the round that you would be paid for, so, you should treat each round with equal importance. As you proceed through the experiment your potential earnings will be displayed in **points**. At the end of the experiment we will convert points to dollars at an exchange rate of 1 point = $0.07. You will also be paid a $5.00 show up fee in addition to any money you earn during the experiment.

The basic idea

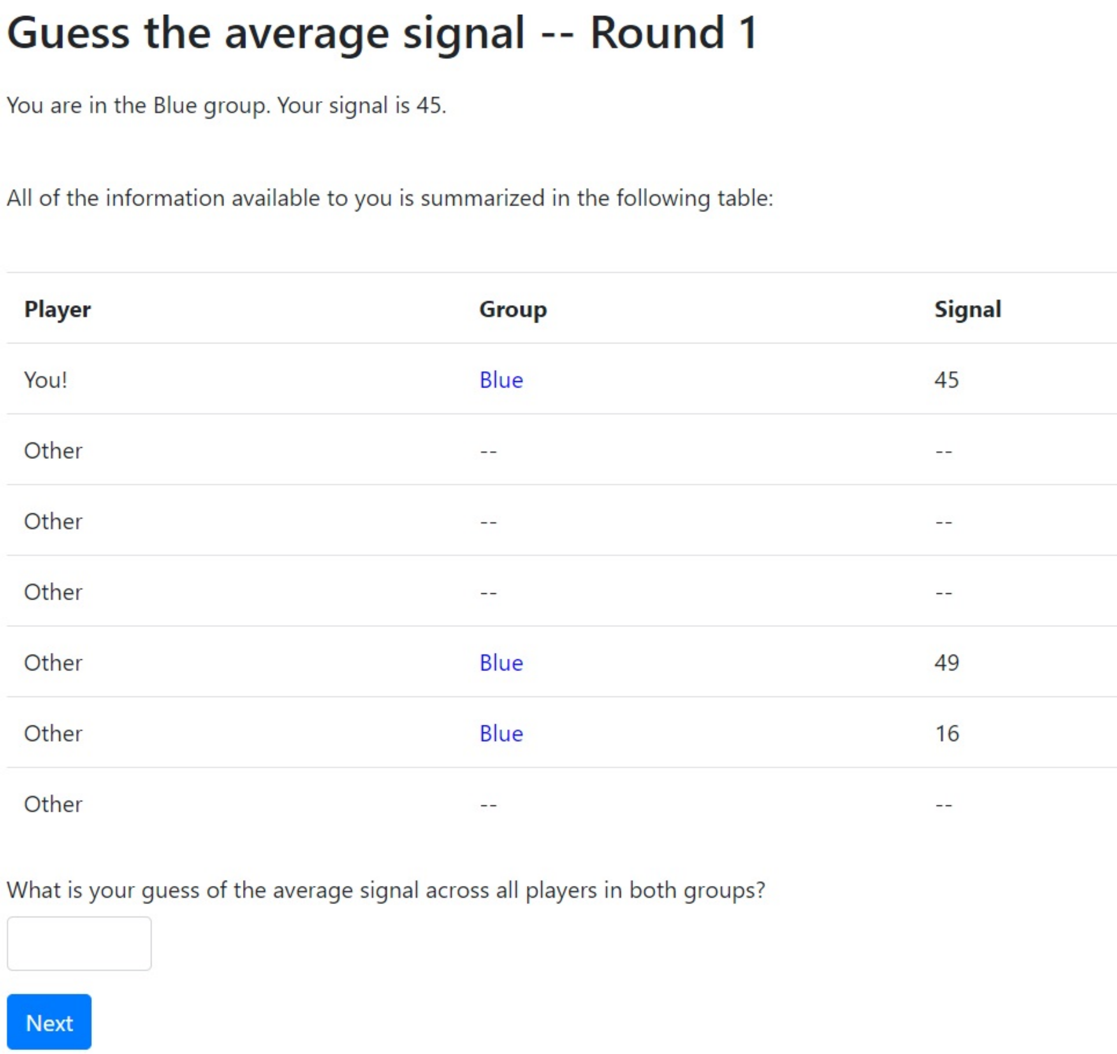
In each round, you will be placed in a set of 7 players.[1] Each of the seven of you will receive a private signal, that will be a randomly drawn integer between 1 and 100. You will have to guess the average of these seven numbers. In order to make this decision you will be able to see the signals that two other those linked players received. Therefore, you will know three of the seven signals in your set when you guess the average of those seven signals. In any round, the more accurate your guess is the more points you can expect to earn.

Groups

Within your set of 7 players, you will be formed into two groups: a **red group** and a **blue group**. If a player has a signal that lies between 1 and 50, then they will be placed in the blue group. If a player has a signal between 51 and 100, then they will be placed in the red group. Thus, a player with a high signal is in the red group, and a player with a low signal is in the blue group.

Guessing the average signal

[1] Note that the other 'players' in Part 1 do not make any decisions that affect you. Because of this, and in order to speed up the experiment, the players are computers.


Guess the average signal -- Round 1

You are in the Blue group. Your signal is 45.

All of the information available to you is summarized in the following table:

Player	Group	Signal
You!	Blue	45
Other	--	--
Other	--	--
Other	--	--
Other	Blue	49
Other	Blue	16
Other	--	--

What is your guess of the average signal across all players in both groups?

Next

Figure 1: Screen shot of the guess page, first 20 rounds

For the first 20 rounds you will be able to see your own signal and color and the signals and colors of two other players, as displayed in Figure 1.

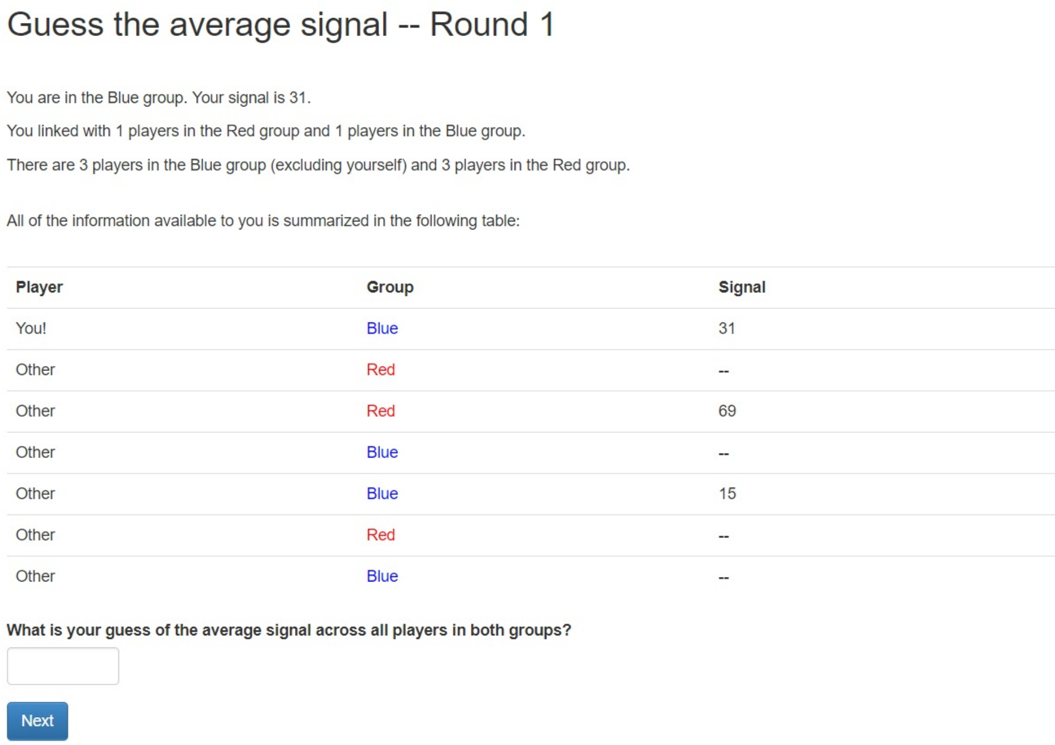
For the first 20 rounds you will be able to see your own signal and color and the signals and colors of two other players and the colors of all players, as displayed in Figure 1.

From round 21 to 40 you will be able to see your own signal and color and the signals and colors of two other players and the colors of all players, as displayed in Figure 2.

From round 21 to 40 you will be able to see your own signal and color and the signals and colors of two other players, as displayed in Figure 2.

For each round there is a large prize, worth 200 points, and a small prize, worth 50 points. **The more accurate your guess is, the higher are the chances of you winning the large prize.**

If your guess is exactly accurate then you will win the large prize with certainty. If your guess is more than $16\frac{2}{3}$ points away from the true average you will earn the small prize. If


Guess the average signal -- Round 1

You are in the Blue group. Your signal is 31.

You linked with 1 players in the Red group and 1 players in the Blue group.

There are 3 players in the Blue group (excluding yourself) and 3 players in the Red group.

All of the information available to you is summarized in the following table:

Player	Group	Signal
You!	Blue	31
Other	Red	--
Other	Red	69
Other	Blue	--
Other	Blue	15
Other	Red	--
Other	Blue	--

What is your guess of the average signal across all players in both groups?

Next

Figure 2: Screen shot of the guess page, rounds 21 to 40

your guess is between 0 and $16\frac{2}{3}$ of the true average, the probability that you will win the large prize is given by following formula:

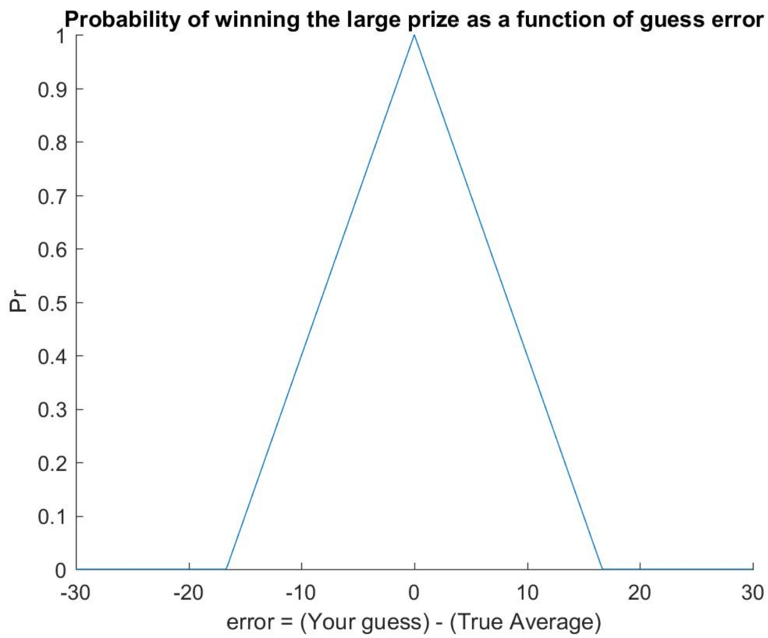
$$1 - \frac{6 * \text{error}}{100}$$

where "error" is the difference between your guess and the true average. The probability of winning the large prize is equivalently displayed in the graph below. Notice that improving your guess accuracy by 1 unit increases your chances of winning the large prize (worth 200 points) by 6 percentage points, and improving your accuracy by 5 units increases your chances of winning the large prize by 30 percentage points.

Payment

Determining which round will be paid

When you arrived, you were given a sealed envelope. DO NOT OPEN THE ENVELOPE UNTIL YOU ARE INSTRUCTED TO. The envelope contains a randomly generated number between 1 and 40, with each number having an equal chance of being in the envelope. The number that is inside the envelope will determine which of the 40 rounds you will be paid for.



Calculating your payment

You will receive the points you earned from your linking choice in the appropriate round plus either the large or small prize.

Your guess in the appropriate round will determine your probability of winning the large prize. After you have completed all 40 rounds, a randomly generated number (uniformly distributed between 0% and 100%) will appear on your screen. You will win the large prize if your percentage chance of winning is larger than the randomly drawn number, otherwise you will win the small prize. This procedure ensures that you will win the large prize with the correct probability.

Survey

At the end of the experiment there will be a short demographics survey. Please fill this in accurately. If you would prefer not to answer any of the questions you may do so, and there will be no penalty for not filling it in.

Frequently Asked Questions

Q1. Is this some kind of psychology experiment with an agenda you haven't told us?

Answer: No. It is an economics experiment. If we do anything deceptive, or don't pay you cash as described, then you can complain to the campus Human Subjects Committee and we will be in serious trouble. These instructions are meant to clarify how you earn money in the experiment, and our interest is in seeing how people make economic decisions.

Q2. Do I have to enter integer (whole) numbers for my guesses?

Answer: No, you may use decimal places if you wish, but you cannot use fractions.

Q3. In part 2, does my guess of my counterpart's linking choice affect the signals I see on the next screen?

Answer: No, the signals you see on the next screen are exactly the same as what your counterpart saw when making their original decision, and are only determined by your counterpart's linking choice.

B Instructions for the Active Choice and Sample Average treatments

Immediately following this page are the Instructions for the Active Choice and Sample Average treatments.

The instructions handed out to subjects at the beginning of the session are reproduced in black text. After 20 rounds, subjects saw a short notification on their screens informing them of a change in procedures. This notification is included in the following as red text enclosed within brackets.

Instructions

You are about to participate in an experiment in the economics of decision-making. If you follow these instructions carefully and make considered decisions you can earn a considerable amount of money, which will be paid to you in cash at the end of the experiment.

Your computer screen will display useful information. Remember that the information on your computer screen is private. Please do not communicate with the other participants at any point during the experiment. If you have any questions, or need assistance of any kind, raise your hand and the experimenter will come and help you.

Please switch your phones off and place them away. The only materials you will need for this experiment are the computer and the calculator in front of you. We will also provide you with some paper if you wish to take notes.

In the experiment you will make many sets of decisions; each set of decisions will be called a *round*. At the end of the experiment you will be paid for one, randomly selected, round. Each round has an equal chance of being chosen as the round that you would be paid for, so, you should treat each round with equal importance. As you proceed through the experiment your potential earnings will be displayed in **points**. At the end of the experiment we will convert points to dollars at an exchange rate of 1 point = $0.07. You will also be paid a $5.00 show up fee in addition to any money you earn during the experiment.

The basic idea

In each round, you will be placed in a set of 7 players.[2] Each of the seven of you will receive a private signal, that will be a randomly drawn integer between 1 and 100. You will have to guess the average of these seven numbers. In order to make this decision, you can choose to be linked to up to two out of the six other players, and you will be able to see the signals that those linked players received. Therefore, you will know three of the seven signals in your set when you guess the average of those seven signals. In any round, the more accurate your guess is the more points you can expect to earn.

Groups

Within your set of 7 players, you will be formed into two groups: a **red group** and a **blue group**. If a player has a signal that lies between 1 and 50, then they will be placed in the blue group. If a player has a signal between 51 and 100, then they will be placed in the red group. Thus, a player with a high signal is in the red group, and a player with a low signal is in the blue group.

[2]Note that the other 'players' do not make any decisions that affect you. Because of this, and in order to speed up the experiment, the players are computers.

Forming links

In each period, you shall form up to 2 links to other players. You may choose whether you wish to form two links with players within the same color group as you (either red or blue), one link with a red group player and one link with a blue group player, or two links with players from the differing color group. You will earn points for each link you form, and the number of points you earn may vary depending on whether you form the link with a player from your own group or the other group. When you are making your decisions, the number of points that you will earn from each choice will be clearly displayed on your screen, as in the example below.

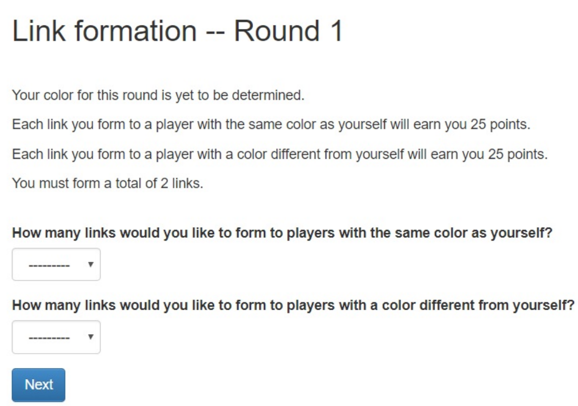
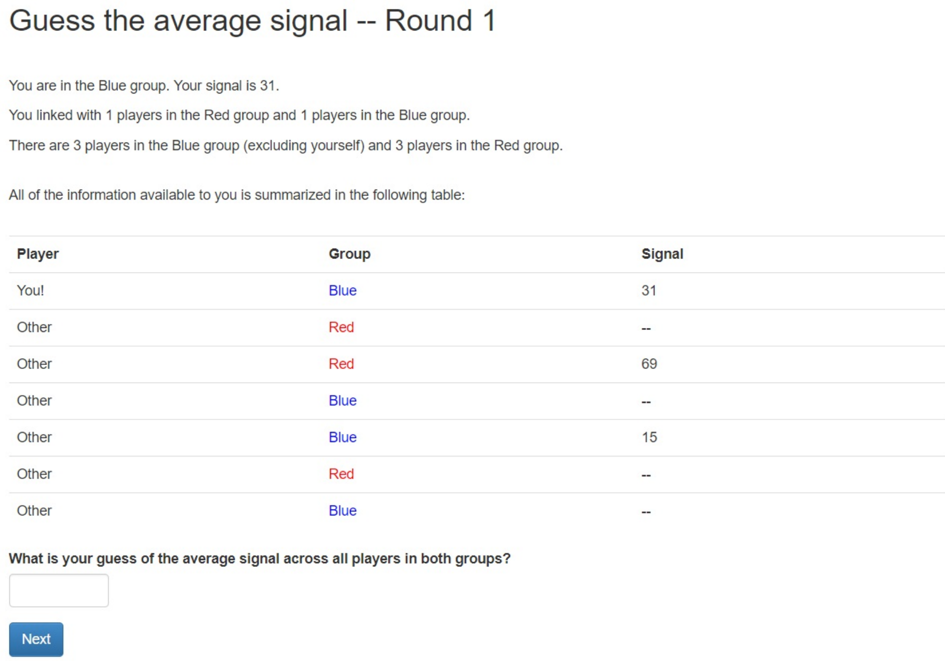


Figure 3: Screen shot of the link choice screen

If there are not enough players available for you to link with in a particular group (e.g. you request to link with two players in the same group as you, but you are the only player in your color group) then you will not be able to form those links. In this case, the computer will automatically allocate you two links that are available. For example, in the sample screen above if you elect to form two links with players from your own group, but there is only one other player in your group, then you would form 1 link with each group.



Guess the average signal -- Round 1

You are in the Blue group. Your signal is 31.

You linked with 1 players in the Red group and 1 players in the Blue group.

There are 3 players in the Blue group (excluding yourself) and 3 players in the Red group.

All of the information available to you is summarized in the following table:

Player	Group	Signal
You!	Blue	31
Other	Red	--
Other	Red	69
Other	Blue	--
Other	Blue	15
Other	Red	--
Other	Blue	--

What is your guess of the average signal across all players in both groups?

Next

Figure 4: Screen shot of the guess page

Guessing the average signal

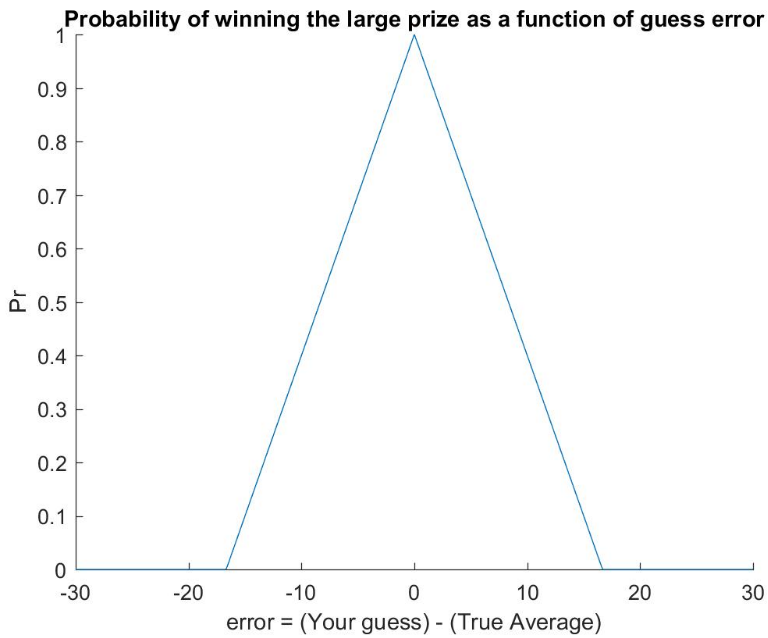
After you have formed the links, for a given round, you will be required to guess the average signal value across all seven players.

For each round there is a large prize, worth 200 points, and a small prize, worth 50 points. **The more accurate your guess is, the higher are the chances of you winning the large prize.**

If your guess is exactly accurate then you will win the large prize with certainty. If your guess is more than $16\frac{2}{3}$ points away from the true average you will earn the small prize. If your guess is between 0 and $16\frac{2}{3}$ of the true average, the probability that you will win the large prize is given by following formula:

$$1 - \frac{6 * \text{error}}{100}$$

where "error" is the difference between your guess and the true average. The probability of winning the large prize is equivalently displayed in the graph below. Notice that improving your guess accuracy by 1 unit increases your chances of winning the large prize (worth 200 points) by 6 percentage points, and improving your accuracy by 5 units increases your chances of winning the large prize by 30 percentage points.



Rounds

There will be 40 rounds in total. After 20 rounds, there will be a slight change in procedures for rounds 21 through 40. Your computer screen will indicate when the change in procedure will occur, and it will disclose what the change is.

[For the remaining rounds the computer will calculate your "guess" automatically. Your "guess" will be the average of the three signals that you observe.]

Payment

Determining which round will be paid

When you arrived, you were given a sealed envelope. DO NOT OPEN THE ENVELOPE UNTIL YOU ARE INSTRUCTED TO. The envelope contains a randomly generated number between 1 and 40, with each number having an equal chance of being in the envelope. The number that is inside the envelope will determine which of the 40 rounds you will be paid for.

Calculating your payment

You will receive the points you earned from your linking choice in the appropriate round plus either the large or small prize.

Your guess in the appropriate round will determine your probability of winning the large prize. After you have completed all 40 rounds, a randomly generated number (uniformly distributed between 0% and 100%) will appear on your screen. You will win the large prize if your percentage chance of winning is larger than the randomly drawn number, otherwise you will win the small prize. This procedure ensures that you will win the large prize with the correct probability.

Survey

At the end of the experiment there will be a short demographics survey. Please fill this in accurately. If you would prefer not to answer any of the questions you may do so, and there will be no penalty for not filling it in.

Frequently Asked Questions

Q1. Is this some kind of psychology experiment with an agenda you haven't told us?

Answer: No. It is an economics experiment. If we do anything deceptive, or don't pay you cash as described, then you can complain to the campus Human Subjects Committee and we will be in serious trouble. These instructions are meant to clarify how you earn money in the experiment, and our interest is in seeing how people make economic decisions.

Q2. Do I have to enter integer (whole) numbers for my guesses?

Answer: No, you may use decimal places if you wish, but you cannot use fractions.