Memory and Partial Consideration in Choice*

Yuval Salant, *Northwestern University*
Jörg L. Spenkuch, *Northwestern University*
David Almog, *Harvard Business School*

August 2026

Abstract

We propose that memory guides choice behavior in unfamiliar environments by shaping which alternatives decision-makers consider. In our model, individuals preferentially consider in-memory alternatives over out-of-memory ones. This mechanism generates a "memory premium," whereby decision-makers disproportionately choose in-memory options, even when they are objectively inferior. Our theory also predicts a "memory speedup." Conditional on being chosen, in-memory alternatives are selected more quickly than out-of-memory ones. We test and confirm these predictions in a unique observational data set and a tightly controlled laboratory experiment. Taken together, our findings point to a role for memory in choice that goes beyond the subjective evaluation of alternatives.

*The first two authors are listed in random order. This manuscript supersedes our previous working paper entitled "The Memory Premium" (Salant et al. 2025). We gratefully acknowledge helpful comments from Ido Ben Artzi, Ariel Rubinstein, David Smerdon, and Leeat Yariv, as well as from seminar audiences at Chicago Booth, Northwestern, Penn, Penn State, Toronto, USC, the Barcelona Summer Forum, and the Nash 75 Conference at Oxford. We are also grateful to Thibault Duplessis and Theo Wait for providing us with data from Lichess, and for patiently answering our questions. Austin Coffelt and Michael Indenkämpen provided excellent research assistance. The research in this paper was supported in part through the computational resources provided by the Quest high-performance computing facility at Northwestern University. The experiment in this paper has been reviewed by the Northwestern IRB, and was preregistered in the OSF Registries at `https://osf.io/8tkd4`. Correspondence should be addressed to the first two authors at: MEDS Department, Kellogg School of Management, 2211 Campus Dr, Evanston, IL 60208; or by email to `y-salant@kellogg.northwestern.edu` and `j-spenkuch@kellogg.northwestern.edu`.

1. Introduction

Economic actors frequently navigate new environments. Whether a firm adapts to a sudden regulatory shift, a bidder enters an auction with an unfamiliar format, an executive moves across industries, or a consumer encounters a novel product, decision-makers routinely lack direct prior experience with the problem at hand. At the same time, some features of the new environment may evoke memories of decisions in related settings. How do such memories affect choice behavior?

Prior work in economics focuses on how memory affects beliefs, reference points, and subjective assessments of value (see, e.g., Gilboa and Schmeidler 1995; Mullainathan 2002; Gennaioli and Shleifer 2010; Bordalo et al. 2020, 2023).[1] In this paper, we propose a complementary mechanism. Memory may shape not only *how* alternatives are evaluated but also *which* options are considered in the first place. Rather than considering all available alternatives, individuals may disproportionately attend to those that readily come to mind.

Our analysis begins by formalizing this idea in a model of satisficing with memory-dependent partial consideration. In the model, a decision-maker evaluates alternatives one at a time until she finds one that meets her aspiration threshold. Sampling for evaluation favors in-memory alternatives. Relative to out-of-memory options, they appear earlier in the search order and are thus more likely to be considered.

One key prediction of this model is that choice behavior should exhibit a "memory premium." That is, in-memory alternatives are chosen more frequently than out-of-memory ones, even when the latter are objectively better. A second key prediction is the "memory speedup." Conditional on being chosen, in-memory alternatives are selected more quickly than out-of-memory ones.

We test and confirm these predictions in a unique observational data set on choice behavior in unfamiliar environments. In our data, we observe the opening moves of nearly one hundred and fifty thousand experienced chess players as they engage in a game of Chess960 for the first time. Chess960 is a variant of classical chess. It uses the same board and pieces, and it relies on the same rules. The key difference relative to regular chess is that the initial positions of the pieces on players' back ranks are randomized, resulting in 960 possible starting positions.

Chess960 provides an almost ideal setting for studying how memory guides decision-making in unfamiliar environments. In regular chess, players typically invest significant amounts of time and effort into memorizing effective opening moves and counter-strategies. In Chess960, however, the large number of potential starting positions makes rote memorization of opening sequences impractical. Moreover, given that even minor differences in the initial placement of

[1] A related strand of literature studies the implications of constraints on encoding, storage, and recall (see, e.g., Wilson 2014; Azeredo da Silveira and Woodford 2019; Azeredo da Silveira et al. 2024).

pieces can have profound strategic implications, opening moves that tend to perform well in classical chess need not be a good choice in a particular starting position in Chess960. In our analysis, we ask whether players' decisions are nonetheless shaped by memories from regular chess.

There are two main challenges to empirically documenting the memory premium and speedup in a real-world environment. The first one is to carefully control for the quality of the available alternatives. Accounting for quality is important in order to rule out that in-memory alternatives are chosen more frequently because they are better than out-of-memory ones. In the context of Chess960, we can overcome this difficulty by using sophisticated computer algorithms to objectively measure the quality of all available opening moves. These algorithms have proven to be considerably better at analyzing chess positions than even the best grandmasters.

The second challenge is to approximate players' (unobservable) memory. Here, we draw on the richness of our data. Since we observe each player's history of opening moves in all standard chess games on the platform, we can model her memory as a function of this history prior to her first game of Chess960. According to our workhorse definition, a given opening move is "in memory" for a particular player if she previously used that move in at least one standard chess game. This is a conservative definition because, due to decay, not every memory of a move that should, in principle, be retrievable can, in fact, be recalled. We also explore alternative definitions of memory, such as classifying a move as in-memory if it had been played sufficiently often or within a certain period of time.

Irrespective of particulars, we find that, in their first game of Chess960, players are significantly more likely to choose an in-memory opening move than an equally good out-of-memory alternative. Relying on our workhorse definition of memory, the average memory premium equals 3.8–4.5 percentage points (p.p.), or about 75–89% of the mean choice frequency in this setting.[2]

To connect the memory premium to the sampling mechanism in our model, we document that the size of the premium varies systematically with several proxies for the probability of recall. Drawing on well-established findings in psychology and neuroscience (see, e.g., Kahana 2012), we show that the memory premium is associative, subject to repetition and interference effects, and that it decays over time. We also ask whether it matters if memories are "good" or "bad." To this end, we distinguish between moves that players had used more and less successfully in the past. Good memories are associated with a higher premium than bad ones,

[2]For comparison, taking our regression estimates at face value, the implied difference in the choice frequency of moves at the fifth and ninety-fifth percentile of the quality distribution is about 3.2 p.p. The memory premium is, therefore, large enough to overcome even significant differences in the objective quality of moves.

but the difference is modest, and even moves with below-median prior payoffs carry a sizeable premium. In addition, a horse race suggests that repetition and decay are quantitatively more important than prior success. Taken together, our findings imply that the memory premium is not simply an experience effect or a reduced-form preference for familiar alternatives; it is closely linked to the cognitive forces that govern recall.

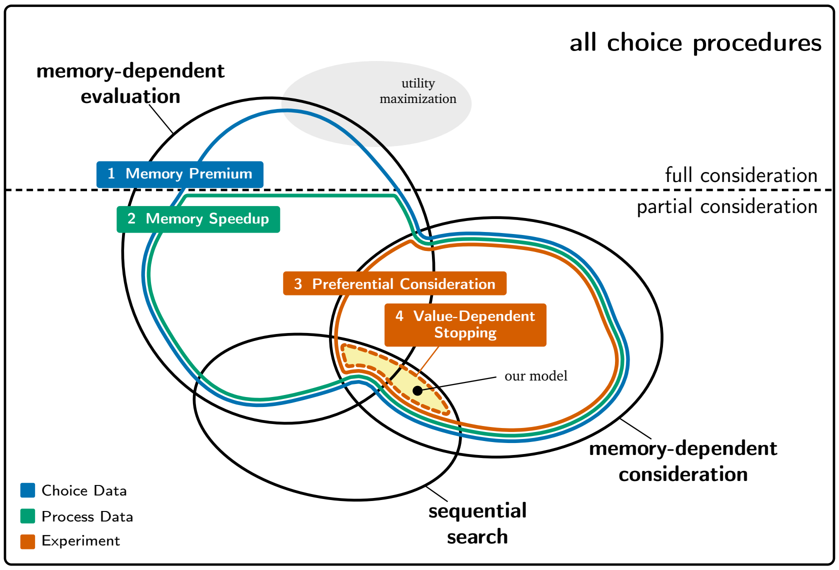
We next turn to the memory speedup. Since our data do not include timestamps for the first move in a game, we analyze how long players deliberate prior to choosing their second and third moves. We classify such a move as in-memory if a player previously selected that move after encountering the same sequence of preceding moves in standard chess. Holding the board position fixed, players select in-memory moves about 10% faster than out-of-memory ones. The speedup is more pronounced in the first game of Chess960, when memories from standard chess are likely especially salient. Like the memory premium, we find that as players gain experience, the speedup diminishes but it does not disappear. We also find evidence that the speedup is mediated by the same cognitive mechanisms that govern recall.

The memory premium and the memory speedup are nonobvious patterns in choice behavior that are predicted by our model. We note, however, that neither the premium nor the speedup constitutes direct evidence of our key idea, i.e., that decision-makers preferentially consider in-memory alternatives. In order to provide such evidence, we conducted a pre-registered laboratory experiment that allows us to observe which options participants actually attend to and in what order.

In our experiment, subjects choose between six investment options, half of which are in-memory on account of having been encountered earlier. By randomly assigning memory status and by *inducing* values and beliefs, our experimental design eliminates any payoff-based reason to favor in-memory alternatives. Nonetheless, both the memory premium and the memory speedup emerge in the experimental data. Beyond replicating these findings in a tightly controlled setting, the experiment yields two additional results. First, in-memory alternatives are more likely to be considered, and they are more likely to be considered first—even when the most promising in-memory option is worse than the average out-of-memory alternative. Second, participants are more likely to terminate their search process after evaluating an option with a high realized value. The first result provides direct support for the idea that memory shapes which alternatives are considered. The second result is suggestive of satisficing, as assumed in our model.

Taking a step back, Figure 1 organizes our findings by what they imply about decision-making in unfamiliar environments. The memory premium shows that memory affects choice. It therefore rules out memory-independent choice procedures; but it does not by itself identify the mechanism through which memory operates.

Figure 1: Conceptual Overview



Notes: Figure maps our empirical findings onto the space of choice procedures. The black box symbolizes all possible choice procedures, with the horizontal line separating full- from partial-consideration procedures. Black ellipses correspond to classes of choice procedures with memory-dependent evaluations, memory-dependent consideration, and sequential search, respectively. The gray oval marks procedures featuring utility maximization, i.e., full-consideration procedures in which the decision-maker selects the alternative with the highest subjective value. The colored contours are nested, so the nth contour encloses every procedure consistent with the first n findings. Color encodes the data behind each finding.

To this end, it is useful to distinguish between two overlapping classes of choice procedures. In the first one, memory affects choice through a *utility channel*, i.e., by shaping the subjective evaluation of alternatives. In the second class of procedures, memory operates through the *consideration channel*. That is, memory shapes which of the available options decision-makers even consider. Although the memory premium and its comparative statics are a natural consequence of preferential consideration, they could be generated by either mechanism.

The memory speedup helps to narrow the set of candidate procedures. Under some assumptions, the observed speedup allows us to rule out any choice procedure that requires full consideration, including memory-dependent utility maximization as an important special case. Holding the choice set fixed, full consideration entails evaluating all available options, irrespective of which one is ultimately selected. As a consequence, deliberation times should not depend on the memory status of the chosen alternative. The memory speedup contradicts this prediction.

Even though the memory speedup points to memory-dependent choice procedures with partial consideration, it does not automatically follow that memory operates through the consideration channel. Our experiment closes this gap. The fact that subjects preferentially evaluate in-memory alternatives shows that the consideration process itself is memory-dependent.

To be clear, we do not dispute the important role of the utility channel in real-world choice behavior. Our findings point to a class of choice procedures in which memory guides sequential consideration, possibly in conjunction with memory-dependent evaluations.

Related Literature. Our work contributes to a growing theoretical literature on the role of memory in decision-making. In early contributions, Gilboa and Schmeidler (1995, 2001) develop a theory of case-based decisions in which remembered outcomes determine the values of the available alternatives, while Mullainathan (2002) models how associative recall and rehearsal shape beliefs. More recently, Bordalo et al. (2020; 2023; 2025; 2026) incorporate memory into models of attention, probability judgments, and reasoning. For example, Bordalo et al. (2020) present a theory of consumer choice in which memories of past purchases determine reference points that affect which attributes of goods are salient. Other important contributions include Nagel and Xu (2022) and Wachter and Kahana (2024) on memory-based beliefs and financial decisions, Fudenberg et al. (2024) on long-run behavior under selective recall, and Kőszegi et al. (2022) on choice and self-esteem. While extant theories differ substantially in primitives and applications, one common feature is that memory influences choice (only) through the utility channel. That is, memory shapes beliefs, norms, salience weights, or other components of the subjective value of alternatives.[3]

Our first contribution is to propose a different, complementary mechanism for memory dependence in choice. In our model, memory does not affect *how* alternatives are evaluated but *which* options are considered in the first place. This mechanism connects the economics of memory to theories of satisficing and limited consideration (e.g., Simon 1955; Manzini and Mariotti 2014). In Simon's classic formulation, the decision-maker evaluates alternatives sequentially, but the source of the order in which alternatives are examined is left open. Memory provides one cognitively plausible source of that order. In other models of limited consideration, consideration probabilities are primitives to be inferred from observed behavior. Linking consideration to memory retrieval gives these probabilities behavioral structure and generates testable predictions based on repetition, recency, environmental cues, and interference.

Our work also contributes to an empirical literature on the effect of memory on belief formation and economic decisions (see, e.g., Afrouzi et al. 2023; Charles 2022, 2025; D'Acunto et al. 2021; Enke et al. 2024; Hartzmark et al. 2021; Jiang et al. 2025). In influential work, Malmendier and Nagel (2011, 2016) show that individuals who lived through the Great

[3]Another strand of literature emphasizes explicit constraints on memory and mental representations (see, e.g., Salant 2011; Wilson 2014; Woodford 2014, 2020; Khaw et al. 2021; Azeredo da Silveira and Woodford 2019; Azeredo da Silveira et al. 2024). Yet even in this class of models, memory and cognitive constraints do not determine consideration.

Depression are less likely to invest in the stock market later in life, and that experiencing high inflation has lasting effects on inflation expectations.[4] Graeber et al. (2024) provide evidence that selective memory explains why the effect of stories on beliefs decays significantly less than the effect of statistics. Guided by our model, we contribute to this literature by documenting the memory premium and the memory speedup. Together with their comparative statics, these regularities provide new empirical moments against which models of memory-dependent choice can be assessed.

Finally, we make a modest technical contribution to the literature on sequential search. Much of this literature assumes that alternatives are drawn from a stationary distribution, either because the market is a continuum or because sampling is with replacement (see, e.g., McCall 1970; Carlson and McAfee 1983; De Los Santos et al. 2012; Eliaz and Spiegler 2016).[5] We instead study stochastic search without replacement from a finite menu with heterogeneous sampling probabilities. The sampling distribution thus changes after every evaluation, which complicates even basic comparative statics. We address these difficulties by coupling and comparing search paths so that their rejected alternatives can be ordered. These arguments yield general comparative statics for choice probabilities and first-order stochastic dominance results for deliberation times.

2. Theory

2.1. *Preliminaries*

Fix a grand set of alternatives $\mathcal{X}$. Alternatives have objective values that depend on the environment. We use $V(x, E)$ to denote the value of alternative $x \in \mathcal{X}$ in environment E. A choice situation is a pair (A, E), where $A \subseteq \mathcal{X}$ denotes the set of available alternatives, and E is the environment in which choice is to be made. Decision-makers (DMs) encountered a fixed choice situation $\hat{S} = (\hat{A}, \hat{E})$ in the past. We refer to $\hat{E}$ as the "familiar" environment and to all other environments as "unfamiliar." The latter may differ in their similarity to $\hat{E}$.

A DM's memory M consists of all experiences she had in the familiar environment. For simplicity, we equate an experience with a chosen alternative. That is, M is a vector $(a_1, a_2, \ldots)$, where $a_k \in \hat{A}$ is the alternative that the DM chose k periods ago from $\hat{S}$.[6] We refer to an alternative x as "in memory" and write $x \in M$ if there exists a k such that $x = a_k$. An alternative x that is not in memory M is said to be "out of memory." In this case, we write $x \notin M$. Given a memory M and an in-memory alternative a, we denote by $M - a$ the memory that is obtained by removing all instances of a from M.

[4] See Malmendier and Wachter (2024) for a survey of how past experiences shape expectations and behavior.
[5] For important exceptions, see Weitzman (1979) and Choi and Smith (2024).
[6] One may enrich the definition of an experience by including additional variables. Since we do not use additional variables in the analysis, we keep the definition concise.

In what follows, we simplify notation by fixing the choice situation $S = (A, E)$ and suppressing the dependence of model objects on S. Unless stated otherwise, every alternative is understood to belong to A, and M denotes the restriction of the DM's memory to experiences involving alternatives in A. Since S is arbitrary, the predictions below apply to any choice situation.

2.2. *Satisficing with Memory-Dependent Consideration*

Following Simon (1955), we assume that the DM satisfices. That is, she sequentially samples alternatives, evaluates them, and makes a choice as soon as she finds one that is "good enough." Sampling for evaluation is memory-dependent in the sense that the sampling probability of in-memory alternatives is larger than that of out-of-memory ones.

Memory-Dependent Consideration. To formalize the idea that the DM preferentially considers in-memory alternatives, we let $\mu_M(x)$ denote the probability that she recalls alternative x from memory M. If $x \in M$ then $\mu_M(x) > 0$; otherwise $\mu_M(x) = 0$. We assume that $\sum_x \mu_M(x) \leq 1$. Any remaining probability corresponds to the event that the DM does not retrieve an in-memory option.

Retrieval probabilities depend on features of memory and the environment that prior work has shown to affect recall. For example, retrieval is more likely when an alternative was chosen more recently or more often, and when the current environment is more similar to the familiar one. To capture the idea that memories interfere with one another, we impose two monotonicity restrictions on how retrieval probabilities change when an alternative enters memory.

The first one is *individual monotonicity*: For any memory M, and every $a \in M$ and $x \in M - a$,

$$\mu_M(x) \;\leq\; \mu_{M-a}(x).$$

In words, adding an alternative to memory creates competition for retrieval. As a consequence, the probability of recalling any other in-memory alternative weakly decreases.

The second restriction is *aggregate monotonicity*: For every M, and every $a \in M$,

$$\sum_{x \in A} \mu_M(x) \geq \sum_{x \in A} \mu_{M-a}(x).$$

That is, the DM is more likely to retrieve a previous experience from memory when there are more experiences to recall. Taken together, individual and aggregate monotonicity imply that, when an alternative enters memory, the increase in its retrieval probability is at least as large as the total decrease in the retrieval probabilities of all other alternatives.

Retrieval probabilities are important because sampling for evaluation depends on "what comes to mind." Let $A_k = A - \{x_1, \ldots, x_{k-1}\}$ denote the set of alternatives that remain after $k-1$ evaluations, so that $A_1 = A$. If the DM did not stop after evaluating $k-1$ alternatives, then the next alternative, x_k, is drawn according to

$$p_M(x \mid A_k) = \begin{cases} \mu_M(x) + \left(1 - \sum_{x' \in A_k} \mu_M(x')\right) \frac{1}{|A_k|} & \text{if } x \in A_k \\ 0 & \text{otherwise} \end{cases} . \tag{1}$$

The DM thus samples without replacement. If one of the available alternatives comes to mind, she evaluates it. If she recalls none of the alternatives in A_k, then all remaining options are equally likely to be considered next.

Satisficing. After drawing an alternative x_k according to p_M, the DM assesses its value. Alternative x is *satisfactory* if its subjective value to the DM, $u(x)$, exceeds her aspiration threshold T. Subjective values may differ from objective values, but we assume that they do not depend on the DM's memory or the order in which alternatives are drawn.[7] We impose this restriction deliberately in order to isolate the effect of memory-dependent partial consideration from memory-dependent evaluations.

If x_k is satisfactory, then the DM stops and chooses it; otherwise, she draws another alternative. The DM continues in this fashion until she finds an alternative that is good enough. If she exhausts the choice set without finding a satisfactory alternative, then she chooses the one with the highest assessed value.[8]

2.3. *Predictions*

Our model makes two key predictions. The first concerns choice probabilities, and the second deliberation times.

To state the first prediction, we follow Rubinstein and Salant (2006) and turn to the random choice function that our choice procedure induces. Given memory M, we write $C_M(x)$ for the probability that the DM chooses alternative x under M, so that C_M is a probability distribution on A. Choice behavior is random because consideration is probabilistic. We define $\Delta_M(a) = C_M(a) - C_{M-a}(a)$ to be the change in the probability of choosing a as a enters memory.

[7]The set of satisfactory alternatives can thus be fixed before the search begins.

[8]If more than one alternative has the highest value, then the DM chooses uniformly among them.

Prediction 1 (Memory Premium): *Satisficing with memory-dependent consideration gives rise to a* memory premium. *That is,* $\Delta_M(a) \geq 0$ *for every memory* M *and every alternative* $a \in M$*. If* A *contains at least two satisfactory alternatives, then the inequality is strict for every satisfactory alternative in* M*.*[9]

To build intuition, recall that a satisfactory alternative a is chosen if and only if it is considered before any other satisfactory option. On the initial draw, adding a to memory raises its own sampling probability while weakly lowering the sampling probability of every other available alternative. It is, therefore, more likely to be evaluated and, ultimately, chosen.

Although instructive, comparing the initial sampling probabilities does not suffice to establish the memory premium. Since the DM samples without replacement, the sampling distribution changes after every evaluation. To resolve this complication, consider a sampling path on which another satisfactory alternative is drawn before a. Along any such path, every sampled alternative differs from a. When a enters memory, the probability of each draw along the path weakly falls. Hence, every path leading to another satisfactory alternative becomes weakly less likely, which makes a weakly more likely to be the first satisfactory alternative that the DM considers. Our proof in Appendix A formalizes this argument and establishes strictness when there are at least two satisfactory alternatives.

The same logic also yields comparative statics. Changes that make a satisfactory alternative easier to retrieve, and thus shift sampling probability away from competing alternatives, increase that alternative's memory premium. The premium should, therefore, be mediated by the cognitive forces that govern recall. For example, the memory premium should increase with repetition and decline as memories decay over time. Section 5 investigates these predictions empirically.

We next develop the second key prediction of our model, which concerns deliberation times. Following Rubinstein (2007, 2016), we interpret deliberation time as a proxy for cognitive effort. Choices that require more effort take longer.

We consider two sources of cognitive effort, and hence of deliberation time. The first is a fixed cost associated with the choice situation, such as the time the DM needs to orient herself or understand the decision problem. The second is variable and depends on which alternatives the DM evaluates. We do not take a stand on what exactly evaluation entails. Depending on the choice procedure, it may involve, for example, assigning a cardinal numerical value to an alternative or comparing its value with an aspiration threshold. Writing $\mathcal{E}$ for the set of

[9]If no alternative is satisfactory, the DM maximizes and $\Delta_M(a) = 0$. If exactly one alternative is satisfactory, it is chosen with probability one regardless of memory.

alternatives that the DM evaluates, we model deliberation time as

$$DT = \tau + t_M(\mathcal{E}), \tag{2}$$

where τ is the fixed component and t_M is the variable one.

We require that expanding the set of evaluated alternatives increases deliberation time. That is, for every memory M,

$$t_M(\mathcal{E}) < t_M(\mathcal{E}') \quad \text{whenever} \quad \mathcal{E} \subsetneq \mathcal{E}'.$$

We also assume that out-of-memory alternatives take at least as long to evaluate as in-memory options. Formally, for every M, and all $x \in M$ and $y \notin M$,

$$t_M(\mathcal{E} + x) \leq t_M(\mathcal{E} + y) \qquad \text{for all } \mathcal{E} \subseteq A - \{x, y\}.$$

Except for these two restrictions, the cost structure is flexible. Deliberation time may depend only on the number of evaluated alternatives, or it may vary with their identities. Moreover, t_M need not be additively separable. In particular, it may include the costs of comparisons among evaluated alternatives, and it may depend on the DM's memory.

Given these assumptions, our model predicts that, conditional on being chosen, in-memory alternatives are selected more quickly than out-of-memory ones.

PREDICTION 2 (Memory Speedup): *Satisficing with memory-dependent consideration gives rise to a* memory speedup. *That is, for any memory M and any satisfactory alternatives $x \in M$ and $y \notin M$ that are each chosen with positive probability, deliberation time conditional on choosing x is weakly shorter than deliberation time conditional on choosing y. It is strictly shorter if there is at least one unsatisfactory alternative in the choice set.*[10]

To develop intuition for why the model predicts a speedup, consider the special case in which the variable component of deliberation time depends only on the number of alternatives that the DM considers. Because the choice situation and memory are held fixed, any reduction in deliberation time must come from fewer evaluations. Since in-memory alternatives receive higher priority in consideration than out-of-memory ones, they tend to appear earlier in the search order and, therefore, tend to be evaluated before the latter.

The proof of Prediction 2 formalizes this intuition and extends it to set-monotone variable deliberation costs. The chief technical complication is that the relevant sampling probabilities at each stage of the search process need to condition on which alternatives have already been

[10] If all alternatives are satisfactory, the DM always stops after the first evaluation.

evaluated and which alternative is ultimately chosen. We show that, for every remaining menu and every unsatisfactory alternative c, a search process ending in $x \in M$ is weakly less likely to continue through c than another search process ending in $y \notin M$. The two processes can thus be coupled so that, for every realization of the coupling, every alternative rejected before choosing x is also rejected before choosing y.

Because both search processes terminate at different alternatives, this subset relation does not, by itself, order their evaluated sets. The cost restrictions above bridge the gap. Replacing the out-of-memory alternative y with the in-memory alternative x cannot increase deliberation time, while evaluating any additional alternatives strictly increases it. Deliberation times conditional on choosing x are, therefore, first-order stochastically smaller than those conditional on choosing y. If at least one alternative is unsatisfactory, then the search process ending at y evaluates strictly more alternatives with positive probability, which yields a strict speedup.

In Appendix A, we consider a different notion of the memory speedup. Instead of fixing the DM's memory and comparing deliberation times conditional on selecting different alternatives, we show that the model also predicts a speedup across memories: holding the chosen alternative fixed, deliberation time is shorter when that alternative is in memory than when it is not. The proof of this prediction is more difficult, but it allows us to dispense with the assumption that out-of-memory alternatives take weakly longer to evaluate. Instead, we require that adding alternatives to memory does not increase deliberation time.[11]

2.4. *Discussion*

In our model, memory affects choice only through the sampling process. As mentioned above, we deliberately abstract from the possibility that memory shapes the subjective evaluations of alternatives in order to isolate the implications of memory-dependent consideration.

A more conventional approach might have been to augment the textbook model of utility maximization by allowing memory to affect the evaluation of alternatives. Specifically, let $u_M(x)$ denote the subjective value the DM assigns to alternative x under memory M. This value may depend on x's objective value $V(x)$ and on the environment.[12]

Alternative Model 1 (Utility Maximization with Memory-Dependent Evaluations): *The DM evaluates every alternative $x \in A$ and chooses one that maximizes $u_M(x)$.*

To see how this model can generate the memory premium, suppose that entering memory raises a's subjective value relative to other alternatives. That is, $u_M(a) - u_M(x) > u_{M-a}(a) -$

[11] In symbols, we assume that $t_M(\mathcal{E}) \leq t_{M-a}(\mathcal{E})$ for every $a \in M$ and every $\mathcal{E} \subseteq A$.

[12] Recall that we suppress environmental dependence from the notation.

$u_{M-a}(x)$, for every $x \neq a$. Entering memory thus makes a more attractive and more likely to be chosen.

The comparative statics of the memory premium follow if the change in a's relative value increases in the probability of recall. Such a relationship between retrieval and value could arise because the DM treats recall of past choices as a signal of value. It could also arise in models of choice under uncertainty. Interpreting $u_M(a)$ as the expected utility of alternative a, a higher likelihood of recall would increase its subjective value if it increases the probability weight placed on states in which a performs well. Either way, memory-dependent utility maximization can explain the memory premium and its comparative statics if easier retrieval makes an alternative more attractive relative to its competitors.

Utility maximization cannot, however, generate the speedup in Prediction 2. Since memory is held fixed, and given that the DM evaluates all of the available options, deliberation time cannot depend on which alternative is ultimately selected. The same conclusion applies to any full-consideration procedure and, more generally, whenever the distribution of evaluated sets is the same conditional on each possible choice.

By contrast, partial-consideration procedures in which memory affects which or how many alternatives are evaluated *can* generate a speedup. Our satisficing model is one leading example. Section 6 discusses other procedures, and Section 8 presents experimental evidence that speaks directly to memory-dependent consideration and satisficing.

3. Setting and Data

3.1. *Chess960*

Chess960—also known as Fischer Random Chess—is a variant of standard chess, which was invented in 1996 by former world champion and legendary grandmaster Bobby Fischer. The game uses the same board, pieces, and basic rules as regular chess. The key difference relative to standard chess is that the initial positions of the pieces on players' home ranks are randomized. That is, pawns are placed in their usual starting positions, but the pieces behind them are shuffled according to two rules: (i) Bishops must be placed on opposite-color squares, and (ii) the king must be positioned between the rooks to enable castling. These constraints result in 960 possible initial board configurations—hence the name Chess960.

Chess960 provides an almost ideal environment to study the role of memory for decision-making in unfamiliar environments. First, there is no uncertainty about the rules of the game; yet first-time players of Chess960 may be unsure about the value of individual alternatives. After all, even minor differences in the initial placement of pieces can have important implications for optimal strategy. Second, individuals' memory can be empirically approximated. Chess players often memorize sequences of effective opening moves and countermoves, on which they rely in standard chess. To the extent that we observe players' choices in standard chess, we can approximate which opening moves they have stored in memory. This, in turn, allows us to relate memory to their opening moves in the first game of Chess960. Third, randomness in the positioning of other pieces on the board provides us with exogenous variation in the quality of a given opening move. This variation is helpful in ruling out that in-memory alternatives are chosen more frequently because they happen to be better than out-of-memory ones. Fourth, data on games of chess and Chess960 are abundant, affording us enough statistical power to document even subtle comparative statics.

3.2. *Data and Descriptive Statistics*

Our data come from `lichess.org`, one of the most popular online chess servers.[13] Funded by donations, Lichess is ad-free and allows anyone to play live chess games at no cost through a high-quality graphical user interface. Every day, Lichess hosts several million games of standard chess and thousands of games of Chess960. For each variant, Lichess distinguishes between rated and casual games, both of which can be played under different time controls. Unlike casual games, rated games are consequential in the sense that their outcomes directly affect users' strength ratings and rankings on the site. They are thus only available to registered users. Since high ratings tend to be a source of pride among chess players, Lichess has a strict policy against computer-assisted play. Figure 2 shows a typical game of Chess960 on Lichess.

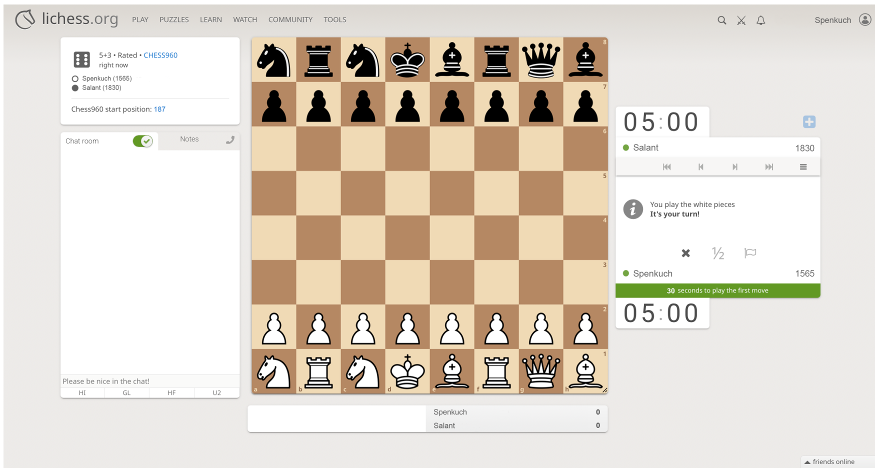
We have data on the universe of games on Lichess from January 2013 through June 2021. The available information includes players' usernames, strength ratings and real-world titles (if any), the type, date and start time of each game, the sequence of moves, as well as the ultimate outcome. For games of Chess960, we also observe the exact starting position. Given the granularity of our data, we can reconstruct every registered user's history of play on the platform. We, therefore, know every opening move of every player in standard chess games as well as in Chess960.

In much of our analysis, we restrict attention to the opening move of approximately 147,000 players as they engage in a game of Chess960 for the first time.[14] There are two restrictions that limit the size of this sample. First, users must play their first game of Chess960 as White, and the game must be rated. While color is randomly assigned (in rated games), about half of first-time players choose to engage in a casual game. We nonetheless impose this restriction because we want to analyze choices with real stakes. Second, prior to their first game of

[13] Our description of these data borrows from Salant and Spenkuch (2026).

[14] Appendix Figure AF.1 verifies that the distribution of starting positions in these games is consistent with random assignment.

Figure 2: Chess960 on Lichess

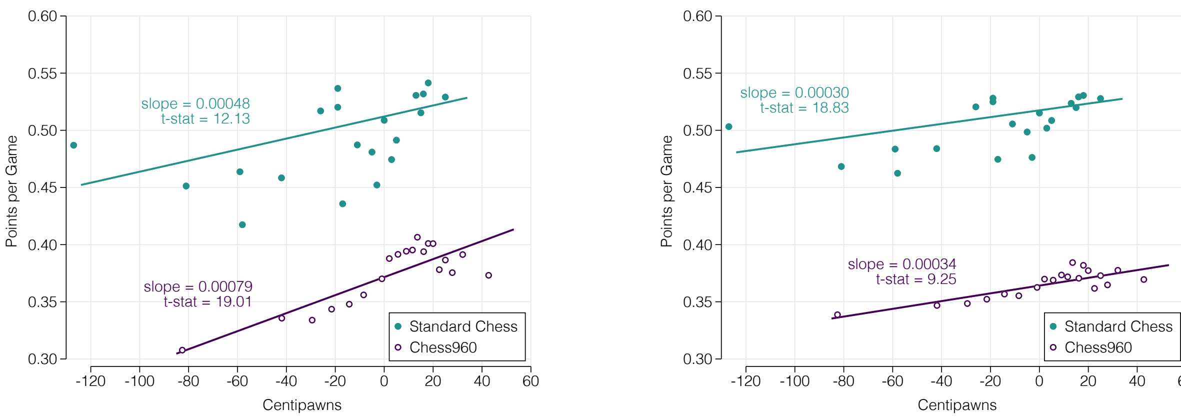


Notes: Figure shows the beginning of a Chess960 game between registered users on `lichess.org`.

Figure 3: Opening Moves and Game Outcomes

(a) Raw Data

(b) Controlling for Player Strength



Notes: Figure shows binscatter plots of the relationship between game outcomes and the Stockfish ratings of opening moves in both standard chess and Chess960. Panel (a) does so based on the raw data, whereas panel (b) presents estimates of the same relationship after controlling for a fixed effect for a player's strength rating in standard chess interacted with that of her opponent. Wins are valued at one point, draws at half a point, and losses at zero.

Chess960, users must have executed at least twenty opening moves in standard chess games on Lichess. This restriction is meant to ensure that we can approximate which moves are stored in memory. In ancillary analyses, we impose alternative restrictions. Reassuringly, different samples yield qualitatively similar results.

Measuring the Quality of Opening Moves. We complement the Lichess data with information on the objective quality of opening moves. To this end, we use the Stockfish software to rate every feasible opening move in every possible starting position. Stockfish is one of the most powerful open-source chess engines. It is far better at analyzing chess positions than even the best grandmasters.[15] Stockfish evaluates board configurations by examining sequences of possible moves (up to a certain depth) using a combination of heuristic evaluation functions and advanced neural networks. It then assigns a numerical score to each available move. A positive score reflects an advantage for White, whereas a negative score reflects an advantage for Black. The size of the advantage is measured in "centipawns," which correspond to one-hundredth of a pawn. For example, a move with a value of +50 leads to a position in which White is ahead by the equivalent of half a pawn.[16] Important for our purposes, Stockfish explicitly supports Chess960.

Figure 3 relies on Stockfish ratings to compare the importance of opening moves in standard chess and in the first game of Chess960. The left panel plots the relationship between the centipawn score of moves and the average outcome of the respective games in the raw Lichess data. Wins are valued at one point, draws at half a point, and losses are worth zero. The panel on the right shows the same relationship after carefully controlling for the strength of both players. In either panel, two patterns stand out. First, novice players of Chess960 win less than half of their games. Second, and more importantly, there is a clear positive relationship between game outcomes and the quality of opening moves. This relationship is at least as strong in Chess960 as in regular chess. Taking the estimated slope in the right panel at face value, choosing an opening move that is approximately one standard deviation (≈ 35 centipawns) better increases White's likelihood of winning by about 1.2 percentage points.

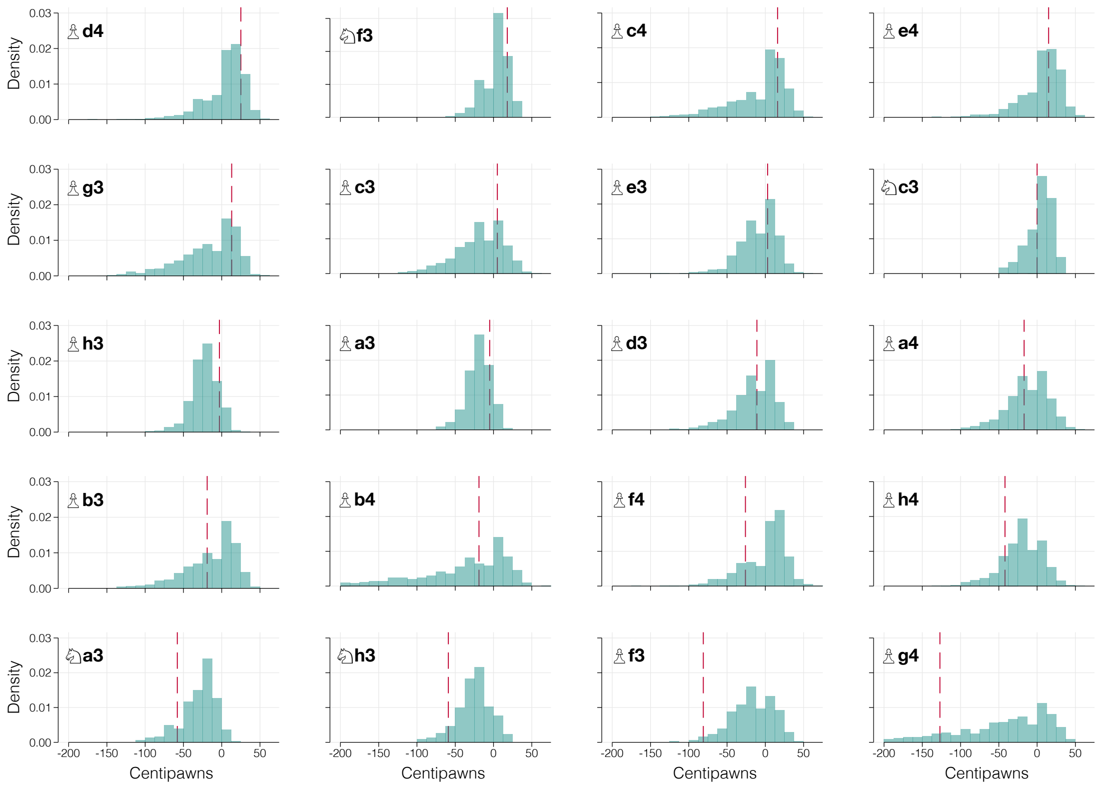
Figure 4 shows that there is a great amount of within-move variation in quality across different starting positions. Each panel in this figure depicts the distribution of Stockfish ratings for a given opening move across starting positions in which the respective move is available. Vertical lines indicate an opening move's centipawn score in standard chess. Evidently, there are many starting positions in which popular opening moves, such as ♙d4 or ♙e4, are rated favorably; but there are also a number of positions in which playing one of these moves would put White at a significant disadvantage. Similarly, opening moves that tend to be bad in standard chess, such as ♙f3 or ♙g4, can, in fact, be quite good in Chess960. Overall, opening moves' centipawn scores in the standard starting position explain less than 6% of their variation in Chess960.

Measuring Memory. To study the connection between memory and choice behavior, we must empirically approximate the content of players' memory, i.e., which moves they are able to recall. According to our workhorse definition, a given opening move is said to be "in

[15] Recent versions of Stockfish can be expected to lose less than one percent of games against any human player.

[16] For comparison, commonly used rules of thumb value each bishop and knight at 300 centipawns, each rook at 500, and the queen at 900 centipawns.

Figure 4: Variation in the Quality of Opening Moves Across Starting Positions



Notes: Figure shows the distributions of centipawn scores of each opening move across all possible starting positions in Chess960. Panels are arranged according to the respective moves' ratings in standard chess, as indicated by the dashed vertical lines in each panel. Only opening moves that are available in standard chess are shown.

memory" for a particular player if she previously chose that move in at least one standard chess game on Lichess. In other words, we model memory based on individuals' entire history of play prior to their first game of Chess960, and we assume that they are able to recall any previous opening move.

An alternative approach would be to model recall as a function of the frequency and recency of play—and perhaps even outcomes. For instance, one might assume that memory is limited to the last one hundred games or to actions that were taken, say, at least twice. The rationale for such alternative definitions is that memory decays and that repetition strengthens recall. A significant drawback of the alternative approach is that it is somewhat arbitrary.

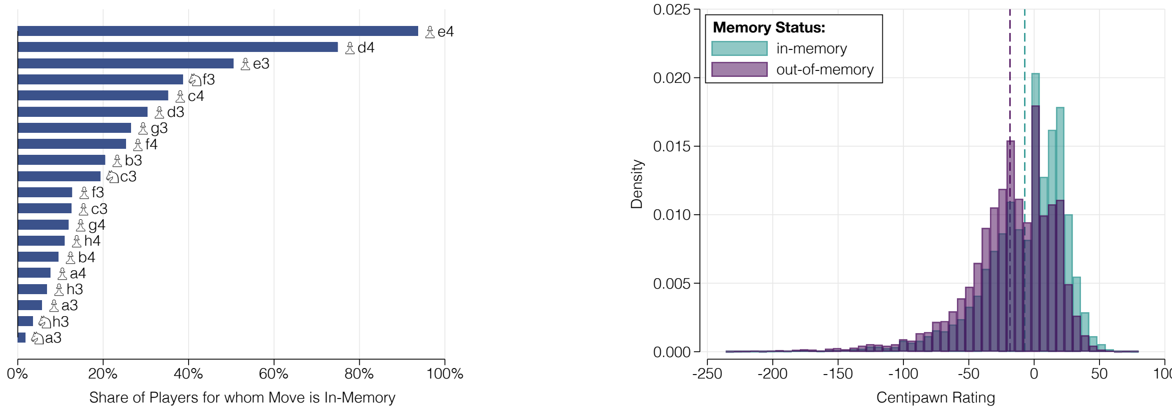
Since we do not know how long an experience remains in memory before it is forgotten, our workhorse definition labels any opening move that might, in principle, be recalled as in-memory. Guided by our model, we then *empirically* explore comparative statics with respect to various proxies for the probability of recall.[17]

[17]In addition, we note that our approach to measuring memory is conservative because our workhorse definition might falsely classify some out-of-memory moves as in-memory. This, in turn, biases our estimates

Figure 5: Memory

(a) Memory Status of Moves

(b) Quality of In- vs. Out-of-Memory Moves



Notes: Panel (a) shows the share of players for whom a particular opening move is in-memory, i.e., who have played this opening move at least once prior to their first game of Chess960. Panel (b) presents the distribution of centipawn ratings in players' first game of Chess960 for moves that are in- and out-of-memory. Vertical lines correspond to the mean rating in each category.

Drawing on our workhorse definition, the left panel in Figure 5 shows the share of players for whom a particular opening move is in-memory. Unsurprisingly, ♙e4 and ♙d4—the most common opening moves in standard chess—are in-memory for the vast majority of players.[18] By contrast, a total of eleven other moves are out-of-memory for more than four out of five individuals. The right panel in Figure 5 presents the distribution of Stockfish ratings across starting positions for players' in- and out-of-memory moves. Although both distributions overlap significantly, in-memory moves have, on average, higher ratings.

Descriptive statistics for the most important remaining variables in our data are shown in Table 1. The average (median) player has engaged in more than 1,000 (270) standard chess games prior to her first game of Chess960. Over the course of these games, she has executed 4.5 (4) different opening moves, which she now holds in memory. The individuals in our analysis thus have substantial experience playing chess. Yet 61% of them lose their first game of Chess960. The latter observation suggests that we, indeed, observe them in an unfamiliar environment.

4. The Memory Premium

Figure 6 provides the first piece of evidence that in-memory alternatives are chosen more frequently than out-of-memory ones. The figure depicts binscatter plots relating move quality to choice frequency. It does so for in-memory and out-of-memory moves. Two key patterns

toward zero. As a robustness check, we estimate the memory premium for several thousand alternative definitions. Virtually all of them yield larger estimates (see Appendix Figure AF.3).

[18] Opening moves that are only available in Chess960 are never classified as in-memory. They are excluded from Figure 5. To gauge the popularity of each opening move in standard chess, see Appendix Figure AF.2.

Table 1: Summary Statistics

Variable	Mean	SD	Percentile 5%	25%	50%	75%	95%	N
A. Player Level								
ELO Rating in Standard Chess	1,555	344	1,026	1,304	1,527	1,792	2,150	147,357
Previous Chess Games	1,030	2,529	48	101	276	889	4,395	147,357
Previous Chess Games as White	521	1,284	24	51	139	448	2,224	147,357
In-Memory Moves	4.98	3.49	1	2	4	7	12	147,357
B. Game Level								
White Wins (%)	35.64	47.89						147,357
Draw (%)	3.31	17.88						147,357
White Loses (%)	61.04	48.77						147,357
Board Similarity	2.10	1.34	0	1	2	3	4	147,357
C. Move Level								
All Moves:								
Choose Move (%)	5.08	21.97						2,898,455
Stockfish Rating	–15.89	35.20	–80	–32	–11	9	27	2,898,455
In-Memory (%)	22.92	42.03						2,898,455
In-Memory Moves:								
Times Played in Chess	111.36	552.94	1	1	5	35	476	664,424
Days Since Last Played	71.08	168.10	0.01	0.94	9.33	56.13	357.78	664,424
Average Past Payoff	0.44	0.32	0.00	0.18	0.49	0.61	1.00	664,424

Notes: Table displays summary statistics for selected variables in our main sample. To be included in this sample, players must play their first game of Chess960 as White, and the game must be rated. We further require that they play at least twenty games of standard chess on Lichess prior to their first game of Chess960. Observations in Panel A correspond to players. Each observation in Panel B corresponds to a player's first game of Chess960. Panel C reports move-level information, so that each observation corresponds to a feasible opening move in a particular game.

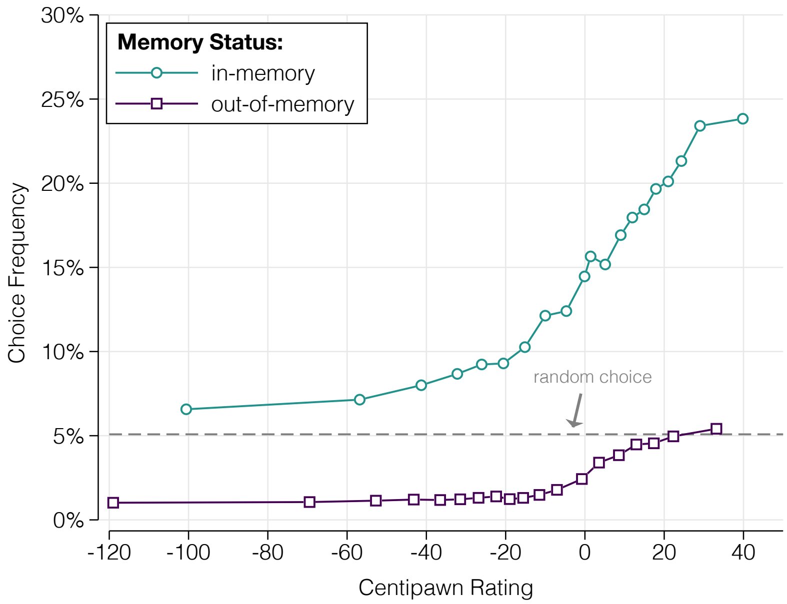
emerge. First, conditional on memory status, choice frequencies are monotonically increasing in quality. Second, comparing moves with similar centipawn ratings, the choice frequency of in-memory alternatives exceeds that of out-of-memory ones by anywhere from five to twenty percentage points. Strikingly, the worst in-memory moves are chosen at a slightly higher rate than the best out-of-memory ones.

To provide a more rigorous test of Prediction 1, we estimate the following linear probability model:

$$Choose_{p,m,s} = \delta\, In\text{-}Memory_{p,m} + \gamma\, Quality_{m,s} + \mu_m + \epsilon_{p,m,s}. \tag{3}$$

Here, $Choose_{p,m,s}$ is an indicator for whether player p selected opening move m in starting position s, $In\text{-}Memory_{p,m}$ denotes whether p holds m in memory, $Quality_{m,s}$ corresponds to the move's objective value, as measured by its Stockfish rating in the relevant starting position, and μ_m is an opening-move fixed effect. The latter accounts for the possibility that some moves may be more likely to be chosen because they accord with popular advice like "open with a center pawn." The coefficient of interest is δ. It is identified by comparing the same opening move across players for whom a particular opening move is and is not in

Figure 6: Choice Frequencies in the Raw Data, by Memory Status



Notes: Figure shows binscatter plots of the relationship between the centipawn ratings and choice frequencies of in- and out-of-memory moves in players' first game of Chess960 as White.

memory, holding the objective quality of the move as well as its latent popularity fixed.

In more granular specifications, we interact the move fixed effect with indicators for the starting position and the size of players' memory, i.e., the number of other in-memory moves. In these specifications, δ is identified by comparing the same opening move in the same starting position across players that hold the same number of other moves in memory. These comparisons fix the quality of the respective move, the set of available alternatives, as well as how many other opening moves a player may recall.[19] They need not, however, involve players that are otherwise similar. In order to address potential concerns about systematic differences across players, we estimate models that also control for player characteristics, such as their prior chess experience and skill. In Appendix E, we additionally leverage within-player variation over time. For now, however, we focus on players' first game of Chess960 and test Prediction 1 using variation in memory across players.

Table 2 presents our estimates of the memory premium. The coefficient in col. (1) indicates that, in the raw data, the average choice frequency of in-memory moves exceeds that of out-of-memory ones by 12.1 p.p. Controlling for the objective quality as well as the latent popularity of a move narrows the difference between both types of alternatives to about 3.8

[19] In Appendix Table AT.1, we replicate our main result controlling for *which* other moves players hold in memory. The implied comparisons come even closer to approximating the definition of the memory premium in Section 2, but yield point estimates and comparative statics that are less precise. Reassuringly, both sets of results are qualitatively equivalent.

Table 2: Memory Premium

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	12.07	3.83	3.83	4.52	4.52	4.51
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.05	0.13	0.16	0.23	0.23	0.23
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

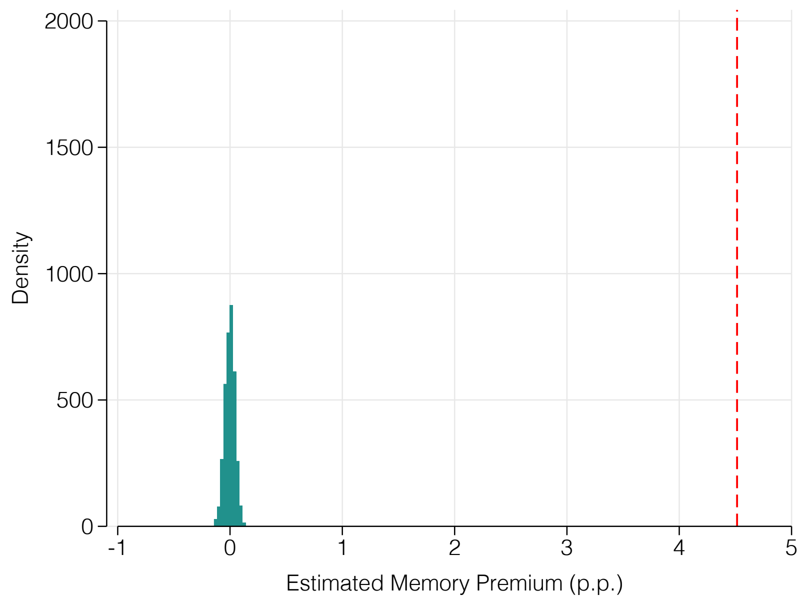
Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (3) by ordinary least squares. The sample consists of all available opening moves in players' first game of Chess960 as White. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player, and are shown in parentheses.

p.p. (col. 2). The point estimate shrinks because opening moves that are generally more likely to be played—e.g., because they help to gain control of the center of the board—are more likely to be in memory. The next column also accounts for the set of available alternatives by including move-by-starting-position fixed effects. Holding the choice set as well as the quality of all available alternatives fixed has no appreciable effect on the coefficient of interest. The specification in col. (4) additionally controls for the size of players' memory, which increases the estimated memory premium to about 4.5 p.p. Columns (5) and (6) add player-level controls, i.e., players' strength ratings in regular chess (interacted with those of their opponents) and their prior experience. Accounting for experience is potentially important because, given our definition of memory, players with longer histories of play tend to have different memories than those with shorter histories. Reassuringly, controlling for systematic differences across players has almost no effect on the estimated coefficients.

In sum, the results in Table 2 support Prediction 1. A given opening move is about 3.8–4.5 p.p. more likely to be chosen when it is in- rather than out-of-memory. Given a mean choice frequency of about 5%, our findings imply a memory premium that is economically very large.

Validation & Robustness. As a first check on the memory premium, we conduct a placebo test. Specifically, we randomly reassign histories of prior play across the individuals in our data. We then re-estimate the regression model in col. (6) of Table 2 based on the newly induced memories. Repeating this process a thousand times, we compare the true point estimate to the distribution of placebo coefficients.

Figure 7: Distribution of Placebo Estimates



Notes: Figure shows the distribution of placebo estimates of the memory premium. Placebo estimates are derived from randomly reshuffling the observed memories across players and estimating the regression model in col. (6) of Table 2 on the same sample of moves. The vertical line corresponds to the original estimate of the memory premium in that column.

Figure 7 shows the results. As one might expect, the median placebo estimate is approximately zero.[20] The remaining placebo coefficients are tightly clustered around the median. As a result, the true point estimate from col. (6) of Table 2 is a clear outlier. It is more than twenty times larger than the largest placebo coefficient, which suggests that players' memory and choice behavior are tightly linked.

In the appendix, we assess the robustness of the memory premium to different definitions of memory. To this end, we say that an opening move is in-memory if, in the last n games, a player has used it at least t times. Each alternative definition is, therefore, characterized by a pair (n, t), where n denotes how many games players can hold in memory (i.e., its length) and t is a usage threshold that governs which moves can be recalled. Using this notation, our workhorse definition of memory corresponds to $n = \infty$ and $t = 1$. For each player in our data, we construct alternative memories for all possible combinations of $n \leq 200$ and $1 \leq t \leq 0.9n$, and for $n = \infty$ with various usage thresholds. We then estimate the memory premium under each alternative definition. The resulting estimates range from about 4.5 to 27 p.p. (see Appendix Figure AF.3). Comparing this range with the coefficients in Table 2

[20]By contrast, the regression model in col. (1) of Table 2 produces placebo estimates that tend to be significantly greater than zero. This observation is consistent with omitted variables bias, which explains why the point estimate decreases upon controlling for centipawn ratings and move fixed effects.

demonstrates that our workhorse definition yields a conservative estimate of the memory premium.

In principle, players may retrieve either individual moves or entire opening strategies from memory. Both possibilities are consistent with our finding that memory shapes which opening moves receive priority. Nonetheless, Appendix D examines the level at which prior experience is represented in memory by asking whether players favor the particular moves they have used before or instead adapt familiar strategies to the new board. Based on the evidence in Appendix Tables AT.2 and AT.3, we conclude that players appear to prioritize individual moves. We also note that, whatever the precise content of players' memories, explanations tied to the particulars of chess cannot account for our experimental results in Section 8.

Heterogeneity. Appendix Table AT.4 explores heterogeneity in the memory premium. While we do not detect meaningful heterogeneity by skill level, we do find that the memory premium is about twice as large for players who have played fewer than the median number of regular chess games on the platform. We also find heterogeneity by relative move quality. The estimated memory premium is significantly higher for opening moves that are close to optimal; yet even suboptimal moves carry a positive premium.[21]

In addition, we investigate heterogeneity across individual moves. Appendix Figure AF.4 plots the estimated memory premium for each of the twenty opening moves in standard chess. All estimates are statistically significant and economically sizeable—especially compared to the mean choice frequencies of the respective moves.[22]

5. Comparative Statics

The memory premium is a quantitatively important empirical regularity that is predicted by our model. The sampling mechanism in the model further predicts that the premium should be larger for alternatives that are easier to retrieve from memory. Drawing on prior work on human memory, we now test this comparative static. Even though the theory abstracts from memory-dependent evaluations, we also ask whether the premium increases when the remembered experiences are "good" rather than "bad."

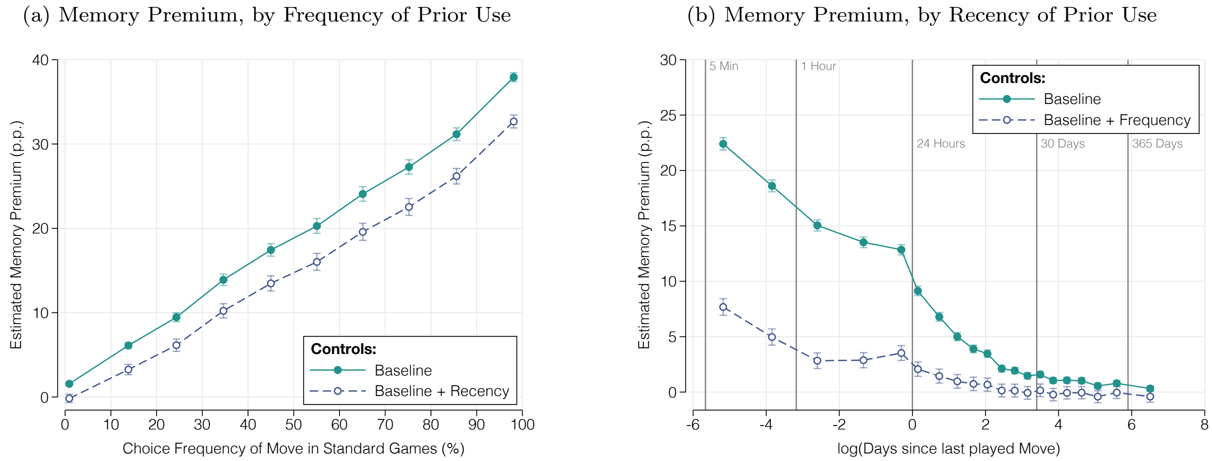
5.1. *Repetition and Decay*

A fundamental property of memory is that it is reinforced by repetition and that it decays over time (see, e.g., Kahana 2012). In the context of our model, these forces imply higher retrieval probabilities—and hence larger memory premia—for moves that players used more

[21] This pattern is predicted by our model if a larger share of close-to-optimal moves is deemed satisfactory.

[22] Appendix Table AT.5 further distinguishes between starting positions in which a given move is and is not close to optimal.

Figure 8: Reinforcement and Decay



Notes: Figure presents estimates of the memory premium by frequency (left panel) and recency (right panel) of prior use of the move in standard chess. Solid lines and markers correspond to estimates based on the regression model and sample in col. (6) of Table 2. Dashed lines and hollow markers correspond to estimates that additionally control for recency (left panel) and frequency (right panel) of prior use. Error bars show 95%-confidence intervals, accounting for clustering by player.

often or more recently in standard chess. We assess this prediction by extending our most saturated regression specification in Table 2 to allow δ to vary with the frequency and recency of prior use.

Figure 8 presents the results. The solid line and markers in the panel on the left show estimates of the memory premium according to the frequency with which the respective player chose the move in regular chess. The dashed line and hollow markers correspond to estimated premia after additionally controlling for how recently the respective move had been played. The evidence in this panel indicates that the memory premium increases greatly with the frequency of prior use.

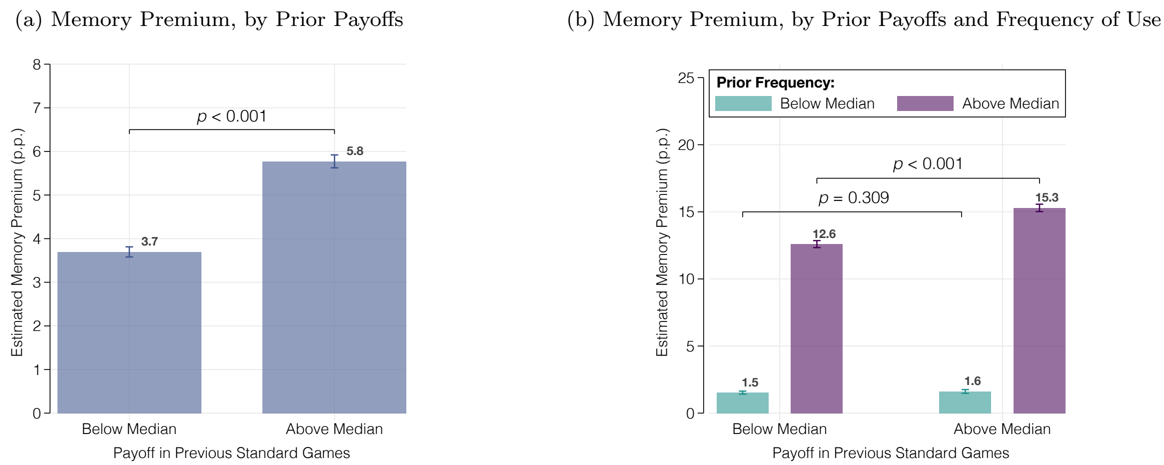
The right panel of Figure 8 shows how the memory premium varies with recency, with and without controlling for the frequency with which a player relied on the respective move in the past. Consistent with our model, the evidence implies that moves that had been played more recently exhibit significantly higher memory premia.[23]

5.2. *Good vs. Bad Memories*

In Figure 9, we differentiate between moves with above- and below-median player-specific payoffs in regular chess. The left panel shows that opening moves with which a player won

[23] Appendix Figure AF.5 replicates the analysis in the right panel of Figure 8 based on "donut definitions" of memory. According to these definitions, a move is said to be in-memory if, and only if, it had been played at least once more than, say, a month prior to a player's first game of Chess960. This robustness check helps to address concerns about short-term serial correlation in the choice of opening moves. For replications of our main result in Table 2 based on donut definitions of memory, see Appendix Tables AT.6–AT.9.

Figure 9: Prior Success



Notes: Figure presents estimates of the memory premium. Panel (a) distinguishes in-memory moves according to whether the player's average payoff from using the move in standard chess is above or below her median average payoff across opening moves. Payoffs are calculated by averaging over the outcomes of all games in which a player previously chose the respective move. Wins are valued at one point, draws at half a point, and losses at zero. Panel (b) further distinguishes according to whether these moves were used more or less frequently than the player's median opening move. Error bars show 95%-confidence intervals, accounting for clustering by player.

more standard games in the past carry a 2.1 p.p. higher memory premium. The right panel additionally differentiates by frequency of prior use. The results suggest that prior success matters only for frequently used moves. The difference in the memory premium is large and statistically significant among frequently used moves, but it vanishes among less frequently used ones.[24]

Unlike the comparative statics with respect to frequency and decay, the effect of prior success is not predicted by our model. The fact that prior success matters suggests that our abstraction from memory-dependent evaluations misses a force that is, in fact, present in the data. At the same time, we note that the difference in the memory premium between frequently and infrequently used moves is much larger than that between moves with "good" and "bad" prior outcomes.

5.3. *Associativity*

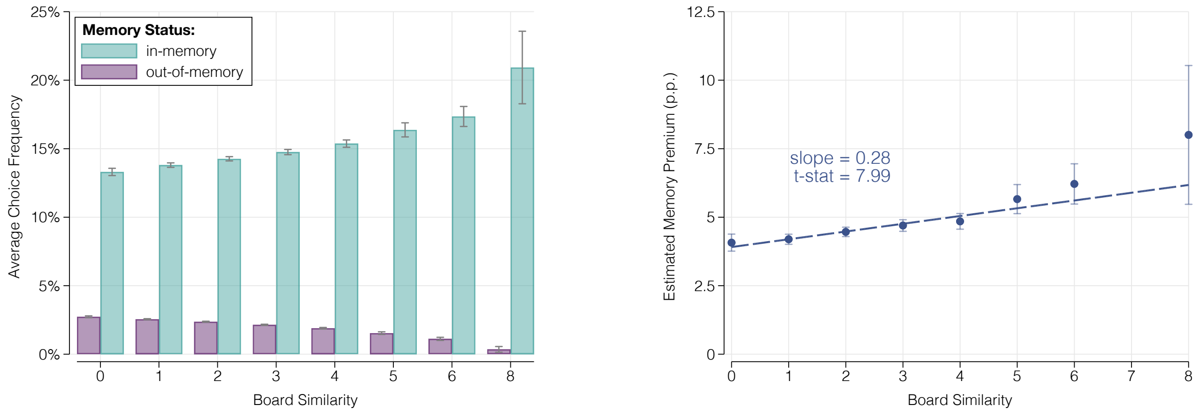
Classic work in psychology establishes that memory is associative in the sense that past experiences are more likely to be recalled when they are related to current environmental cues (see, e.g., Thomson and Tulving 1970; Tulving and Thomson 1973; Smith et al. 1978). Hence, if the memory premium is linked to the probability of recall, then we would expect to estimate higher premia in starting positions with greater resemblance to the standard one.

[24] Appendix Figure AF.6 shows that this pattern is robust to using a more fine-grained split by frequency of prior use. It also shows that the pattern is unchanged when we control for recency.

Figure 10: Memory Premium, by Board Similarity

(a) Raw Data

(b) Conditional on Controls



Notes: Panel (a) shows average choice frequencies of in- and out-of-memory moves by similarity between the starting position in Chess960 and that in standard chess. Similarity is defined as the number of pieces that are placed on the same squares as in standard chess. Panel (b) presents estimates of the memory premium by board similarity. These estimates are based on the regression model and sample in col. (6) of Table 2, allowing for heterogeneity in the memory premium by board similarity. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

To investigate this prediction, we rely on the fact that starting positions in Chess960 are randomized. We define the similarity between a particular position in Chess960 and the standard chess board as the number of pieces that are initially placed on the same squares.[25]

The left panel of Figure 10 partitions the raw data according to this measure of similarity, and shows raw choice frequencies for in- and out-of-memory moves. The right panel shows, for each level of board similarity, regression estimates of the memory premium. Even after controlling for players' prior experience and skill and accounting for the quality of all available alternatives via position-by-move-by-size-of-memory fixed effects, we observe larger memory premia in starting positions that resemble the standard chess board more. The evidence, therefore, suggests that the memory premium is larger when the current choice environment provides stronger cues to recall prior play.

5.4. *Interference*

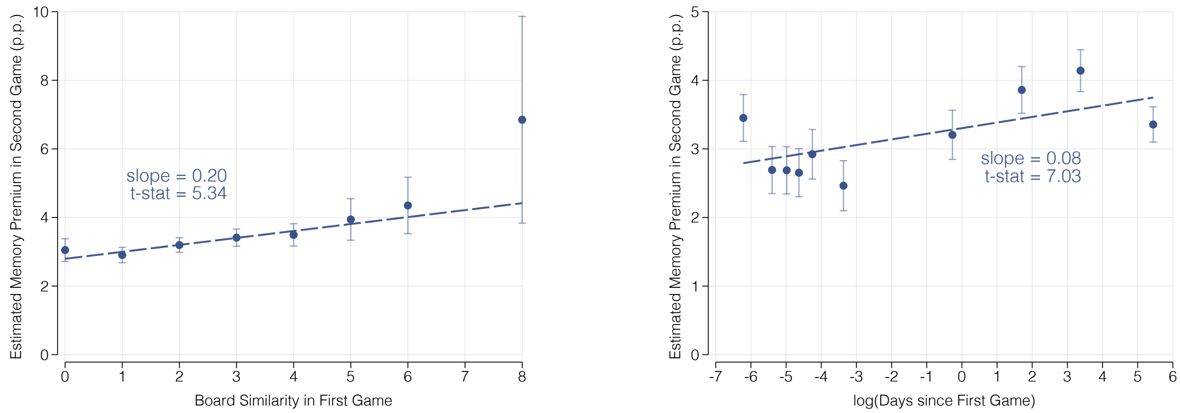
Another property of memory is that experiences compete for recall to the point where new experiences interfere with the retrieval of older ones. To test for interference effects in the memory premium, we take a three-pronged approach. First, we ask whether the memory premium in a player's *second* game of Chess960 as White depends on the similarity between the standard chess board and the starting position in the *first* game. We would expect that memories of the first game of Chess960 are more salient when the starting position was very

[25]Similarity is thus a number between zero (no back-rank piece occupies the same square) and eight (the two positions are identical).

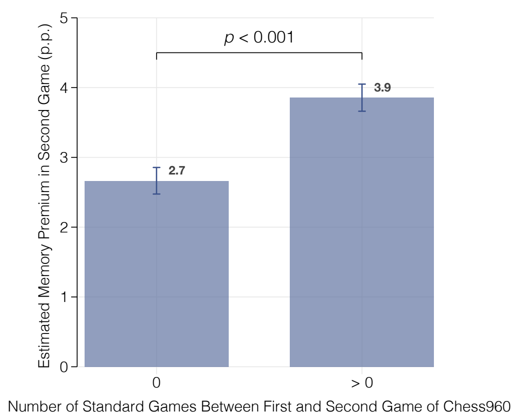
Figure 11: Evidence of Interference

(a) Interference via Similarity

(b) Interference over Time



(c) Interference by Newer Memories



Notes: Panel (a) presents estimates of the memory premium in players' second game of Chess960, by similarity between the starting position in their first game of Chess960 and the starting position in standard chess. Panel (b) shows estimates of the memory premium in players' second game of Chess960, by decile of elapsed time between their first and second Chess960 games. Panel (c) distinguishes between the memory premium in the second game of Chess960 when players have and have not played any standard chess games between their first and second games. The estimates in panels (a) and (b) are based on the regression model in col. (6) of Table 2, allowing for heterogeneity in the memory premium along the relevant dimension. The estimates in panel (c) additionally control for board similarity in the first game of Chess960, as well as the time that elapsed between the first two games. The sample in each panel consists of all available opening moves in players' second game of Chess960, restricting attention to cases in which each of the first two games was rated and played as White. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

different from the standard one. They should thus provide greater competition when it comes to recalling experiences from standard chess. If correct, then the memory premium in the second game ought to be an increasing function of the similarity between the standard starting position and that in the player's first game of Chess960. Second, we ask whether the memory premium in the second game of Chess960 as White increases in the temporal distance to the first one. A long delay after the first game will cause the associated memories to decay. They should, therefore, interfere less with the retrieval of other experiences. Third, we investigate whether the memory premium in the second game increases in the number of standard chess games between the first game of Chess960 and the second one, conditional on board similarity in the first game and the time that has elapsed since. The rationale for our third test is that newer experiences from standard chess will directly interfere with recalling the first game of Chess960 while strengthening the recall of experiences from previous standard games. This, in turn, should increase the memory premium.

Figure 11 provides evidence in support of all three predictions. Measuring board similarity in the same way as before, panel (a) of Figure 11 shows that the memory premium in the second game is significantly smaller when the starting position in the first game of Chess960 was very different from the standard one. Panel (b) presents estimates of the memory premium in the second game as a function of the time between the first game and the second one. The positive slope indicates that the memory premium increases as more time elapses. Panel (c) compares estimates of the memory premium when players have and have not engaged in at least one game of standard chess between their first two games of Chess960.[26] The memory premium is significantly smaller when they have not. The evidence in panels (a)–(c) is, therefore, consistent with the prediction that the memory premium is larger when competing memories are less likely to interfere with recall.

5.5. *Experience*

Finally, we explore how the memory premium evolves as individuals become more familiar with the new choice environment. Our sample for this analysis includes all individuals who played at least fifty games of Chess960 as White, regardless of which color they were assigned in their first game and whether that game was rated. There are 16,731 such players in our data. For each of their first fifty games, we separately estimate the most saturated regression model in Table 2. In doing so, we distinguish between the memory premium in all games (incl. casual ones) and in rated games only. Drawing on all games affords us slightly more statistical power but comes with the drawback of lower stakes.

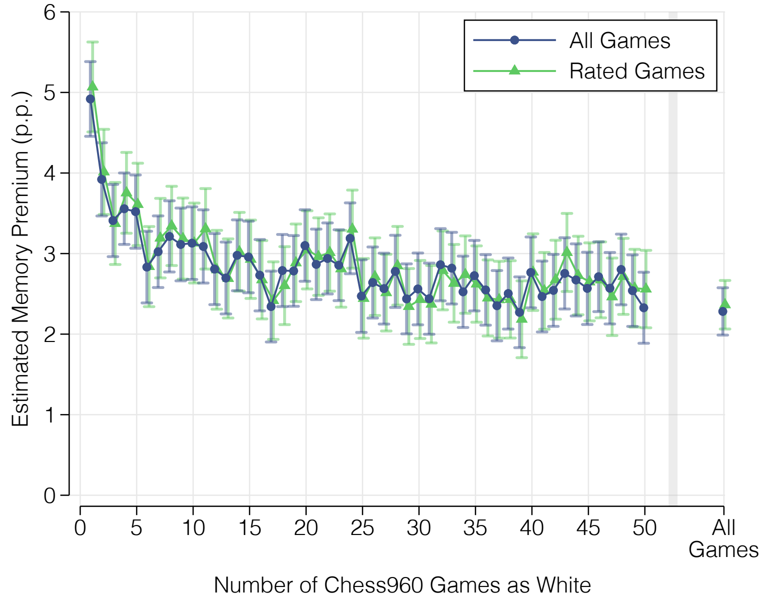
Figure 12 presents the results. Irrespective of whether we consider all games or only rated ones, two patterns stand out. First, despite modest overlap in samples, the estimated memory premium in players' first game of Chess960 as White is quantitatively similar to that in Table 2. Second, the memory premium decreases rapidly as players gain experience; but it does not disappear. In fact, after about twenty-five games, the estimated memory premium stabilizes around 2.5 p.p., or about 50% of the average choice frequency in this setting.[27]

In Appendix E, we draw on the same sample of players and games to provide further evidence on the link between memory and choice behavior. Using both an RD and a difference-in-differences design, we show that the choice frequency of an opening move in Chess960 increases significantly after that move has been played for the first time in a game of regular chess. The evidence from these research designs is, therefore, consistent with the idea that retrievability from memory has a causal effect on choice.

[26] About 60% of individuals in our data play zero standard games between their first two games of Chess960.

[27] In Appendix Figure AF.7, we replicate the analysis in Figure 12 based on a definition of memory that is based on players' choices in prior games of Chess960. Although we estimate a significantly larger premium for this definition of memory, we observe a qualitatively similar decline as players gain experience.

Figure 12: Dynamics of the Memory Premium



Notes: Figure shows two series of estimates of the memory premium in each of players' first fifty games of Chess960 as White. The first series restricts attention to rated games, whereas the second series considers both rated and unrated games. All estimates are based on the regression model in col. (6) of Table 2. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

6. Potential Mechanisms

The memory premium and its comparative statics are predicted by our model because the DM preferentially considers alternatives that readily come to mind. As discussed in Section 2, however, the same choice patterns can also be explained by models with memory-dependent evaluations. In particular, utility maximization can generate the memory premium, as well as its comparative statics, if the subjective utility assigned to an alternative increases in the probability that the alternative is retrieved from memory. Choice data alone can, therefore, not establish that memory affects decision-making by guiding consideration.

To distinguish between different potential mechanisms, we proceed in two steps. First, we ask whether the DM engages in full or partial consideration, i.e., whether she evaluates all available alternatives or only a subset of them. Second, focusing on partial consideration procedures, we ask whether memory affects only the subjective evaluation of alternatives or also which alternatives are evaluated in the first place.

6.1. *Full vs. Partial Consideration*

We rely on deliberation times to differentiate between full- and partial-consideration procedures. Under full consideration, the DM evaluates every alternative, regardless of which option she ultimately selects. Given that the amount of cognitive work is fixed, deliberation times should not systematically vary with the memory status of the chosen alternative.

If, however, the DM evaluates only a subset of the available alternatives, then that set may depend on which alternative she selects. In our satisficing model, easy retrieval from memory moves alternatives, on average, earlier in the search order. Conditional on being chosen, a satisfactory in-memory alternative is, therefore, reached after fewer evaluations than a satisfactory out-of-memory option. Together with the assumption that in-memory alternatives are no slower to evaluate, this generates the memory speedup in Prediction 2.

We test this prediction in Section 7. We find that deliberation times are meaningfully shorter when players choose an in-memory option. Under our maintained assumptions on deliberation time, this finding is inconsistent with full-consideration procedures in general and utility maximization in particular.

We note, however, that the memory speedup is not unique to our model. Other models with partial consideration may generate the same pattern. A leading example is choice from consideration sets. In this class of models, the DM first draws a subset of alternatives and then chooses the best option from that set (see, e.g., Manzini and Mariotti 2014). Such procedures can generate the memory speedup if certain features of the consideration set correlate with the memory status of the chosen alternative.[28] The memory speedup, therefore, points to partial consideration, but it is not diagnostic of the mechanism through which memory affects choice behavior within that class of procedures.

6.2. *Memory-Dependent Utility vs. Memory-Dependent Consideration*

In particular, a partial-consideration procedure may select or order alternatives independently of memory while allowing memory to shape their subjective evaluations. For example, a satisficer may sample alternatives in a memory-independent way yet be more likely to stop at an in-memory option because easy retrieval raises the alternative's perceived value. Such a procedure can generate both the memory premium and the speedup even though memory retrieval does not affect consideration.

Thus, even in a partial-consideration procedure, memory can affect choice either through the utility or the consideration channel. Both mechanisms might even operate simultaneously. To provide direct evidence for the idea that memory determines which alternatives the DM considers, Section 8 presents results from a controlled laboratory experiment that shuts down the utility channel and makes the consideration process observable. In the experimental data,

[28] For instance, conditional on choosing an in-memory alternative, the DM may have drawn a smaller consideration set; or the consideration set may contain a greater share of in-memory options, which may be easier to evaluate.

in-memory alternatives are more likely to be considered and, crucially, more likely to be considered first—even when the most promising in-memory alternative has a lower expected value than the average out-of-memory option. These results imply that memory guides the consideration process. They thus speak to the mechanism through which memory affects choice.[29]

The experiment also helps to distinguish between different choice procedures with partial consideration. In sequential procedures such as satisficing, the information revealed during the evaluation process determines whether search continues; in canonical choice-from-consideration-sets models, the alternatives to be evaluated are determined prior to the DM assessing any one of them. The experimental data are inconsistent with the latter.

7. The Memory Speedup

To distinguish between full- and partial-consideration procedures, we turn to deliberation times. As explained above, full-consideration procedures imply no difference in deliberation time as a function of which alternative is ultimately selected. By contrast, our satisficing model and other partial-consideration procedures predict that in-memory alternatives are chosen faster than out-of-memory ones.

One complication in testing this prediction is that Lichess does not record timestamps for individual moves. The data contain only players' clock readings, and the clock does not start until after White and Black have both executed their first move. Since we cannot recover deliberation times for the opening move of White, we analyze its second and third moves instead. For these decisions, the clock readings in our data reveal how long players deliberate between observing Black's preceding move and making a choice of their own.

Analyzing later moves requires us to refine our definition of memory. We classify a given second or third move as in-memory if, in standard chess, the player previously selected that move after encountering the same sequence of preceding moves. In the appendix, we document that the memory premium extends beyond the first move. In fact, the estimated premium is even larger for moves two and three (see Appendix Table AT.10).[30]

Figure 13 shows that, in the raw data, deliberation times are shorter when the chosen move is in memory. The left panel in this figure plots the empirical CDFs of deliberation times in players' first game of Chess960. We observe that, for both the second and third moves, the distribution of deliberation times for in-memory choices is first-order stochastically dominated by that for out-of-memory ones. The right panel shows raw differences in mean deliberation

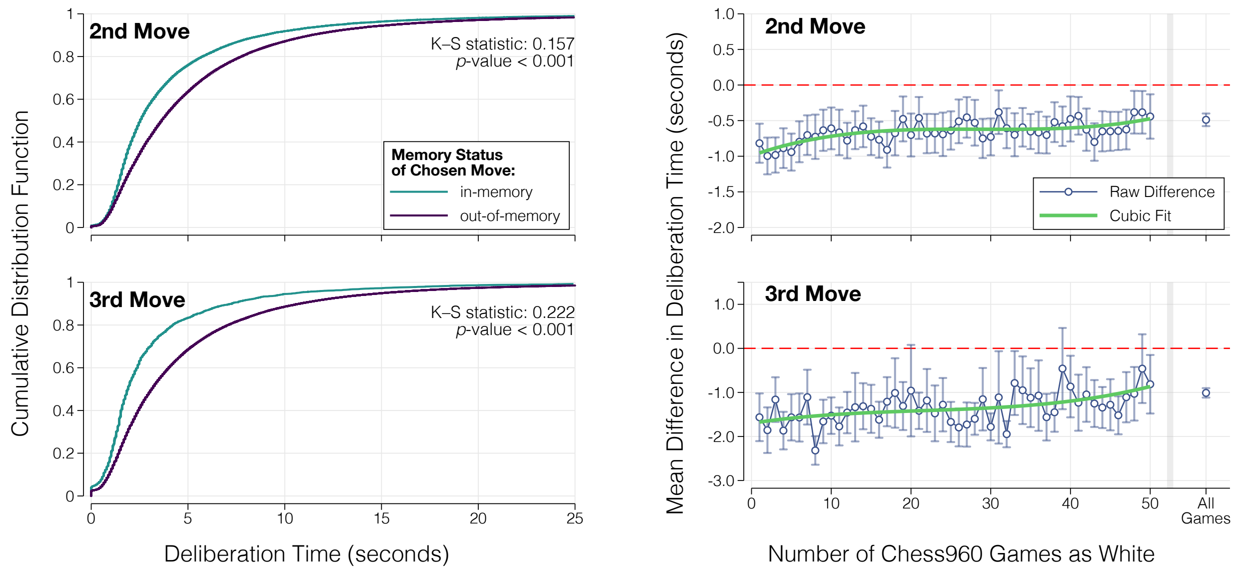
[29] To be clear, the experimental results do not rule out memory-dependent evaluations in other settings.

[30] We have also explored alternative definitions of memory, such as classifying a move as in-memory if the player had previously chosen that move irrespective of the preceding sequence of moves. Reassuringly, such alternative definitions produce qualitatively similar results.

Figure 13: Deliberation Times in the Raw Data, by Memory Status of Chosen Move

(a) CDF of Deliberation Times in the First Game of Chess960

(b) Mean Difference Between In- and Out-of-Memory Choices



Notes: Panel (a) plots empirical CDFs of deliberation times for in- and out-of-memory choices in players' first game of Chess960 as White. For both second and third moves, the reported Kolmogorov-Smirnov statistics are from tests of the null hypothesis of equal distributions. Panel (b) shows the difference in mean deliberation time between in- and out-of-memory choices for each of players' first fifty games of Chess960 as White. The sample in this panel is restricted to players who played at least fifty games of Chess960 as White. Deliberation time is measured from Black's preceding move to White's response. A given move is classified as in-memory if the player previously selected that move after encountering the same sequence of preceding moves in standard chess.

time between in- and out-of-memory choices for players' first fifty games of Chess960. The vast majority of these differences are negative and statistically significant. The fitted curves suggest that the memory speedup diminishes over time—consistent with the idea that the role of previous memories declines as players gain experience with the new environment—but it does not disappear.

To quantify the memory speedup while accounting for systematic differences in the difficulty of the choice problem, we estimate variants of the following econometric model:

$$\log(Deliberation\ Time_{p,g,b}) = \beta\, Choose\ In\text{-}Memory\ Move_{p,g,b} + \gamma X_{p,g} + \lambda_b + \varepsilon_{p,g,b}. \tag{4}$$

Here, $Choose\ In\text{-}Memory\ Move_{p,g,b}$ is an indicator for whether player p executed an in-memory move from board position b in game g, $X_{p,g}$ is a vector of player- and game-specific covariates, including p's previous experience playing standard chess and the time controls of the match, and λ_b denotes a board-position fixed effect. In this baseline specification, the coefficient of interest, β, is identified by comparing in- and out-of-memory choices from the same position. Because both the player and the selected move may vary within a board position, these comparisons pool variation across chosen alternatives and across memories.

Table 3: Memory Speedup

A. Second Move

	Log Deliberation Time					
	(1A)	(2A)	(3A)	(4A)	(5A)	(6A)
Chosen Move In-Memory	–0.31	–0.19	–0.11	–0.10	–0.09	–0.10
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.03)
Sample	First Game	First Game	All Games	All Games	All Games	All Games
Board, Move & Player Fixed Effects:						
Board Position	Yes	No	Yes	No	Yes	No
Board Position × Move	No	Yes	No	Yes	No	No
Player	No	No	No	No	Yes	No
Player × Board Position	No	No	No	No	No	Yes
Additional Fixed Effects:						
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Own × Opponent Strength Decile	Yes	Yes	Yes	Yes	Yes	Yes
Standard Chess Experience	Yes	Yes	Yes	Yes	Yes	Yes
Chess960 Experience	No	No	Yes	Yes	Yes	Yes
Mean Deliberation Time (seconds)	5.29	5.29	3.47	3.47	3.47	3.47
R^2	0.38	0.51	0.43	0.50	0.51	0.72
N	116,998	116,998	3,353,446	3,353,446	3,353,446	3,353,446

B. Third Move

	Log Deliberation Time					
	(1B)	(2B)	(3B)	(4B)	(5B)	(6B)
Chosen Move In-Memory	–0.27	–0.23	–0.15	–0.11	–0.13	–0.09
	(0.05)	(0.08)	(0.01)	(0.01)	(0.01)	(0.04)
Sample	First Game	First Game	All Games	All Games	All Games	All Games
Board, Move & Player Fixed Effects:						
Board Position	Yes	No	Yes	No	Yes	No
Board Position × Move	No	Yes	No	Yes	No	No
Player	No	No	No	No	Yes	No
Player × Board Position	No	No	No	No	No	Yes
Additional Fixed Effects:						
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes
Own × Opponent Strength Decile	Yes	Yes	Yes	Yes	Yes	Yes
Standard Chess Experience	Yes	Yes	Yes	Yes	Yes	Yes
Chess960 Experience	No	No	Yes	Yes	Yes	Yes
Mean Deliberation Time (seconds)	5.05	5.05	3.42	3.42	3.42	3.42
R^2	0.57	0.67	0.55	0.61	0.61	0.77
N	117,415	117,415	3,695,051	3,695,051	3,695,051	3,695,051

Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (4) by ordinary least squares. The upper panel focuses on second moves, while the lower panel presents estimates for third moves. In each panel, the sample in the first two columns is restricted to players' first game of Chess960 as White. The samples in the remaining columns include second and third moves in any games of Chess960 as White for players who have played at least fifty such games. Deliberation time is measured from Black's preceding move to White's response. Choices with a deliberation time of zero are excluded. The corresponding moves were preselected by White, i.e., before Black even chose their move. For estimated mean differences including these choice problems, see Figure 13. Standard errors are clustered by player, and reported in parentheses.

In more saturated specifications, we additionally interact λ_b with a fixed effect for the selected move. These specifications hold both the choice problem and the chosen alternative fixed. The identifying variation thus closely corresponds to our notion of a "speedup across memories," which is stated in Prediction 3 (see Appendix A). Exploiting the panel nature of our data, we also estimate models that add a player fixed effect in order to control for any systematic differences in the speed of decision-making across individuals. Our most saturated specification includes a player-by-board-position fixed effect. In this model, all identifying variation comes from players that choose different moves as they encounter the same choice problem more than once. By holding both the player—and thus, to a first approximation, her memory of opening moves—and the choice problem fixed, this variation provides the closest empirical counterpart to the memory speedup in Prediction 2.

Table 3 presents the resulting estimates for moves two (upper panel) and three (lower panel). Within each panel, columns (1) and (2) focus on players' first game of Chess960 as White. The estimates therein imply that in-memory moves are chosen 17% to 27% faster than out-of-memory ones.[31] Columns (3) and (4) expand the sample to include all games of Chess960 as White. Mirroring the evidence in Figure 13, the average memory speedup decreases to about 9% to 14% as individuals gain experience with the unfamiliar environment. The last two columns exploit variation within players. Column (5) includes separate player and board-position fixed effects, while column (6) includes player-by-board-position fixed effects. Controlling for systematic differences across choice problems as well as players, our estimates of the memory speedup range from 8% to 12%.

Broadly summarizing, the results in Table 3 are consistent with the predicted speedup. In-memory moves are chosen significantly faster than out-of-memory ones.

Comparative Statics. Since a higher retrieval probability tends to move an alternative earlier in the search order, we empirically investigate whether the memory speedup is larger for alternatives that are easier to recall. Appendix Table AT.11 presents evidence in support of this prediction. Like the memory premium, the speedup is larger for moves that players have used more frequently and more recently in standard chess, in starting positions that more closely resemble the standard chess board, and when competing memories create less interference. The cognitive forces that govern recall thus mediate both the memory premium and the memory speedup.

8. Experiment

The memory speedup points to memory-dependent partial-consideration procedures. Even within this class of procedures, however, memory may operate through the subjective evaluation of alternatives rather than through the consideration process itself. Distinguishing between these channels is difficult. Standard data capture outcomes but provide little information on how decision-makers arrive at their choices, and models with memory-dependent utility

[31] Note, $e^{-0.19} \approx 0.83$ and $e^{-0.31} \approx 0.73$.

can accommodate a wide range of behavior. To overcome these challenges, we conducted a pre-registered laboratory experiment that shuts down the utility channel and makes the consideration process directly observable.[32]

8.1. *Experimental Design*

Setup. In the experiment, subjects complete twenty rounds of a simple choice task. We describe the task as choosing one out of six "companies" for a hypothetical investment. Each company is identified by a unique name and color, and appears in only one round. Subjects' bonus payments are determined by the value of the company they choose in one randomly selected round. A company's value depends on its "profit" and "growth":

$$\text{Value} = \text{Profit} \times \text{Growth Factor}.$$

Profit levels are integers between 2 and 10, while growth factors are two-digit decimals between 0.5 and 1.5. Both are distributed uniformly and independently across companies, all of which is known to participants.

Each round of the choice task consists of two stages, separated by a brief wait phase. In the first stage, which we call the "research stage," subjects are shown three of the six companies in their choice set, together with information on these companies' "number of stores," "revenue per store," and "total expenses." Subjects know that which companies appear in the research stage is entirely random. Based on the provided information, subjects are instructed to calculate the three companies' profits:

$$\text{Profit} = (\text{Stores} \times \text{Revenue per Store}) - \text{Total Expenses}.$$

Note, a company's profit corresponds to its expected value to the participant prior to learning the growth factor.

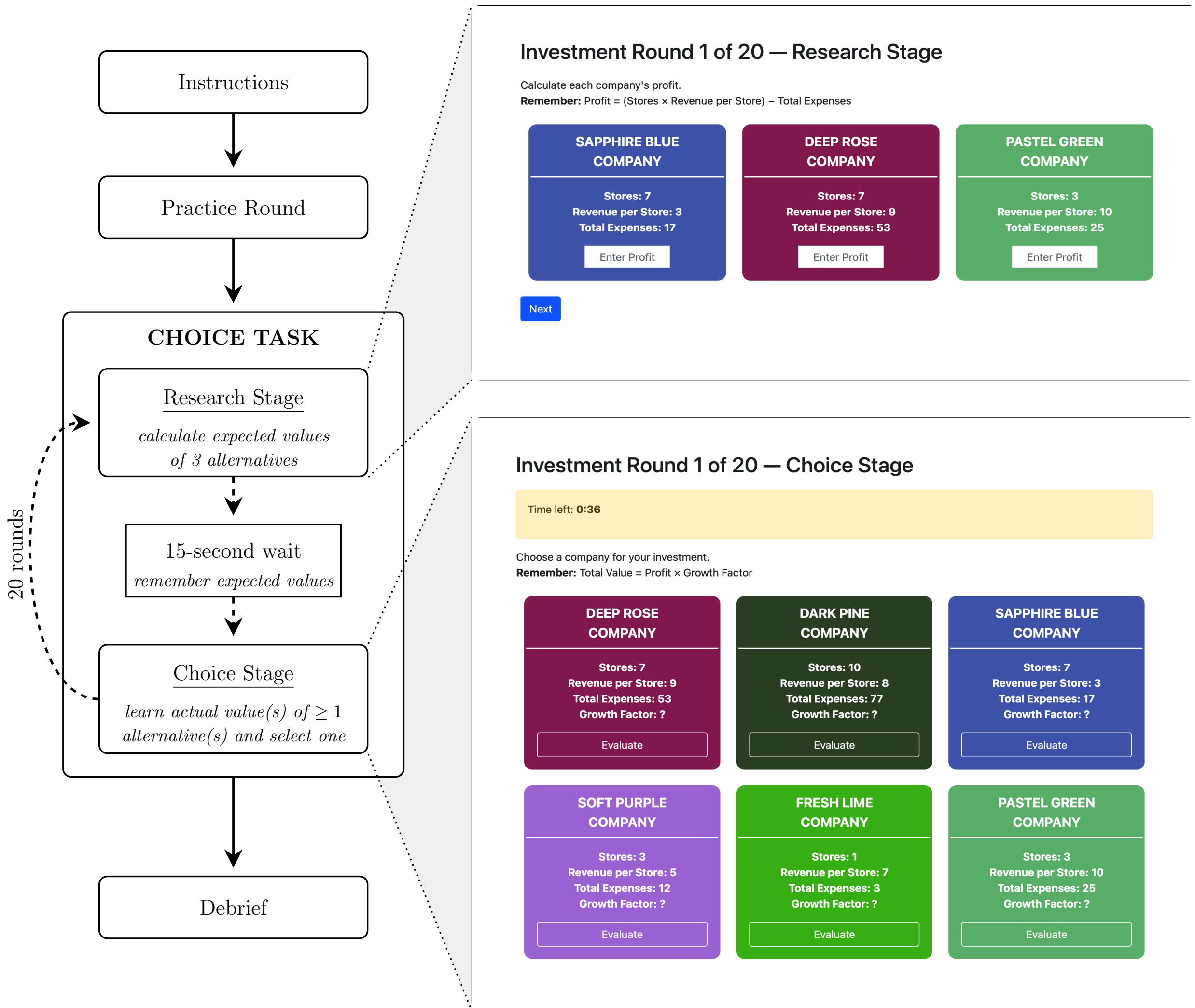
Upon correctly calculating profits, subjects wait for fifteen seconds before proceeding to the "choice stage."

In the choice stage, subjects are shown all six companies simultaneously.[33] At this point, subjects have enough information to calculate every company's profit, but they do not yet know the relevant growth factors. To see the growth factor of a particular company, subjects need to click on that company's "evaluate" button. Subjects may elicit the growth factors

[32] The preregistration for this experiment is available in the OSF Registries at `https://osf.io/8tkd4`. Figure 15 in the main text presents the results from the pre-registered analyses. Additional analyses that were not pre-registered are shown in Figure 16 and Appendix Tables AT.12–AT.16.

[33] The positioning of companies on the screen is random.

Figure 14: Experimental Design



Notes: Figure shows a high-level summary of our experimental design, together with screenshots from each stage of the choice task. The upper screenshot depicts the "research stage" of each round, in which participants are asked to calculate and memorize the expected value of three alternatives. The lower screenshot depicts the "choice stage." In the choice stage of a given round, subjects click on alternatives' "evaluate button" to reveal additional information about the respective option before making a final choice. The choice set always includes the three alternatives that subjects encountered during the "research stage" of the same round, as well as three new ones. In either stage, the position of alternatives on the screen is randomly determined.

of as many companies as they wish, and they can do so in any order. The only constraints are that a company must be evaluated before it can be chosen, and that a choice must be made before the time limit for the choice stage expires. The time limit varies between ten and forty-five seconds, and is distributed uniformly and independently across rounds.[34]

[34] Participants do not to make a timely choice in about 3.9% of problems. The analysis in the main text excludes all choice tasks for which this is the case. Appendix Table AT.13 shows that the results are robust to including these observations. In addition, we have conducted a follow-up experiment in which we did not impose a binding time constraint. Again, the results are qualitatively equivalent (see Appendix Table AT.15).

Design Rationale. Figure 14 provides a high-level summary of the experimental design, together with screenshots from both stages of the choice task. We settled on this design because we wish to isolate the role of the consideration channel in generating the memory premium. By *inducing* values and beliefs, our design eliminates any potential differences in the subjective utility of alternatives. The utility channel, therefore, cannot account for the emergence of a memory premium in the experiment.

Another key feature of the design is that it allows us to monitor the choice process. Building on Johnson et al. (2002), we track which alternatives participants evaluated and in what order. We can, therefore, examine whether in-memory alternatives receive priority in consideration.

Finally, our two-stage design randomly induces memories of some alternatives in the choice set but not others. By instructing participants to remember the alternatives in the research stage and by restricting the delay before the choice stage to only fifteen seconds, we ensure that subjects are able to retrieve and rely on the relevant memories—if they wish. The results below should thus be interpreted as speaking to memory dependence in choice behavior under idealized conditions, when imperfect recall is unlikely to be a first-order issue.

Implementation. The experiment took place on a custom-built website on July 12, 2025. We recruited 501 participants from Prolific, promising a $7.50 base payment plus a performance-based bonus of up to $15. The median participant earned a total of $16.18 and spent about 34.5 minutes on the experiment. For additional details and for a copy of the instructions to participants, see Appendix F.[35]

8.2. *Experimental Results*

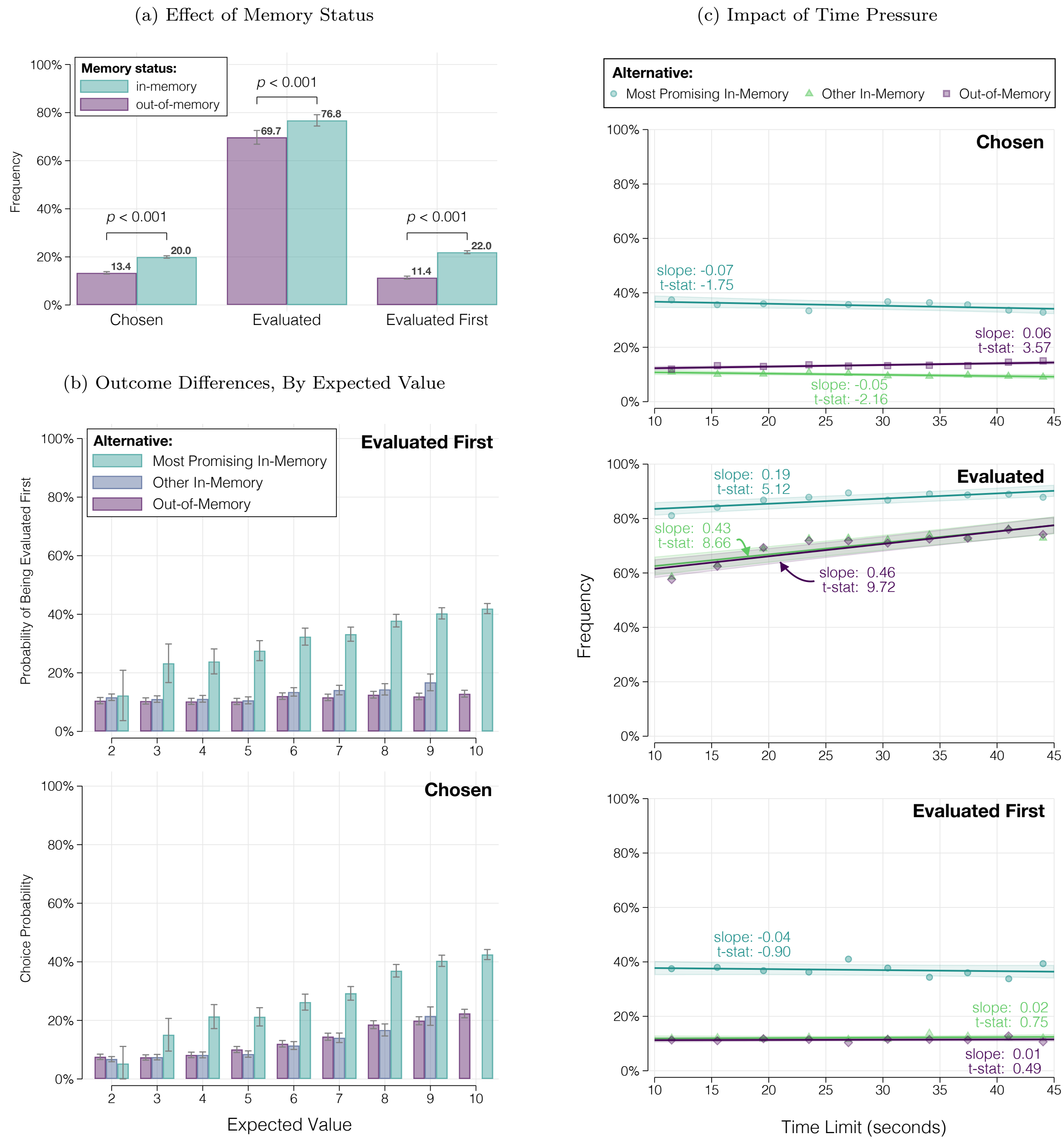
Figure 15 visualizes the key results of the experiment. The top-left panel shows that in-memory alternatives are, on average, 6.6 p.p. more likely to be chosen than out-of-memory ones, even though there are no systematic differences in either expected or realized values. Relative to a mean choice frequency of 16.7%, the memory premium in this environment is economically large.

The top-left panel also shows that, on average, in-memory options are 7.1 p.p. more likely to be evaluated than out-of-memory ones; and they are 10.6 p.p. more likely to be evaluated first. In other words, subjects are not only more likely to consider alternatives that they have previously encountered, but they tend to start by considering these alternatives.

The results in the bottom-left panel of Figure 15 demonstrate that these findings are driven by the "most-promising" of the in-memory alternatives, i.e., the one with the highest expected

[35]More than 98% of participants report that the instructions they received were clear. For summary statistics on the participant population and selected outcomes, see Appendix Table AT.12.

Figure 15: Experimental Results



Notes: Figure presents results from our experiment. Panel (a) compares mean outcomes for alternatives that appeared in the research stages of the choice task and were, therefore, in-memory with those that were not. Panels (b) and (c) distinguish between the in-memory alternative with the highest expected value, the remaining two in-memory alternatives, and out-of-memory options. Panel (b) compares mean outcomes for each type of alternative, by expected value. Panel (c) shows how outcomes depend on the time constraint in the choice stage. Error bars and shaded regions correspond to 95%-confidence intervals, accounting for clustering by participant. For results that control for the display position and that further distinguish between the second- and third-most promising in-memory alternative, see Appendix Tables AT.13 and AT.14.

value. Participants are substantially more likely to evaluate this alternative first, even when its expected value is only three, four, or five—and therefore below that of a randomly sampled out-of-memory option. Even at these values, the most-promising in-memory alternative is evaluated first in about 23–28% of choice problems.

The right panel of Figure 15 investigates how the observed behavior depends on the time constraint. Three patterns stand out: (i) Subjects' tendency to evaluate the most promising in-memory alternative first does not depend on the time limit. (ii) When given more time, subjects evaluate a greater number of in- and out-of-memory alternatives, but especially the latter. (iii) They thus become somewhat more likely to ultimately select an out-of-memory option. Importantly though, the evidence suggests that memory affects the consideration process, irrespective of how fast choices need to be made.[36]

Appendix Table AT.16 investigates response times. The evidence therein indicates that participants complete the choice stage 8–13% faster when they select an in- rather than an out-of-memory alternative—confirming the memory speedup. Consistent with the mechanism in our model, fewer evaluations explain a significant share of this difference.

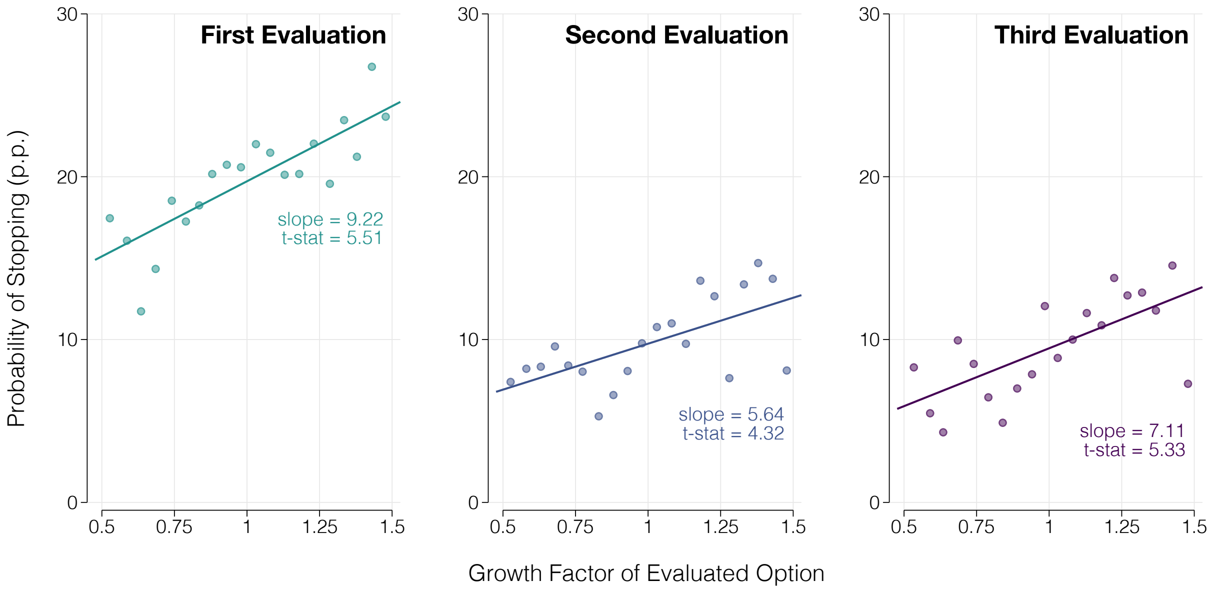
8.3. *Discussion*

There are two natural explanations for why participants begin by evaluating the most-promising in-memory alternative. One is that this alternative tends to have a high expected and, therefore, realized value. The other is that beginning with an in-memory alternative economizes on mental effort because its "profit" has already been calculated. Both explanations are plausible when the expected value of the remembered alternative is sufficiently high. Neither, however, explains why participants continue to prioritize it when its expected value is only three or four. At these values, participants could obtain a higher expected payoff by considering a randomly selected out-of-memory alternative first, without performing any additional calculations. The persistence of this pattern at low expected values points to an independent procedural role for memory in the consideration process.

The experimental data can also shed some light on the specific procedure that governs partial consideration. Two leading candidates are satisficing and maximization from consideration sets. The key difference between these procedures is that satisficing is fully sequential. At each stage of the search process, the decision-maker relies on the newly revealed information to decide whether or not to continue. Under maximization from consideration sets, however, the set of alternatives that she evaluates is determined before the evaluation process begins.

[36] As an additional robustness check, we conducted a follow-up experiment without a binding time constraint in the choice stage ($N = 201$). The results from this experiment demonstrate that subjects prioritize the most promising in-memory alternative even when they do not have to make a decision quickly (see Appendix Table AT.15).

Figure 16: Evidence of Sequential Decision-Making



Notes: Figure shows binscatter plots of the relationship between alternatives' realized growth factors and the probability that participants stop searching after the first, second, and third evaluation. Stopping after a given evaluation is defined as making a choice without evaluating another alternative. Markers correspond to averages within twenty equal-sized bins. Solid lines show OLS fits based on the underlying choice-problem-level data. The sample in each panel consists of all choice problems in which participants reached the respective number of evaluations.

In the context of our experiment, maximization from consideration sets therefore predicts that the decision to stop searching should not depend on alternatives' realized growth factors.

The evidence in Figure 16 contradicts this prediction. The three panels in this figure plot the probability of stopping search after the first, second, and third evaluation against the realized growth factor of the respective alternatives. In each panel, we observe a strong positive relationship. Participants are significantly more likely to stop after evaluating an alternative with a higher growth factor. This pattern is consistent with satisficing but not maximization from consideration sets.

Taken together, our results point to a procedural explanation for memory dependence in choice. Memory affects which alternatives decision-makers consider and in what order. Although our findings establish a role for memory in the consideration process, it is important to note that they do not rule out memory-dependent evaluations. In other settings, subjective utility differences may reinforce both the memory premium and the memory speedup.

9. Conclusion

We propose that memory guides choice in unfamiliar environments by shaping which alternatives decision-makers consider. We formalize this idea in a model of satisficing in which sampling for evaluation favors in-memory alternatives. Our model predicts two key patterns in choice behavior. First, in-memory alternatives are chosen more frequently than out-of-memory ones, even when the latter are objectively better. We dub this phenomenon the "memory premium," and confirm its existence and its comparative statics in a unique observational data set. Second, we test and confirm the prediction that choice behavior exhibits a "memory speedup." That is, conditional on being chosen, in-memory alternatives are selected more quickly than out-of-memory ones.

Although these findings follow naturally from the memory-dependent consideration channel in our model, they do not by themselves establish that mechanism. In order to close this gap, we conduct a tightly controlled, pre-registered laboratory experiment. By randomly assigning memory status and inducing values and beliefs, the experimental design eliminates any systematic differences in the subjective utility of in- and out-of-memory alternatives. It also allows us to observe which options participants consider, in what order, and when they stop searching. We find that in-memory alternatives are more likely to be considered and more likely to be considered first. We also find that stopping is more likely after encountering a high-value alternative. The experimental data thus provide direct evidence that memory guides the consideration process, while the stopping behavior is consistent with satisficing.

Taken together, our findings suggest a broader view of how past experiences influence decision-making in unfamiliar environments. Memory may not only affect *how* alternatives are evaluated but also *which* options are considered in the first place. By directing attention toward familiar alternatives, memory can shorten the decision process and reduce the number of options that need to be considered. Whether such a shortcut is useful depends on the value of familiar alternatives in the new environment. When familiar alternatives tend to perform well across settings, relying on memory to structure consideration can help to economize on cognitive effort without materially reducing choice quality. By contrast, when past experience is a poor guide to current payoffs, memory-dependent consideration can crowd out exploration of better options and, thereby, slow down learning.

Our findings may also have implications for decision-making in familiar environments. Bordalo et al. (2025) highlight a basic advertising puzzle: Why do firms continue to promote well-known, intrinsically similar brands when consumers are already well informed? In their model, ads act as retrieval cues. By prompting recall of pleasurable consumption experiences, advertisements raise the predicted utility of consuming the associated brand and, thereby, increase demand. Our model of satisficing with memory-dependent consideration suggests a complementary mechanism. Even if advertisements leave consumers' subjective evaluations unchanged, they may help to keep advertised brands top of mind, move them earlier in the search order, and thus make them more likely to be considered.[37] Advertising is only one

[37]To be clear, we do not argue that memory affects choice *only* through consideration. The utility and

application of this mechanism. By changing what comes to mind, reminders, comparison tools, and similar interventions can alter what is considered and ultimately chosen, even without affecting beliefs or preferences.

Looking beyond particular applications, our results speak to a basic question about choice behavior: If decision-makers do not attend to all available alternatives, what determines which options they consider? Our evidence suggests that memory retrieval is one important part of the answer. A natural next step is to identify other forces that shape consideration, understand how they interact with memory, and determine when the resulting choice process facilitates or impedes adaptation.

References

Afrouzi, H., S. Kwon, A. Landier, Y. Ma, and D. Thesmar (2023). Overreaction in Expectations: Evidence and Theory. *Quarterly Journal of Economics 138*(3), 1713–1764.

Azeredo da Silveira, R., Y. Sung, and M. Woodford (2024). Optimally Imprecise Memory and Biased Forecasts. *American Economic Review 114*(10), 3075–3118.

Azeredo da Silveira, R. and M. Woodford (2019). Noisy Memory and Over-Reaction to News. *AEA Papers and Proceedings 109*, 557–561.

Bordalo, P., G. Burro, K. Coffman, N. Gennaioli, and A. Shleifer (2025). Imagining the Future: Memory, Simulation, and Beliefs. *Review of Economic Studies 92*(3), 1532–1563.

Bordalo, P., G. Burro, N. Gennaioli, G. Nacamulli, and A. Shleifer (2025). Ads as Cues. NBER Working Paper No. 34387.

Bordalo, P., J. Conlon, N. Gennaioli, S. Kwon, and A. Shleifer (2023). Memory and Probability. *Quarterly Journal of Economics 138*(1), 265–311.

Bordalo, P., N. Gennaioli, G. Lanzani, and A. Shleifer (2026). A Cognitive Theory of Reasoning and Choice. *Quarterly Journal of Economics 141*(3), 1921–1963.

Bordalo, P., N. Gennaioli, and A. Shleifer (2020). Memory, Attention and Choice. *Quarterly Journal of Economics* (3), 1399–1442.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *Review of Economic Studies 91*(6), 3253–3285.

Calonico, S., M. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Carlson, J. and P. McAfee (1983). Discrete Equilibrium Price Dispersion. *Journal of Political Economy 91*(3), 480–493.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The Effect of Minimum Wages on Low-Wage Jobs. *Quarterly Journal of Economics 134*(3), 1405–1454.

Charles, C. (2022). Memory and Trading. mimeographed, London School of Economics.

consideration channels may operate simultaneously and reinforce each other.

Charles, C. (2025). Memory Moves Markets. *Review of Financial Studies 38*(6), 1641–1686.

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

Choi, M. and L. Smith (2024). The Economics of Web Search. mimeographed, University of Wisconsin–Madison.

De Los Santos, B., A. Hortaçsu, and M. Wildenbeest (2012). Testing Models of Consumer Search Using Data on Web Browsing and Purchasing Behavior. *American Economic Review 102*(6), 2955–2980.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2021). Exposure to Grocery Prices and Inflation Expectations. *Journal of Political Economy 129*(5), 1615–1639.

Eliaz, K. and R. Spiegler (2016). Search Design and Broad Matching. *American Economic Review 106*(3), 563–586.

Enke, B., F. Schwerter, and F. Zimmermann (2024). Associative Memory, Beliefs and Market Interactions. *Journal of Financial Economics 157*, 103853.

Fudenberg, D., G. Lanzani, and P. Strack (2024). Selective-Memory Equilibrium. *Journal of Political Economy 132*(12), 3978–4020.

Gennaioli, N. and A. Shleifer (2010). What Comes to Mind. *Quarterly Journal of Economics 125*(4), 1399–1433.

Gilboa, I. and D. Schmeidler (1995). Case-Based Decision Theory. *Quarterly Journal of Economics 110*(3), 605–639.

Gilboa, I. and D. Schmeidler (2001). *A Theory of Case-Based Decisions.* Cambridge, UK: Cambridge University Press.

Graeber, T., C. Roth, and F. Zimmermann (2024). Stories, Statistics, and Memory. *Quarterly Journal of Economics 139*(4), 2181–2225.

Hartzmark, S., S. Hirshman, and A. Imas (2021). Ownership, Learning, and Beliefs. *Quarterly Journal of Economics 136*(3), 1665–1717.

Jiang, Z., H. Liu, C. Peng, and H. Yan (2025). Investor Memory and Biased Beliefs: Evidence from the Field. *Quarterly Journal of Economics 140*(4), 2749–2804.

Johnson, E., C. Camerer, S. Sen, and T. Rymon (2002). Detecting Failures of Backward Induction: Monitoring Information Search in Sequential Bargaining. *Journal of Economic Theory 104*(1), 16–47.

Kahana, M. (2012). *Foundations of Human Memory.* New York, NY: Oxford University Press.

Khaw, M., Z. Li, and M. Woodford (2021). Cognitive Imprecision and Small-Stakes Risk Aversion. *Review of Economic Studies 88*, 1979–2013.

Kőszegi, B., G. Loewenstein, and T. Murooka (2022). Fragile Self-Esteem. *Review of Economic Studies 89*(4), 2026–2060.

Malmendier, U. and S. Nagel (2011). Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? *Quarterly Journal of Economics 126*(1), 373–416.

Malmendier, U. and S. Nagel (2016). Learning from Inflation Experiences. *Quarterly Journal of Economics 131*(1), 53–87.

Malmendier, U. and J. Wachter (2024). Memory of Past Experiences and Economic Decisions. In M. Kahana and A. Wagner (Eds.), *Oxford Handbook of Human Memory, Volume II: Applications*. New York, NY: Oxford University Press.

Manzini, P. and M. Mariotti (2014). Stochastic Choice and Consideration Sets. *Econometrica 82*(3), 1153–1176.

McCall, J. (1970). Economics of Information and Job Search. *Quarterly Journal of Economics 84*(1), 113–126.

Mullainathan, S. (2002). A Memory-Based Model of Bounded Rationality. *Quarterly Journal of Economics 117*(3), 735–774.

Nagel, S. and Z. Xu (2022). Asset Pricing with Fading Memory. *Review of Financial Studies 35*(5), 2190–2245.

Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *Economic Journal 117*(523), 1243–1259.

Rubinstein, A. (2016). A Typology of Players: Between Instinctive and Contemplative. *Quarterly Journal of Economics 131*(2), 859–890.

Rubinstein, A. and Y. Salant (2006). A Model of Choice from Lists. *Theoretical Economics 1*(1), 3–17.

Salant, Y. (2011). Procedural Analysis of Choice Rules with Applications to Bounded Rationality. *American Economic Review 101*(2), 724–748.

Salant, Y. and J. Spenkuch (2026). Complexity and Satisficing: Theory with Evidence from Chess. *Review of Economic Studies 93*(2), 1296–1322.

Salant, Y., J. Spenkuch, and D. Almog (2025). The Memory Premium. NBER Working Paper No. 33649.

Scharnagl, R. (2004). *Fischer-Random-Schach (FRC / Chess960): Die revolutionäre Zukunft des Schachspiels*. Hamburg: Books on Demand GmbH.

Simon, H. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics 69*(1), 99–118.

Smith, S. M., A. Glenberg, and R. A. Bjork (1978). Environmental Context and Human Memory. *Memory & Cognition 6*(4), 342–353.

Thomson, D. and E. Tulving (1970). Associative Encoding and Retrieval: Weak and Strong Cues. *Journal of Experimental Psychology 86*(2), 255–262.

Tulving, E. and D. Thomson (1973). Encoding Specificity and Retrieval Processes in Episodic Memory. *Psychological Review 80*(5), 352–373.

Wachter, J. and M. Kahana (2024). A Retrieved-Context Theory of Financial Decisions. *Quarterly Journal of Economics 139*(2), 1095–1147.

Weitzman, M. (1979). Optimal Search for the Best Alternative. *Econometrica 47*(3), 641–654.

Wilson, A. (2014). Bounded Memory and Biases in Information Processing. *Econometrica 82*(6), 2257–2294.

Woodford, M. (2014). Stochastic Choice: An Optimizing Neuroeconomic Model. *American Economic Review 104*(5), 495–500.

Woodford, M. (2020). Modeling Imprecision in Perception, Valuation, and Choice. *Annual Review of Economics* (12), 579–601.

Online Appendix

Contents

List of Figures

List of Tables

Appendix A: Proofs and Another Speedup Prediction

After stating a few lemmas, this appendix proves Predictions 1 and 2. It then states and proves an additional speedup prediction that is discussed in Section 2 and tested in Section 7. Throughout, we focus on the case in which the chosen alternatives are satisfactory.

A.1. *Setup*

Fix the choice situation. For comparisons across memories, fix a satisfactory alternative $a \in A \cap M$ and write $M_1 := M$ and $M_0 := M - a$ for the two memories being compared. Let μ_i denote the retrieval function under M_i, so that $\mu_0(a) = 0$. For a menu $B \subseteq A$ containing a, with $m := |B|$, define

$$L_i(B) := 1 - \sum_{y \in B} \mu_i(y), \qquad p_i(x \mid B) := \mu_i(x) + \frac{L_i(B)}{m},$$

so that $p_i(\cdot \mid B)$ is the sampling distribution on B.

For every $x \in A - a$, $\mu_1(x) \leq \mu_0(x)$. If $x \in M_0$, this follows from individual monotonicity. Otherwise, both retrieval probabilities are zero. For every $C \subseteq A - a$, define the cumulative retrieval loss $\Lambda(C) := \sum_{x \in C}(\mu_0(x) - \mu_1(x))$. Because $\mu_1(x) \leq \mu_0(x)$, $\Lambda(C) \geq 0$. Moreover, aggregate monotonicity gives $\Lambda(A - a) \leq \mu_1(a)$. Because every summand is nonnegative, $\Lambda(C) \leq \Lambda(A - a)$ for every $C \subseteq A - a$. Hence,

$$(5) \qquad 0 \leq \Lambda(C) \leq \mu_1(a) \qquad \text{for every } C \subseteq A - a.$$

Consequently, every menu B containing a satisfies $L_0(B) - L_1(B) = \mu_1(a) - \Lambda(B - a)$. Equation (5) therefore implies

$$(6) \qquad 0 \leq L_0(B) - L_1(B) \leq \mu_1(a).$$

Partition the choice set into the satisfactory alternatives $\mathcal{G}$ and the unsatisfactory alternatives $\mathcal{U}$, and let $s := |\mathcal{G}| \geq 1$. The search stops at the first satisfactory draw, so before stopping it eliminates only unsatisfactory alternatives. Every menu it reaches therefore has the form

$$B = \mathcal{G} \cup V, \qquad V \subseteq \mathcal{U}.$$

We call such a menu *relevant*. Throughout, $V = B \cap \mathcal{U}$ denotes the unsatisfactory alternatives remaining in B.

Let $R_i(B)$ denote the expected number of evaluations from B until the first satisfactory draw, and let $q_i^z(B)$ denote the probability that $z \in \mathcal{G}$ is that draw. We abbreviate $q_i(B) := q_i^a(B)$. Since one evaluation is always spent and an unsatisfactory draw restarts the search from a smaller menu,

$$R_i(B) = 1 + \sum_{c \in V} p_i(c \mid B) R_i(B - c), \qquad R_i(\mathcal{G}) = 1. \tag{7}$$

A.2. *Lemmas*

The first two lemmas describe how sampling probabilities respond to the two changes used throughout the proofs. Lemma 1 shows that adding a to memory shifts sampling probability toward a. Lemma 2 shows that deleting an alternative c from a menu redistributes its sampling probability equally among the alternatives that remain.

LEMMA 1 (Shift toward a): *For every $x \in B - a$, $p_1(x \mid B) \leq p_0(x \mid B)$. If $m \geq 2$, then $p_1(a \mid B) > p_0(a \mid B)$, and hence $p_1(x \mid B) < p_0(x \mid B)$ for at least one $x \in B - a$.*

PROOF: For every $x \in B - a$, the pointwise retrieval comparison above gives $\mu_1(x) \leq \mu_0(x)$, while (6) gives $L_1(B) \leq L_0(B)$. Thus, both components of $p_i(x \mid B) = \mu_i(x) + L_i(B)/m$ weakly fall when a enters memory, proving the first claim.

For a, we have $\mu_1(a) > \mu_0(a) = 0$. Thus,

$$p_1(a \mid B) - p_0(a \mid B) = \mu_1(a) - \frac{L_0(B) - L_1(B)}{m} \geq \mu_1(a)\left(1 - \frac{1}{m}\right) > 0$$

where the weak inequality is a consequence of the upper bound in (6) and the strict one is because $m \geq 2$. Since the two sampling distributions each sum to one, this strict gain must be matched by a strict loss for at least one $x \in B - a$. *Q.E.D.*

LEMMA 2 (Equal redistribution): *If $m \geq 2$, then for every $c \in B$ and every $x \in B - c$,*

$$p_i(x \mid B - c) = p_i(x \mid B) + \frac{p_i(c \mid B)}{m - 1}.$$

PROOF: Deleting c leaves the remaining $m - 1$ alternatives to absorb $p_i(c \mid B)$, since their sampling probabilities must sum to one. Their retrieval probabilities are unaffected by the deletion, so the adjustment occurs entirely through the leftover component. Because this component is distributed uniformly, all $m - 1$ remaining alternatives gain the same amount. Each therefore gains $p_i(c \mid B)/(m - 1)$. *Q.E.D.*

The next lemma concerns expected search length. It shows that deleting an unsatisfactory alternative from a menu weakly shortens the search.

LEMMA 3 (Deletion shortens the search): *For every relevant menu B and every $c \in V$, $R_i(B) \geq R_i(B - c)$, with strict inequality if $p_i(c \mid B) > 0$.*

Proof: We proceed by induction on $|V|$. If $|V| = 1$, then $V = \{c\}$ and $R_i(B) - R_i(B-c) = p_i(c \mid B)$, which proves the claim.

Now suppose $|V| \geq 2$. Call the search from B search I and the search from $B - c$ search II. Couple their first draws as follows. For each $x \in B - c$, let both searches draw x with probability $p_i(x \mid B)$. With the remaining probability $p_i(c \mid B)$, let search I draw c and let search II draw uniformly from its $m - 1$ alternatives. By Lemma 2, this construction gives each search the correct first-draw distribution.

If both searches draw the same satisfactory alternative, both stop. If they draw the same unsatisfactory alternative x, they continue from $B - x$ and $B - \{c, x\}$, respectively. By the induction hypothesis, search I's continuation is weakly longer.

It remains to consider the event on which search I draws c. Because c is unsatisfactory, search I continues from $B - c$. Search II draws uniformly from $B - c$. If it draws a satisfactory alternative, it stops; if it draws an unsatisfactory alternative x, it continues from $B - \{c, x\}$, which is weakly shorter than continuing from $B - c$ by the induction hypothesis. Thus, search I is weakly longer. If $p_i(c \mid B) > 0$, then with positive probability search I draws c while search II draws a satisfactory alternative, so the inequality is strict. *Q.E.D.*

The next lemma links choice probabilities to expected search length. It shows that the choice probability of any satisfactory alternative is affine in expected search length, with slope equal to its retrieval edge over the average satisfactory alternative.

Lemma 4 (Choice probability is affine in search length): *For $z \in \mathcal{G}$, let $e_i^z := \mu_i(z) - \frac{1}{s}\sum_{w \in \mathcal{G}} \mu_i(w)$ denote its retrieval edge over the average satisfactory alternative. Then, for every relevant menu B, $q_i^z(B) = \frac{1}{s} + e_i^z R_i(B)$.*

Proof: Fix $z, w \in \mathcal{G}$. At every step of the search for a good enough alternative, the remaining menu C is relevant and therefore contains both z and w. Because their leftover components coincide,

$$p_i(z \mid C) - p_i(w \mid C) = \mu_i(z) - \mu_i(w).$$

This difference in sampling probabilities contributes to the corresponding difference in choice probabilities only if that step of the search is reached. Summing over steps therefore amounts to multiplying $\mu_i(z) - \mu_i(w)$ by $R_i(B)$, because the expected number of evaluations equals the sum of the probabilities that each successive step occurs. Hence

$$q_i^z(B) - q_i^w(B) = \big[\mu_i(z) - \mu_i(w)\big] R_i(B).$$

Summing this identity over $w \in \mathcal{G}$ and using $\sum_{w \in \mathcal{G}} q_i^w(B) = 1$ gives

$$sq_i^z(B) - 1 = \left[s\mu_i(z) - \sum_{w \in \mathcal{G}} \mu_i(w)\right] R_i(B) = se_i^z R_i(B).$$

Dividing by s and rearranging gives the result. *Q.E.D.*

The final lemma in this subsection provides the coupling used to prove the two deliberation-time predictions. It shows that if one conditioned search is no more likely to continue, then the ordered list of alternatives it rejects can be coupled to be a prefix of the corresponding list for the other search.

LEMMA 5 (Coupling): *Consider two conditioned searches indexed by* $j \in \{0, 1\}$*. From a relevant menu* $B = \mathcal{G} \cup V$*, search* j *continues through* $c \in V$ *to* $B - c$ *with probability* $h_j(c \mid B)$ *and stops at its terminal satisfactory alternative with the remaining probability. If*

$$h_1(c \mid B) \leq h_0(c \mid B)$$

for every relevant menu B *and every* $c \in V$*, then the searches can be coupled so that the ordered list of unsatisfactory alternatives rejected by search 1 is a prefix of the corresponding list for search 0. If the inequality is strict for some* c *at a menu reached jointly with positive probability, then the prefix is proper with positive probability.*

PROOF: At any relevant menu B reached by both searches, couple their next transitions as follows. For each $c \in V$, set

$$\Pr(\text{both continue from } B - c) = h_1(c \mid B)$$

and

$$\Pr(\text{search 1 stops and search 0 continues from } B - c) = h_0(c \mid B) - h_1(c \mid B).$$

With the remaining probability $1 - \sum_{c \in V} h_0(c \mid B)$, let both searches stop at their respective terminal alternatives. These probabilities are nonnegative by hypothesis and sum to one.

Under this construction, search 1 continues through c with probability $h_1(c \mid B)$, while search 0 does so with probability $h_1(c \mid B) + [h_0(c \mid B) - h_1(c \mid B)] = h_0(c \mid B)$. Thus both searches have the correct marginal continuation laws.

Whenever both searches continue, they reject the same alternative c and arrive at the same menu $B - c$, where the construction is repeated. If search 1 stops first, search 0 continues under its own law. Hence the ordered list of alternatives rejected by search 1 is a prefix of the list rejected by search 0.

If $h_1(c \mid B) < h_0(c \mid B)$ at a menu reached jointly with positive probability, then search 1 stops while search 0 continues through c with positive probability. The prefix is therefore proper with positive probability. *Q.E.D.*

A.3. *Proof of Prediction 1*

We prove the stronger claim that $q_1(B) \geq q_0(B)$ for every relevant menu B, with strict inequality when $s \geq 2$. This menu-level comparison will also be used in the proof of Prediction 3. The prediction follows because A is itself relevant.

Fix a relevant menu B. Any path on which a satisfactory alternative $z \neq a$ is chosen consists of a sequence of distinct unsatisfactory alternatives followed by z. Along such a path, a remains available and every alternative drawn differs from a. Lemma 1 therefore implies that the sampling probability of each draw, conditional on the preceding draws, weakly falls from M_0 to M_1. The probability of every such path thus weakly falls. Summing over all such paths gives $1 - q_1(B) \leq 1 - q_0(B)$, and hence $q_1(B) \geq q_0(B)$.

Now suppose $s \geq 2$. If $p_1(y \mid B) < p_0(y \mid B)$ for some satisfactory $y \neq a$, the one-step path (y) has strictly lower probability under M_1 than under M_0. Otherwise, $p_1(y \mid B) = p_0(y \mid B)$ for every satisfactory $y \neq a$. The strict part of Lemma 1 then gives an unsatisfactory $v \in B - a$ such that $p_1(v \mid B) < p_0(v \mid B)$. Fix any satisfactory $y \neq a$. By Lemma 2,

$$p_i(y \mid B - v) = p_i(y \mid B) + \frac{p_i(v \mid B)}{m - 1}.$$

Hence $p_1(y \mid B - v) < p_0(y \mid B - v)$. The two-step path (v, y) therefore has positive probability under M_0 and strictly lower probability under M_1. Thus, in either case, at least one path ending at a satisfactory alternative other than a falls strictly in probability. Consequently, $1 - q_1(B) < 1 - q_0(B)$ and $q_1(B) > q_0(B)$. *Q.E.D.*

A.4. *Proof of Prediction 2*

Hold memory M fixed and suppress the memory index. Let $x \in \mathcal{G} \cap M$ and $y \in \mathcal{G} - M$ be chosen with positive probability, so $\mu(x) > 0 = \mu(y)$. For $z \in \{x, y\}$, define

$$\widehat{p}^z(c \mid B) := \frac{p(c \mid B) q^z(B - c)}{q^z(B)} \qquad \text{for every } c \in V,$$

and call the corresponding conditioned search the z-search. Write e^z for the retrieval edge defined in Lemma 4, with the memory index suppressed. Then

$$q^z(B) = \frac{1}{s} + e^z R(B) \qquad \text{and} \qquad e^x - e^y = \mu(x) - \mu(y) > 0.$$

The conditioned searches are well defined at every relevant menu. Indeed, $q^x(B) \geq p(x \mid B) \geq \mu(x) > 0$. Moreover, $e^y \leq 0$ because $\mu(y) = 0$, and repeated application of Lemma 3 gives $R(B) \leq R(A)$, so Lemma 4 implies $q^y(B) \geq q^y(A) > 0$.

To obtain the result, we wish to apply Lemma 5. We therefore need to show that the x-search is no more likely than the y-search to continue through any unsatisfactory alternative. Fix a relevant menu B and $c \in V$. Using the definition of $\widehat{p}^z$ and Lemma 4,

$$\begin{aligned} \widehat{p}^x(c \mid B) - \widehat{p}^y(c \mid B) &= p(c \mid B)\left[\frac{q^x(B-c)}{q^x(B)} - \frac{q^y(B-c)}{q^y(B)}\right] \\ &= \frac{p(c \mid B)(e^x - e^y)}{s\, q^x(B) q^y(B)}\big[R(B-c) - R(B)\big] \le 0. \end{aligned} \tag{8}$$

Thus the conditional continuation probabilities satisfy the ordering required by Lemma 5.

Apply Lemma 5, with the x-search in the role of search 1 and the y-search in the role of search 0. Let H be the set of unsatisfactory alternatives rejected by the x-search, and let J be the additional alternatives rejected by the y-search. Their evaluated sets are $H \cup \{x\}$ and $H \cup J \cup \{y\}$, respectively.

Because $x \in M$ and $y \notin M$, replacing y with x while holding H fixed cannot increase the variable component of deliberation time. Moreover, expanding $H \cup \{y\}$ by adding J cannot decrease it. Hence

$$t_M(H \cup \{x\}) \le t_M(H \cup \{y\}) \le t_M(H \cup J \cup \{y\}).$$

Because the fixed component is common, we obtain that deliberation time conditional on choosing x is weakly shorter on every outcome of the coupling, implying that it is first-order stochastically smaller than deliberation time conditional on choosing y.

For strictness, suppose $\mathcal{U} \neq \varnothing$. Because $q^y(A) > 0$ and $\mu(y) = 0$, there exists $c \in \mathcal{U}$ such that $p(c \mid A) > 0$. Indeed, if $p(y \mid A) > 0$, then the leftover component is positive, so every unsatisfactory alternative has positive sampling probability. If $p(y \mid A) = 0$, any positive-probability path ending at y must begin with an unsatisfactory alternative c satisfying $p(c \mid A) > 0$. For this c, Lemma 3 gives $R(A - c) < R(A)$. Hence the inequality in (8) is strict at $B = A$. Lemma 5 then gives $J \neq \varnothing$ with positive probability. On those outcomes, $H \cup \{y\} \subsetneq H \cup J \cup \{y\}$, so expanding the evaluated set strictly increases the variable component of deliberation time. This proves the strict speedup. *Q.E.D.*

A.5. *An Additional Speedup Prediction*

Unlike Prediction 2, which holds memory fixed and compares different chosen alternatives, the next prediction holds the chosen alternative fixed and compares two memories that differ only in whether that alternative is in memory or not.

PREDICTION 3 (Speedup Across Memories): *Satisficing with memory-dependent consideration gives rise to a* speedup across memories. *That is, for any memory M and any alternative $a \in A \cap M$ that is chosen with positive probability under both M and $M - a$, deliberation time conditional on choosing a is weakly shorter under M than under $M - a$. It is strictly shorter if a is satisfactory and the choice set contains at least one unsatisfactory alternative.*[1]

For this prediction, we can dispense with the assumption that out-of-memory alternatives take weakly longer to evaluate. Instead, holding the set of evaluated alternatives fixed, we require that

[1]If all alternatives are satisfactory, the DM stops after the first evaluation under either memory.

adding an alternative to memory does not increase the variable component of deliberation time

$$t_M(\mathcal{E}) \leq t_{M-a}(\mathcal{E}) \qquad \text{for every } a \in M \text{ and every } \mathcal{E} \subseteq A.$$

To prove the prediction, we first establish a lemma showing that adding a to memory weakly reduces the marginal delay caused by any unsatisfactory alternative.

LEMMA 6 (Memory decreases marginal delay): *For every relevant menu B and every $c \in V$, let $\Delta_i(B,c) := R_i(B) - R_i(B-c)$. Then, $0 \leq \Delta_1(B,c) \leq \Delta_0(B,c)$.*

PROOF: The lower bound follows from Lemma 3. For the upper bound, consider any relevant menu C. The unconditioned search under M_i continues through an unsatisfactory alternative $x \in C$ with probability $p_i(x \mid C)$. Lemma 1 gives $p_1(x \mid C) \leq p_0(x \mid C)$. Treating all satisfactory draws as a single stopping outcome, the construction in the proof of Lemma 5, applied with $h_i(x \mid C) = p_i(x \mid C)$, couples the searches so that the alternatives rejected under M_1 form a prefix of those rejected under M_0. Because each search makes one terminal satisfactory evaluation after its rejected alternatives, the prefix relation implies $R_1(C) \leq R_0(C)$. Thus, $D(C) := R_0(C) - R_1(C) \geq 0$ for every relevant menu C.

We prove $\Delta_1(B,c) \leq \Delta_0(B,c)$ by induction on $n := |V|$.

For $n = 1$, $B - c = \mathcal{G}$. Hence, $\Delta_i(B,c) = p_i(c \mid B)$, and the result follows from Lemma 1.

Now let $n \geq 2$ and suppose the result holds for every relevant menu containing fewer than n unsatisfactory alternatives. Before making the main induction step, we define

$$H_i(B,c) := R_i(B-c) - \frac{1}{m-1} \sum_{x \in V-c} R_i(B - \{c,x\})$$

for every $c \in V$.

The construct H_i has two useful properties. First, it is nonnegative. To see this, observe that by Lemma 3, every term in the sum defining $H_i(B,c)$ is at most $R_i(B-c)$. Because the sum contains $n-1$ terms and $m-1 = s+n-1$, we obtain that

$$H_i(B,c) \geq \frac{s}{m-1} R_i(B-c) \geq 0.$$

Second, the difference $H_0(B,c) - H_1(B,c)$ is also nonnegative. To see this, we observe that for every $x \in V - c$, the induction hypothesis applied at $B - c$ gives

$$D(B-c) - D(B - \{c,x\}) = \Delta_0(B-c,x) - \Delta_1(B-c,x) \geq 0.$$

Together with $D(B-c) \geq 0$ proved above, this implies

$$H_0(B,c) - H_1(B,c) = D(B-c) - \frac{1}{m-1} \sum_{x \in V-c} D(B - \{c,x\})$$
$$\geq \frac{s}{m-1} D(B-c) \geq 0.$$

Thus $0 \leq H_1(B,c) \leq H_0(B,c)$.

We are now ready for the main induction step. Subtracting (7) at $B-c$ from (7) at B and using Lemma 2 gives

$$\begin{aligned}\Delta_i(B,c) &= p_i(c \mid B) \left[R_i(B-c) - \frac{1}{m-1} \sum_{x \in V-c} R_i(B - \{c,x\}) \right] \\ &\quad + \sum_{x \in V-c} p_i(x \mid B) \left[R_i(B-x) - R_i(B - \{c,x\}) \right] \\ &= p_i(c \mid B) H_i(B,c) + \sum_{x \in V-c} p_i(x \mid B) \Delta_i(B-x,c). \qquad (9)\end{aligned}$$

Every term on the right-hand side of (9) is weakly smaller for $i = 1$ than for $i = 0$. For the first term, this follows from $0 \leq H_1(B,c) \leq H_0(B,c)$ and $p_1(c \mid B) \leq p_0(c \mid B)$, where the latter inequality follows from Lemma 1. For every $x \in V - c$, Lemma 1 also gives $p_1(x \mid B) \leq p_0(x \mid B)$, while Lemma 3 and the induction hypothesis give

$$0 \leq \Delta_1(B-x,c) \leq \Delta_0(B-x,c).$$

Hence $\Delta_1(B,c) \leq \Delta_0(B,c)$, completing the induction. *Q.E.D.*

Proof of Prediction 3. Fix a satisfactory alternative a with $q_0(A) > 0$. Observe the analogy to the proof of Prediction 2. Under M_1, a plays the role of the in-memory alternative x in that proof; under M_0, a plays the role of the out-of-memory alternative y. We can therefore use the conditioned-search construction and well-definedness argument from that proof. For $i \in \{0,1\}$, call the search under M_i, conditional on choosing a, search i, and write

$$\widehat{p}_i(c \mid B) = \frac{p_i(c \mid B) q_i(B-c)}{q_i(B)} \qquad (10)$$

for its probability of continuing through $c \in V$ from a relevant menu B.

The key departure from Prediction 2 is that the sampling process is held fixed there, whereas changing memory here changes both p_i and q_i. Lemma 1 gives $p_1(c \mid B) \leq p_0(c \mid B)$. Hence, to establish the weak ordering of the conditional continuation probabilities, it suffices to establish the ranking of the q_i-ratios. We then show directly that $\widehat{p}_1(c \mid A) < \widehat{p}_0(c \mid A)$ for some $c \in \mathcal{U}$ when $\mathcal{U} \neq \varnothing$.

Weak ordering of the q_i-ratios. Fix a relevant menu B and $c \in V$. Write $\Delta_i := \Delta_i(B, c)$ and $e_i := e_i^a$. Since $R_i(B - c) = R_i(B) - \Delta_i$, Lemma 4 gives

$$\frac{q_1(B-c)}{q_1(B)} \le \frac{q_0(B-c)}{q_0(B)} \iff 1 - \frac{e_1\Delta_1}{q_1(B)} \le 1 - \frac{e_0\Delta_0}{q_0(B)} \iff \frac{e_1\Delta_1}{q_1(B)} \ge \frac{e_0\Delta_0}{q_0(B)}. \tag{11}$$

To establish the last inequality in (11), and, hence, the desired ratio ordering, observe that Lemma 6 and the menu-level comparison established in the proof of Prediction 1 give

$$0 \le \frac{\Delta_1}{q_1(B)} \le \frac{\Delta_0}{q_0(B)}.$$

Because $e_i = \mu_i(a) - \frac{1}{s}\sum_{w\in\mathcal{G}} \mu_i(w)$, we have $e_0 \le 0$ as $\mu_0(a) = 0$. Moreover,

$$\begin{aligned} s(e_1 - e_0) &= s\big[\mu_1(a) - \mu_0(a)\big] - \sum_{w\in\mathcal{G}}\big[\mu_1(w) - \mu_0(w)\big] \\ &= (s-1)\mu_1(a) + \sum_{w\in\mathcal{G}-a}\big[\mu_0(w) - \mu_1(w)\big] \\ &= (s-1)\mu_1(a) + \Lambda(\mathcal{G} - a). \end{aligned}$$

Because $s \ge 1$ and $\mu_1(a) > 0$, the first term is nonnegative and is strictly positive when $s \ge 2$. The second term is nonnegative by (5). Thus $e_1 \ge e_0$, with strict inequality when $s \ge 2$.

We therefore obtain

$$\begin{aligned} \frac{e_1\Delta_1}{q_1(B)} &\ge \frac{e_0\Delta_1}{q_1(B)} \quad \text{because } e_1 \ge e_0 \text{ and } \frac{\Delta_1}{q_1(B)} \ge 0, \\ &\ge \frac{e_0\Delta_0}{q_0(B)} \quad \text{because } e_0 \le 0 \text{ and } \frac{\Delta_1}{q_1(B)} \le \frac{\Delta_0}{q_0(B)}. \end{aligned}$$

Thus, the final inequality in (11), and hence the desired ratio ordering, follows.

Strict ordering. Suppose $\mathcal{U} \ne \varnothing$. If $s = 1$, then a is the only satisfactory alternative, so $q_i(B) = 1$ for every relevant menu B. Hence (10) gives

$$\widehat{p}_i(c \mid A) = p_i(c \mid A).$$

The strict part of Lemma 1 provides some $c \in A - a$ such that $p_1(c \mid A) < p_0(c \mid A)$. Since a is the only satisfactory alternative, $c \in \mathcal{U}$, establishing the strict ordering.

Now suppose $s \ge 2$. An argument analogous to the one made in the proof of Prediction 2, can be used to show that some $c \in \mathcal{U}$ must satisfy $p_0(c \mid A) > 0$. Fix such a c. Because the conditioned searches are well defined at every relevant menu, $q_0(A - c) > 0$, and hence $\widehat{p}_0(c \mid A) > 0$.

If $p_1(c \mid A) = 0$, then (10) immediately gives $\widehat{p}_1(c \mid A) = 0 < \widehat{p}_0(c \mid A)$. Suppose instead that $p_1(c \mid A) > 0$. Lemma 3 gives $\Delta_1(A, c) > 0$. Because $e_1 > e_0$ when $s \ge 2$, the first inequality in the comparison used to establish the weak ordering is strict at $B = A$. Hence, the final inequality in (11) is also strict at $B = A$, so

$$\frac{q_1(A - c)}{q_1(A)} < \frac{q_0(A - c)}{q_0(A)}.$$

Together with $0 < p_1(c \mid A) \leq p_0(c \mid A)$, and (10), this implies $\widehat{p}_1(c \mid A) < \widehat{p}_0(c \mid A)$.

Thus, we have shown that $\widehat{p}_1(c \mid B) \leq \widehat{p}_0(c \mid B)$ at every relevant menu, with strict inequality at $B = A$ for some $c \in \mathcal{U}$ when $\mathcal{U} \neq \varnothing$.

To complete the proof, apply Lemma 5, and let $\mathcal{E}_i$ denote the evaluated set under search i. Because both searches terminate at a, $\mathcal{E}_1 \subseteq \mathcal{E}_0$ for every realization of the coupled searches. Because A is the initial menu, the strict continuation inequality makes this inclusion strict with positive probability. Because M_1 differs from M_0 only by adding a to memory, holding $\mathcal{E}_1$ fixed cannot increase the variable component of deliberation time. Moreover, expanding $\mathcal{E}_1$ to $\mathcal{E}_0$ cannot decrease it. Hence

$$t_{M_1}(\mathcal{E}_1) \leq t_{M_0}(\mathcal{E}_1) \leq t_{M_0}(\mathcal{E}_0).$$

The second inequality is strict whenever $\mathcal{E}_1 \subsetneq \mathcal{E}_0$. Because the fixed component is common, we obtain the result. *Q.E.D.*

Appendix B: Data Appendix

The chess data in this paper were furnished by `lichess.org`. Every day, Lichess hosts several million games of standard chess and thousands of games of Chess960. In the summer of 2021, we contacted the founder of Lichess and purchased a database extract covering *all* standard and variant games that were played on the platform between January 2013 and June 2021.[2] The raw data were provided to us in the human-readable PGN format and include basic facts about each game (including players' usernames and ratings, date and time of the game, time controls, ultimate outcome, etc.), as well as the exact sequence of moves. The total size of the raw data is about 6.9TB. In order to reconstruct users' history of play, we process these data using automated scripts and extract all opening moves of every registered player.

The sample for our main analysis restricts attention to the opening moves of 147,000 players in their first game of Chess960 on the platform. In order to be included in this sample, users must have executed at least 20 opening moves in standard chess games on Lichess prior to their first game of Chess960. This restriction helps to ensure a minimal level of accuracy in approximating players' memory. We further require that their first game of Chess960 was rated, and that they played White. We focus on moves as White because we want to analyze their initial choices in an unfamiliar decision problem, and we require that the game be rated in order to ensure that these are real-stakes decisions.

[2] Lichess regularly releases database extracts for rated chess games at `https://database.lichess.org`. These publicly available extracts do not, however, include casual games.

In other parts of our analysis, we study the moves of all registered users who played a total of at least fifty games of Chess960 as White, regardless of which color they were assigned in their first game and whether that game was rated. There are 16,731 such players in our data.

In order to measure the quality of moves, we use the Stockfish chess engine to rate every feasible move in every possible starting position. All ratings are calculated using a single-threaded instantiation of Stockfish version 16.1, with depth set to 40, 4,096MB hash, the default neural net, and access to the 3-,4-,5-, and 6-man Syzygy tablebases.

Appendix C: Additional Results and Robustness Checks

This appendix reports additional results and robustness checks. The order in which these are discussed corresponds to the order in which they are referenced in the main text, followed by exhibits that are discussed only in the appendix.

Appendix Figure AF.1 plots the frequency of every possible starting position in Chess960 in our main sample of players' first game of Chess960 as White. The reported χ^2-test indicates that the observed frequencies are consistent with the null hypothesis of uniform random assignment.

Appendix Figure AF.2 shows the frequency with which the players in the main sample chose each opening move in standard chess games on Lichess prior to their first game of Chess960.

Appendix Table AT.1 replicates Table 2 in the text controlling for *which* other moves players hold in memory. Specifically, the regression models in this table replace the move-by-starting-position-by-size-of-memory fixed effect in Table 2 with a fixed effect for the combination of the move, starting position, and the set of other in-memory moves.

Appendix Figure AF.3 depicts estimates of the memory premium under alternative definitions of memory. Each alternative definition corresponds to a pair (n, t), for which a given move is said to be in-memory if it has been played at least t times (x-axis) in the last n games of standard chess (y-axis).

Appendix Tables AT.2 and AT.3 provide estimates of the memory in the "flipped" standard board position (i.e., when only the king and queen sit on opposite squares) and in board positions in which the king and queen are placed near the corners of the board (i.e., on squares a1, b2, g1, or h1). The evidence in these tables suggests that players retrieve individual opening moves rather than strategies from memory.

Appendix Table AT.4 explores heterogeneity in the memory premium by skill level, prior experience playing standard chess, and quality of moves. The evidence therein shows higher memory premia among inexperienced players and for moves that are close to optimal. We do not detect any differences, however, by players' skill level.

Appendix Figure AF.4 plots the estimated memory premium for each of the twenty opening moves in standard chess according to the specification in col. (6), Table 2.

Appendix Table AT.5 reports estimates of the memory premium, distinguishing between all starting positions, as well as positions in which the move is and is not close to optimal (i.e., within 20 centipawns of the optimal move in a given position). All estimates in this table are based on the regression model and sample in col. (6) of Table 2, allowing for heterogeneity in the memory premium across opening moves.

Appendix Figure AF.5 and Appendix Tables AT.6–AT.9 respectively replicate the analyses in Figure 8(b) and Table 2 in the text based on "donut definitions" of memory. According to these definitions, a move is said to be in-memory if, and only if, it had been played at least once sufficiently long prior to a player's first game of Chess960.

Appendix Figure AF.6 replicates the analysis in Figure 9(a) in the text, differentiating between moves that players used more or less successfully in prior standard games on Lichess, as measured by the outcomes of the respective games. Specifically, we distinguish between moves with above- and below-median *player-specific* average payoffs in prior use in standard chess.

Appendix Figure AF.7 replicates the analysis in Figure 12 based on a definition of memory that is based on players' choices in prior games of Chess960. Although we estimate a significantly larger premium for this definition of memory, we observe a qualitatively similar decline as players gain experience.

Appendix Table AT.10 reports estimates of the memory premium for the second and third moves. If anything, these estimates indicate that the memory premium for these moves exceeds that for the first one.

Appendix Table AT.11 reports estimates of the comparative statics of the memory speedup with respect to proxies for repetition, decay, similarity, and interference. The evidence in this table suggests that the memory speedup is moderated by these cognitive forces, which are known to affect the probability of recall.

Appendix Table AT.12 presents descriptive statistics for our experimental data, including participant demographics provided by Prolific.

Appendix Tables AT.13 and AT.14 show that our experimental results are robust to controlling for alternatives' expected value, their display position, the time limit, as well as to including rounds in which participants did not submit timely choices. Per the pre-registration, our analysis in the main text excludes all observations from such rounds.

Appendix Table AT.15 reports results from the untimed robustness experiment described in Appendix F. Though smaller, the coefficients in this table are qualitatively similar to the differences shown in Figure 15 in the main text.

Appendix Table AT.16 reports correlations between deliberation times, the number of evaluated alternatives, and optimization errors in our main experiment.

Appendix Figure AF.9 presents evidence from an RD design exploiting the fact that some moves don't "enter" players' memory until after their first game of Chess960. For additional details on this research design, see Appendix E.

The upper panel of Table AT.17 shows that the discontinuity in Figure AF.9 is robust to different estimators, bandwidths, and samples. The lower panel of Table AT.17 complements our RD estimates with evidence from a difference-in-differences approach. For additional details on these research designs, see Appendix E.

Appendix Figure AF.8 shows the density of observations around the threshold in the regression discontinuity design in Figure AF.9 and Table AT.17 in Appendix E.

Appendix D: Do Players Remember Moves or Strategies?

The results in Section 4 demonstrate that first-time players of Chess960 are significantly more likely to choose opening moves that they have previously used in regular chess. This finding raises a natural question about the content of players' memories. Players may retrieve particular moves, or they may retrieve broader opening strategies and adapt them to the new board. Both possibilities are consistent with memory shaping which opening moves receive priority, but they differ in the level at which prior experience is represented. For example, a player may choose ♙e4 because she recalls that particular move or because she recalls a broader strategy of establishing control over the center of the board. Similarly, she may retrieve ♘f3 as an individual move or instead recall the Zukertort Opening and its aim of developing the king-side knight while preserving strategic flexibility. In this appendix, we use variation across starting positions in Chess960 to distinguish between these possibilities.

We first examine players' choices in starting position 534.[3] In this position, only the king and queen are swapped, making it a mirror image of the starting position in regular chess. If players retrieve and adapt familiar opening strategies, they should often select the mirror counterparts of the moves associated with those strategies. For example, a player who recalls the Zukertort Opening should choose ♘c3 instead of ♘f3.[4] If players instead retrieve particular moves, they should continue

[3]For the official numbering of starting positions in Chess960, see Scharnagl (2004).

[4]The correlation between the Stockfish ratings of moves in the standard starting position and the ratings of their mirror counterparts in position 534 is 0.93. It is not exactly equal to one because Stockfish's internal heuristics do not directly exploit symmetry; and because castling paths in starting position 534 differ by one tempo. By contrast, the correlation between the same moves' ratings in the standard opening position and position 534 is only 0.33.

to favor the moves they previously used in regular chess. The results in Table AT.2 suggest that players continue to favor previously used moves.

The coefficients in odd-numbered columns of this table show that, depending on controls, the average memory premium in starting position 534 equals about 3.1–3.8 p.p. Even-numbered columns run horse races between the memory status of a given move and that of its mirror counterpart (e.g., ♘c3 $\rightleftharpoons$ ♘f3). Taken at face value, the coefficients in these columns imply that players are somewhat more likely to choose a particular move when its mirror counterpart is in memory on account of having been played in regular chess. The relevant coefficients, however, are statistically insignificant and much smaller than the corresponding estimates of the memory premium. Moreover, accounting for the memory status of the mirror move changes the estimated memory premium by less than 0.2 p.p. (6.4%). The evidence thus points to the retrieval of particular moves rather than to the adaptation of familiar opening strategies

Moreover, Appendix Table AT.3 shows that the memory premium is quantitatively robust to restricting attention to starting positions in which the king and queen are placed near the corners of the board (i.e., on squares a1, b1, g1, or h1). These positions substantially alter the strategic role of many familiar opening moves. Nevertheless, players continue to favor the particular moves they previously used in regular chess, especially center-pawn moves (i.e., ♙d3, ♙d4, ♙e3, and ♙e4). Taken together, the evidence in Appendix Tables AT.2 and AT.3 leads us to conclude that players retrieve individual moves rather than entire opening strategies. It thus supports the move-level representation of memory used throughout the analysis in the main text.

Appendix E: Evidence from Causal Research Designs

Does being retrievable from memory *cause* higher choice frequencies? In order to provide evidence in support of a causal relationship between memory and choice behavior, we consider a dynamic definition of memory. According to this definition, a move is said to be in-memory if, and only if, a player had chosen it at least once in a standard chess game on Lichess prior to the *current* game of Chess960. A move can thus enter a player's memory between her first game of Chess960 and subsequent ones.[5] Does the frequency with which the respective move is chosen in Chess960 increase afterwards?

Focusing on players who played at least fifty games of Chess960 as White and drawing on the 41,990 instances in which an opening move enters these players' memories after their first game, Appendix Figure AF.9 plots the average choice frequency of these moves before and after they were first used in a regular chess game.[6] Consistent with the idea that players experiment more with out-of-memory alternatives as they gain experience, choice frequencies start out below average and increase over time. Importantly though, immediately after a move is used for the first time in a

[5] By contrast, our (static) workhorse definition of memory considers only opening moves in standard games prior to the *first* game of Chess960 (see Section 3).

[6] Appendix Figure AF.8 shows the density of observations around the threshold. There is no apparent discontinuity.

standard game, the frequency with which it is chosen in Chess960 jumps by approximately 0.3 p.p., or about 6%.

The upper panel of Table AT.17 shows that the discontinuity in Figure AF.9 is robust to different estimators, bandwidths, and samples. For instance, the bias-corrected RD estimator of Calonico et al. (2024), in conjunction with the corresponding optimal bandwidth, yields point estimates of 0.37 p.p. in the sample of all games and 0.36 p.p. when we restrict attention to rated games only. Relying on the standard linear model and alternative bandwidths produces estimates ranging from 0.22 to 0.29 p.p., all of which are statistically significant.

To be clear, these estimates average over situations in which a given opening move had truly never been used before—either in regular chess or Chess960—and those in which it had not been played in standard games on Lichess but was, in fact, used in games outside of the platform or in an individual's first few games of Chess960. In the latter cases, we should not think of the move as literally entering memory. Rather, it seems more appropriate to interpret the event of a move's first use in a standard game on Lichess as reinforcing prior experiences, which, in turn, improves recall. Either way, the observed discontinuity suggests that retrievability from memory affects choice behavior.

The lower panel of Table AT.17 complements our RD estimates with evidence from a difference-in-differences approach. The former research design identifies the effect of moves "entering memory" by comparing average choice frequencies *just* before and after the event. Some of the relevant variation, therefore, comes from comparisons across different players and moves. The latter design compares changes in the choice frequency of moves that enter players' memories to contemporaneous changes in the use of other moves. That is, the difference-in-differences approach leverages variation *within* "treated" and "untreated" player-move combinations. Reassuringly, both research designs yield qualitatively equivalent answers.

It is noteworthy, though, that our causal estimates of the memory premium are an order of magnitude smaller than those in Table 2 in the main text. This is not surprising. On average, the moves that identify the estimates in Table AT.17 enter players' memories between their 16th and 17th game of Chess960 as White. By this time, the average memory premium has already decreased by about half (see Figure 12). Moreover, being played for the first time in a standard chess game on Lichess constitutes a very weak treatment. For comparison, if we restrict attention to the 16th and 17th game of Chess960 and use the regression model in col. (6) of Table 2 to estimate the memory premium for moves that had been played only once before in regular chess, then we obtain a point estimate of 0.44 p.p. (with a standard error of 0.16). There is, therefore, no contradiction between the causal estimates in Table AT.17 and our estimates of the average memory premium in Table 2.

Appendix F: Experimental Details

As explained in the text, our pre-registered main experiment ran on July 12, 2025. We recruited 501 participants via Prolific, promising a $7.50 base payment and a performance-based reward of up to

$15. The experiment itself took place on a custom website that we built using oTree (Chen et al. 2016). To accommodate up to 50 simultaneous participants, we hosted the site on a "Performance-L" instance on Heroku.

The experiment had five phases: 1. Consent; 2. Instructions; 3. Practice Round; 4. Choice Task (2 stages, 20 rounds); 5. Debrief (2 questions).

The heart of the experiment consisted of twenty rounds of a simple choice task. As illustrated in Figure 14 in the text, the choice task had two stages separated by a 15-second wait. We described the task as choosing one out of six fictitious companies for a hypothetical investment. Each company is identified by a unique color and a corresponding name. No company appears in more than one round. Colors and names were generated by artificial intelligence software that was instructed to create a palette of 126 visually distinct colors with "snappy, memorable" names.

In the first stage of the choice task, which we refer to as the research stage, participants are shown three out of the six companies in their choice set, along with information that allows them to determine the profits of each of these alternatives through simple calculations. Profits correspond to the expected value of the alternative prior to the respective company's growth factor being revealed. Whether or not a particular alternative appears in the research stage is random. Participants further know that profits are round numbers between two and ten, and that any of these values is equally likely.

To proceed to the second stage of the choice task, which we call the choice stage, participants must correctly calculate the profits of all three companies, after which they are made to wait for fifteen seconds. In the choice stage, participants are shown all six companies in their choice set, along with sufficient information to calculate the expected value of any alternative. The ordering of alternatives on the screen is random.

To determine an alternative's actual value, participants must elicit a final piece of information. To do this, they need to click on the respective company's Evaluate button. Clicking this button reveals the company's growth factor. It also enables another button that allows participants to select the respective alternative. Subjects may evaluate as many companies as they wish, in any order. The only constraint we impose is that a company must be evaluated before it can be chosen, and that a selection must be made before the time limit for that round expires.

Subjects had between ten and forty-five seconds to complete the choice stage. The time limit was randomized and distributed uniformly and i.i.d. across rounds. As indicated in the preregistration, the results in the main text do not consider rounds in which participants failed to make a timely choice, which occurred in about 3.9% of tasks. We obtain results that are qualitatively equivalent when we include these observations, coding none of the available alternatives as having been chosen (see Appendix Table AT.13).

As a robustness check, we conducted a follow-up experiment with 201 participants, in which we did not impose a binding time constraint.[7] The follow-up experiment ran on July 26, 2025, and involved

[7]For technical reasons, the choice stage in the follow-up experiment ended after ten minutes, which

no other changes. It was, therefore, not pre-registered. Results from the follow-up experiment are qualitatively similar and are presented in Appendix Table AT.15.

Subjects earned a performance-based bonus equal to the value of the alternative they selected in one randomly determined round. The median participant earned a total of $16.18 (incl. a $7.50 base payment) and spent about 34.5 minutes on the experiment.[8]

Appendix Table AT.12 presents descriptive statistics for our experimental data, including participant demographics provided by Prolific. On the following pages, we reproduce the text that was shown to participants, with horizontal lines demarcating screens.

corresponds to more than ten times the maximal time limit in the main experiment. In our view, subjects who fail to submit a choice within ten minutes are likely not paying attention to the task.

[8]These numbers refer to participants in the main experiment. Subjects in the follow-up experiments faced the same incentives, took about 45 minutes to complete the experiment, and earned $16.57, on average.

Experimental Screens

Research Study

Research Study: Prolific Choice Experiments (STU00224136)
Principal Investigators: Dr. Yuval Salant; Dr. Jörg Spenkuch
Supported By: This research is funded by Northwestern University.

Welcome! The purpose of this study is to better understand how memory influences decision-making. We are very grateful for your help!

To participate in this study you must be at least 18 years old and reside in the U.S. We also ask that you only participate if you are using a computer. Tablets and phones are not supported. Below we provide additional information about this study in order to help you decide whether you'd like to participate.

To get started, you need to provide your consent by pressing the PROCEED button at the bottom of this page.

What should I expect?
Your participation is voluntary. If you choose to participate, you will be asked to perform basic mathematical calculations and to choose between different hypothetical investment options. We estimate that it will take about 45 minutes to complete the study.

Will I be paid?
We will reward completion of this study with a base payment of $7.50, and a performance-based reward up to $15. Your performance-based reward will depend on your ability to identify good investment options.

Are there any risks?
We foresee little risk from participating in this study, and we do not guarantee that you will receive any benefits beyond the payments described above.

How will my information be used?
The information collected in this study will be exclusively used for research purposes. All data will be handled and stored in accordance with Northwestern University policy. There is minimal risk that participants might be identified from the information provided. The research team will take extensive precautions to keep all data secure in order to protect confidentiality. As part of this effort, your actual identity will remain unknown to the researchers conducting this study. The results of this research may be published, but only in anonymized form.

Who can I talk to?
If you have questions, concerns, or complaints, you can contact the Principal Investigators at

prolific@u.northwestern.edu. This research has been reviewed and approved by an Institutional Review Board (IRB) — an IRB is a committee that protects the rights of people who participate in research studies. You may contact the IRB by phone at +1 (312) 503-9338 or by email at irb@northwestern.edu if:

- Your questions, concerns, or complaints are not being answered by the research team.
- You cannot reach the research team.
- You want to talk to someone besides the research team.
- You have questions about your rights as a research participant.
- You want to get information or provide input about this research.

By proceeding to the next screen, you are consenting to participate in this study.

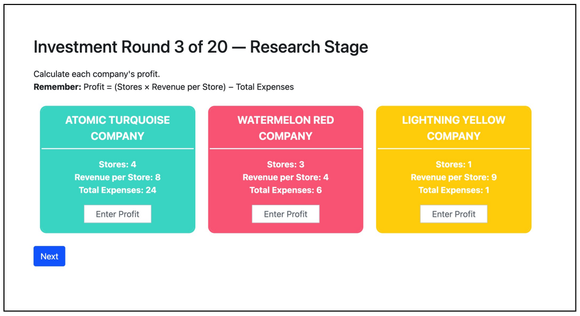
Instructions — Page 1 of 5

In this study, you will be asked to make hypothetical investment decisions. Each decision involves choosing one out of six fictitious companies. Prior to making a choice, you will be given the opportunity to memorize information that partially determines the value of three of the six companies.

There are twenty rounds in this study. In each round, you will make one investment decision. At the end of the study, we will pick one decision at random, and you will be paid the value of the company that you chose in that round. Your performance-based bonus, therefore, depends on the quality of your decisions.

To see an example of what exactly you're being asked to do in each investment round, please proceed to the next screen.

Instructions — Page 2 of 5



Each round in this study consists of three stages: a research stage, a wait phase, and a choice stage.

In the **research stage**, you will be given information on three companies. As in the screenshot above, the information on each company includes its number of stores, the revenue per store, and its total expenses.

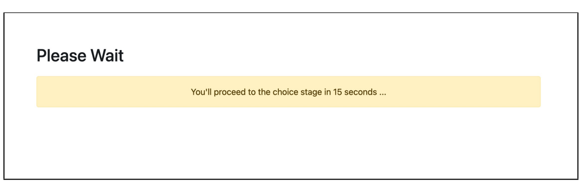
Based on this information, you'll be asked to calculate each company's profit. Profit is determined according to the following formula:

$$\text{Profit} = (\text{Stores} \times \text{Revenue per Store}) - \text{Total Expenses}$$

For example, applying this formula to the Lightning Yellow Company in the screenshot above, its profit equals 8 [$= (1 \times 9)$ - 1]. The profit of the Watermelon Red Company, however, is 6 [$= (3 \times 4)$ - 6].

You will not be able to proceed to the next screen until you have entered the correct profit numbers for all three companies. Profits will play an important role in the choice stage.

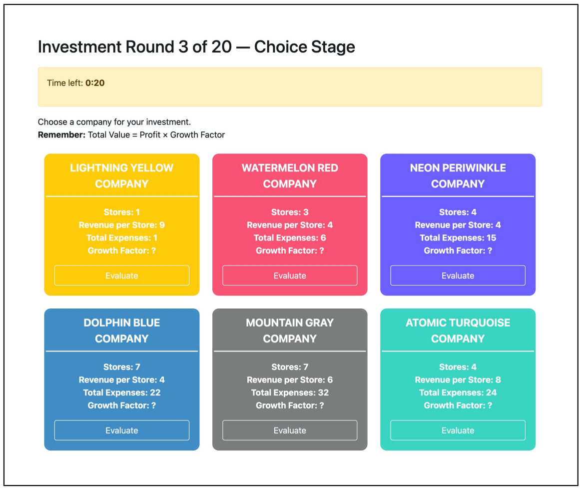
Instructions — Page 3 of 5



Every research stage is followed by a brief **wait phase**. The length of the wait is 15 seconds.

While you wait, you are being asked to remember the profit of each of the three companies that you encountered during the research stage.

Instructions — Page 4 of 5



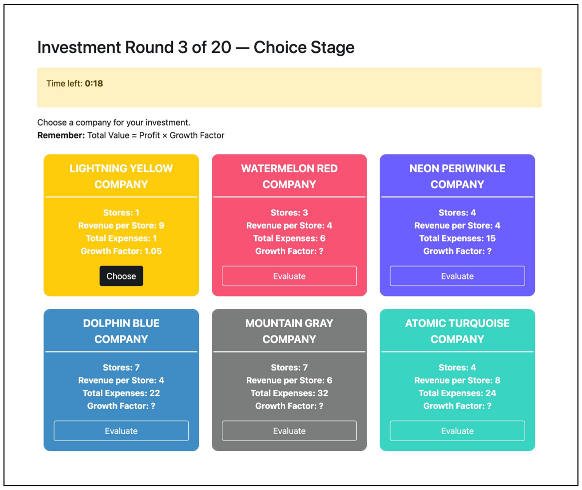
In the **choice stage** of each round, you are being asked to invest in one out of six companies, three of which you have already encountered during the research stage.

The total value of a given company depends on its profitability and growth. Companies that are growing rapidly have a *growth factor* greater than one, whereas companies that are likely to go out of business have a growth factor smaller than one. More specifically, the value of a given company is determined according to the formula:

$$\text{Value} = \text{Profit} \times \text{Growth Factor}$$

While you may remember the profit numbers for the companies that you encountered in the research stage, you will need to click on the Evaluate button in order to discover a company's growth factor.

Instructions — Page 5 of 5



In the example above, the growth factor of the Lightning Yellow Company is 1.05. Since the company's profit is 8 [$= 1 \times 9 - 1$], its total value equals 8.4 [$= 8 \times 1.05$].

You may click on the evaluate buttons of as many companies as you like. After evaluating a particular company, you can select that company for your investment by clicking on the respective Choose button.

For each investment round, this website randomly determines how much time you have to choose a company. In some rounds, you might have as little as 10 seconds, whereas for others you may take up to 45 seconds.

Every investment round ends as soon as you choose a company, or if you run out of time. The software will then calculate the total value of the company you have chosen, if any, and display the result.

Remember, at the very end of the study, we will pick one investment round at random, and you will be paid the value of the company that you chose in that round. If you did not choose a company in that round, then you will not receive a bonus payment.

Practice Round

We will next guide you through one practice round without a time limit, so that you can familiarize yourself with the interface.

Your performance-based reward is independent of what you do in this practice round.

Practice Round — Research Stage

Calculate each company's profit.

Remember: Profit = (Stores × Revenue per Store) - Total Expenses

Company Name	Company Name	Company Name
Stores: ____ Revenue per Store: ____ Total Expenses: ____	Stores: ____ Revenue per Store: ____ Total Expenses: ____	Stores: ____ Revenue per Store: ____ Total Expenses: ____
Enter Profit	Enter Profit	Enter Profit

Calculate and memorize the profit of each of the companies above. You will encounter the same three companies again in the choice stage. Any incorrect entries must be corrected before you will be able to proceed to the next stage.

Practice Round — Wait

You'll proceed to the choice stage in ____ seconds ...

While you wait, try to remember the profits of the three companies that you encountered during the research stage.

Practice Round - Choice Stage

Time Left: [Note clock]

Choose a company for your investment.

Remember: Value = Profit × Growth Factor

Company Name Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate	**Company Name** Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate	**Company Name** Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate
Company Name Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate	**Company Name** Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate	**Company Name** Stores: ____ Revenue per Store: ____ Total Expenses: ____ Growth Factor: ? Evaluate

You must now choose the company in which you want to invest. Before selecting a company, however, you need to reveal its growth factor by clicking on the Evaluate button. You may evaluate as many companies as you'd like before making a final choice.

Your Choice

Company Name

Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ____

Total Value: ____

This screen shows you the total value of the company that you have chosen. Thus, if you were to be paid based on your choice in this round, you would have earned a performance bonus of _____ USD.

Ready?

This concludes the practice round. By clicking on the Next button below, you will proceed to the first of 20 investment rounds.

In each investment round, you will encounter three unfamiliar companies during the research stage. In the corresponding choice stage, you will encounter the same three companies plus three new ones. Whether a particular company appears in both stages or only in the choice stage is entirely random.

For each company, we've chosen the number of stores, revenue per store, and total expenses randomly, subject to the constraint that profit is a round number between 2 and 10. Similarly, the software randomly picks a growth factor between 0.5 and 1.5 for every company.

You can thus expect to see any value of profits between 2 and 10, and any growth factor between 0.5 and 1.5 with equal probability.

Investment Round 1 of 20 — Research Stage

Calculate each company's profit.
Remember: Profit = (Stores × Revenue per Store) - Total Expenses

Company Name	Company Name	Company Name
Stores: ____ Revenue per Store: ____ Total Expenses: ____	Stores: ____ Revenue per Store: ____ Total Expenses: ____	Stores: ____ Revenue per Store: ____ Total Expenses: ____
Enter Profit	Enter Profit	Enter Profit

Please Wait

You'll proceed to the choice stage in ____ seconds ...

Investment Round 1 of 20 — Choice Stage

Time Left: ____

Choose a company for your investment.

Remember: Value = Profit × Growth Factor

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ?
Evaluate

Your Choice

Company Name
Stores: ____
Revenue per Store: ____
Total Expenses: ____
Growth Factor: ____

Total Value: ____

⋮ ⋮ ⋮

⋮ ⋮ ⋮

Results

Round	Your Choice	Stores	Revenue	Expenses	Growth	Total Value
1	Navy Blue	5	5	22	0.72	2.16
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮
20	Spring Green	1	7	5	0.57	1.14

You will be paid based on round ____. You have, therefore, earned a performance-based bonus of ____ USD. This bonus is in addition to your base payment.

Thank You

Thank you for contributing to our research. Before you go, we have just one more question that will help us improve this study.

Did you feel like the instructions you received were clear?

- Yes
- No

If you have any other comments, please enter them below. We would be very interested in hearing your feedback.

Goodbye!

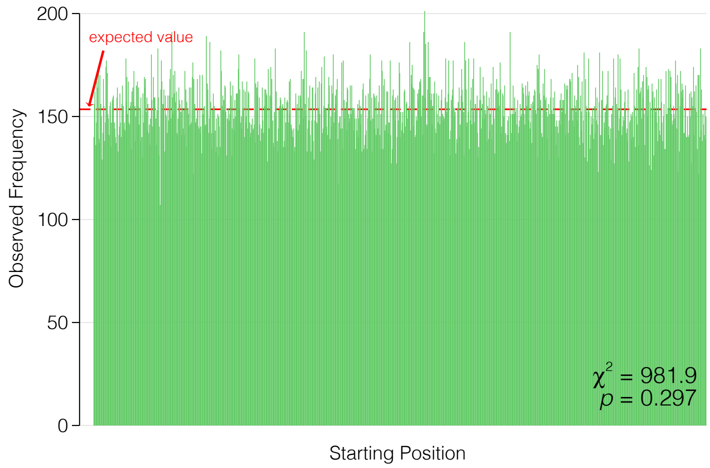
Thanks again for participating in our study!

Your Prolific completion code is: ____

Please click on the button below to be redirected to Prolific.

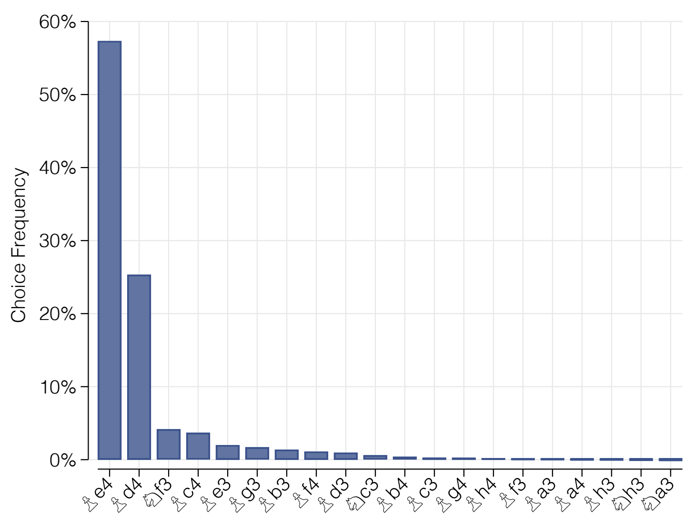
Appendix Figures

Appendix Figure AF.1: Distribution of Starting Positions in Chess960



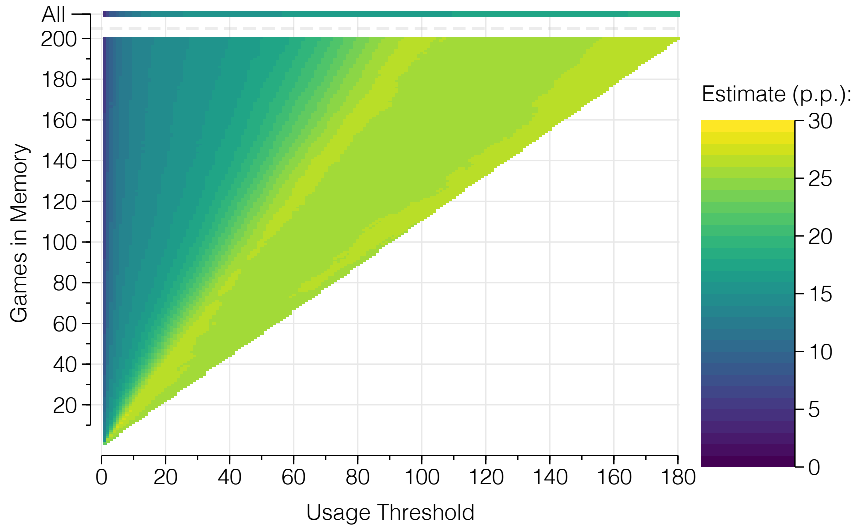
Notes: Figure shows the frequency of each starting position in our main sample of players' first game of Chess960. The reported χ^2 test refers to the null hypothesis of no differences across starting positions.

Appendix Figure AF.2: Choice Frequency of Opening Moves in Standard Chess



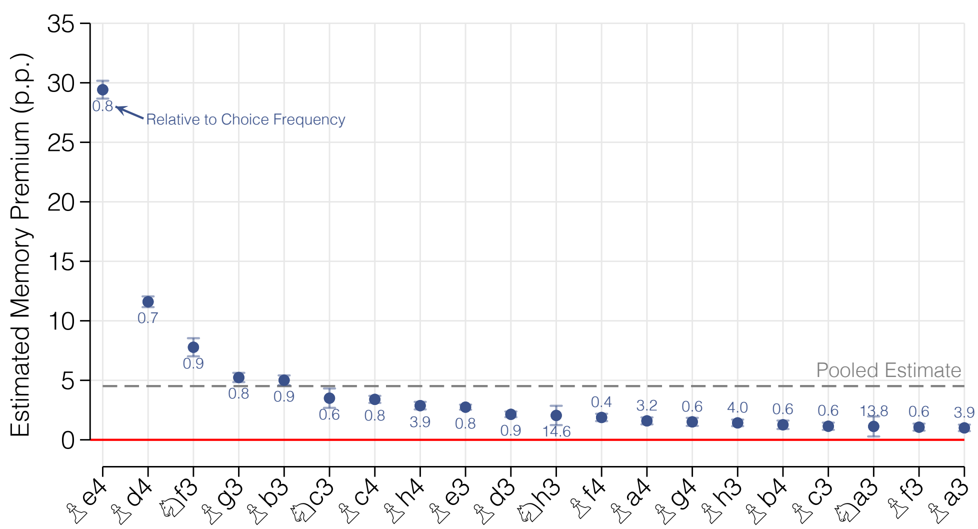
Notes: Figure shows the frequency with which the players in our main sample chose each opening move in prior games of standard chess.

Appendix Figure AF.3: Estimated Memory Premia under Alternative Definitions of Memory



Notes: Figure shows estimates of the memory premium under alternative definitions of memory. Each alternative definition corresponds to a pair (n, t), for which a given move is said to be in-memory if it has been played at least t times (x-axis) in the last n games of standard chess (y-axis). All estimates are based on the regression model and sample in col. (6) of Table 2, excluding players who have not played sufficiently many prior games for any moves to be classified as in-memory.

Appendix Figure AF.4: Heterogeneity in the Memory Premium Across Opening Moves

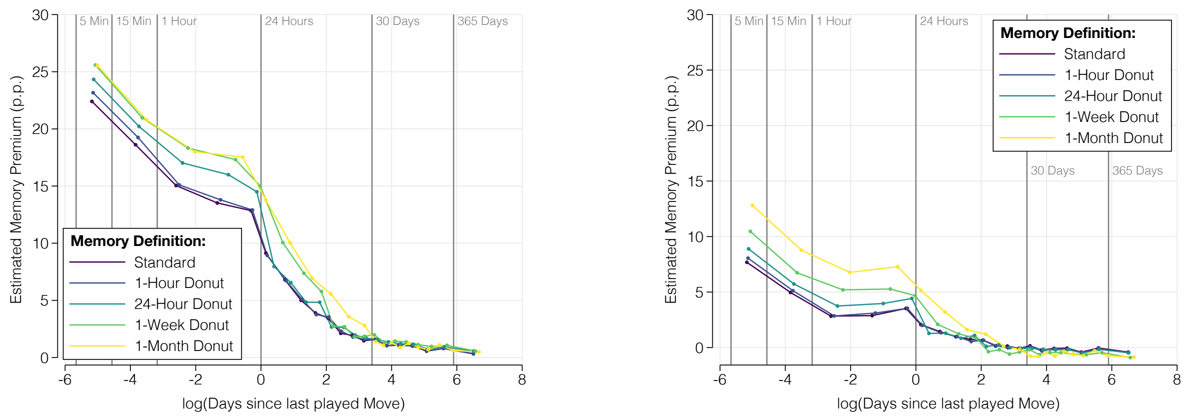


Notes: Figure shows estimates of the memory premium for each opening move in standard chess. All estimates are based on the regression model and sample in col. (6) of Table 2, allowing for heterogeneity in the memory premium across opening moves. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

Appendix Figure AF.5: Decay in the Memory Premium, Based on Donut Definitions of Memory

(a) Baseline Set of Controls

(b) Baseline and Frequency Controls

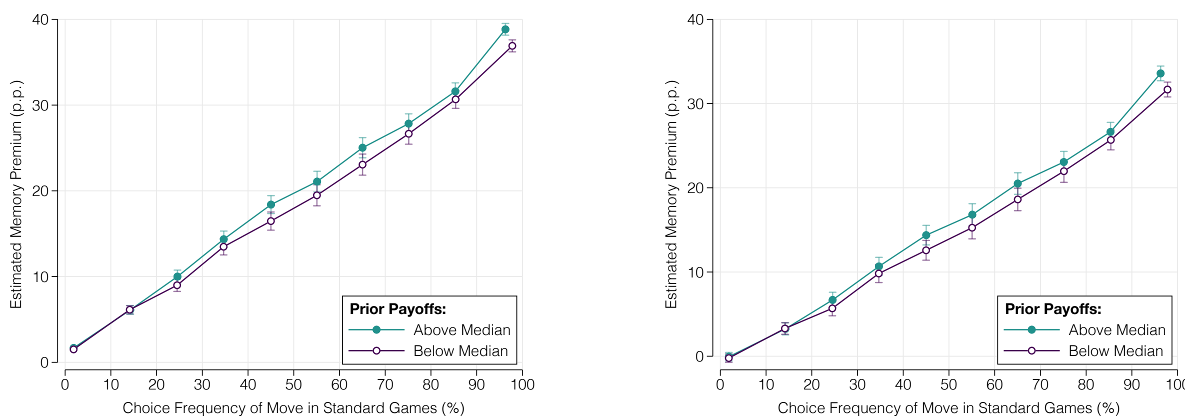


Notes: Figure replicates the analysis in Figure 8(b) in the text based on "donut definitions" of memory. According to these definitions, a move is defined as in-memory if, and only if, it had been played in at least one standard chess game on Lichess that occurred sufficiently long prior to a player's first game of Chess960. Panel (a) shows estimates of the memory premium (according to these alternative definitions) based on the regression model in col. (6) of Table 2. The estimates in panel (b) additionally control for the frequency with which a given move has previously been used. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

Appendix Figure AF.6: Horse Race Between Frequency and Prior Success

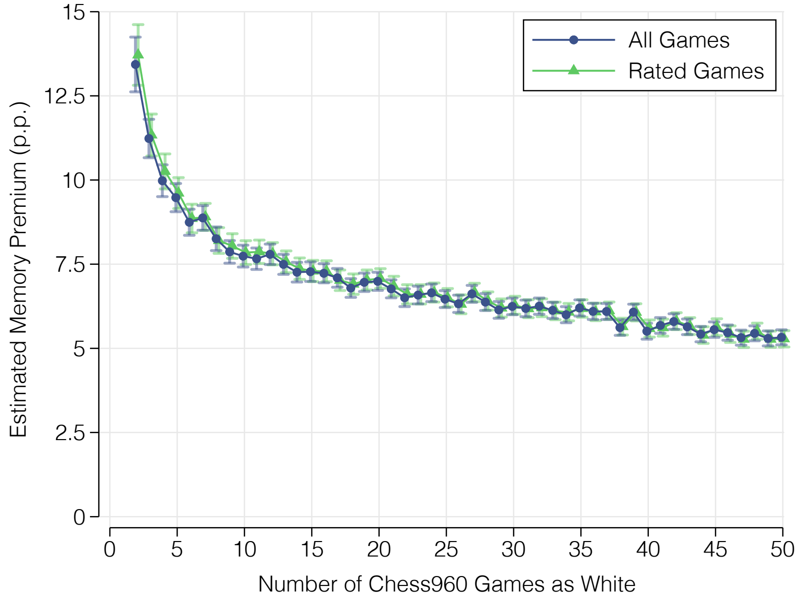
(a) Baseline Set of Controls

(b) Baseline and Recency Controls



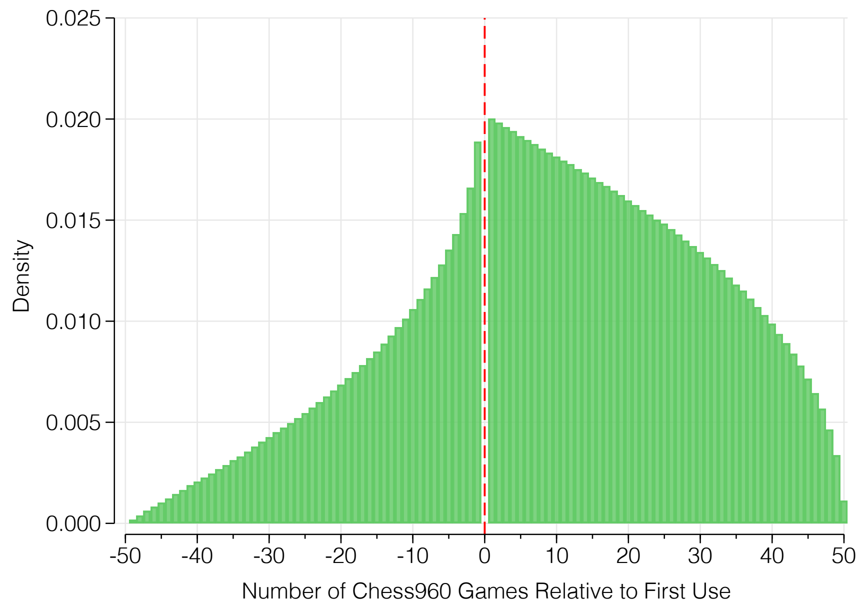
Notes: Figure replicates the analysis in Figure 8(a) in the text, differentiating between moves with above- and below-median player-specific average payoffs in prior use in standard chess games on Lichess. Payoffs are calculated by averaging over the outcomes of all games in which a player previously chose the respective move. Wins are valued at one point, draws at half a point, and losses at zero. Panel (a) shows estimates of the memory premium based on the regression model in col. (6) of Table 2. The estimates in panel (b) additionally control for how long ago the respective move has last been used in standard chess. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

Appendix Figure AF.7: Dynamics of the Memory Premium, Based on Prior Moves in Chess960



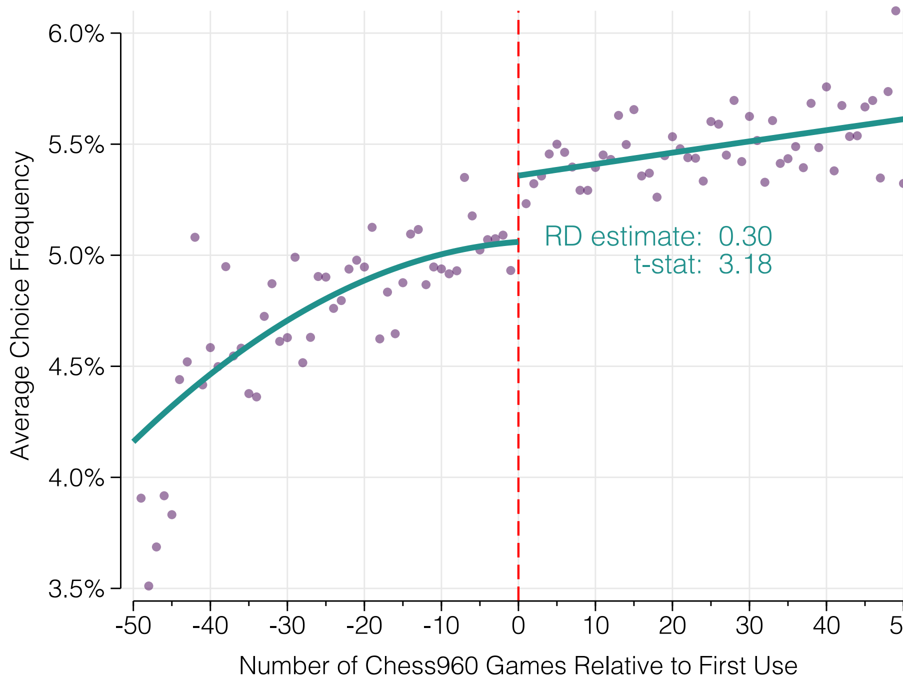
Notes: This figure replicates Figure 12 in the text for a definition of memory that is based on players' choices in previous games of Chess960.

Appendix Figure AF.8: Density of Observations around RD Threshold



Notes: Figure shows the density of observations around the threshold in the regression discontinuity design in Appendix Figure AF.9.

Appendix Figure AF.9: Evidence from an RD Design



Notes: Figure plots average choice frequencies of opening moves in Chess960 as a function of the number of Chess960 games as White relative to the respective move's first use in standard chess. Each bin averages over one game. Opening moves that had been used prior to a player's first game of Chess960 are excluded. Solid lines correspond to a second-order polynomial, fitted separately on each side of the threshold.

Appendix Tables

Appendix Table AT.1: Memory Premium, Controlling for All Other In-Memory Moves

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	12.07	3.83	3.83	4.94	4.94	4.95
	(0.04)	(0.04)	(0.04)	(0.09)	(0.09)	(0.09)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Set of Other In-Memory Moves	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Number of Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.05	0.13	0.16	0.37	0.37	0.37
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

Notes: See Table 2 in the text. The only difference between this table and Table 2 is that cols. (4)–(6) control for the set of other in-memory moves rather than the size of this set.

Appendix Table AT.2: Memory Premium, Flipped Standard Board

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
Move In-Memory (÷ 100)	3.13	2.93	3.76	3.63	3.73	3.62
	(1.04)	(1.29)	(1.57)	(1.61)	(1.58)	(1.62)
Mirror Move In-Memory (÷ 100)		0.65		1.87		1.87
		(1.24)		(1.50)		(1.51)
Fixed Effects:						
Move	Yes	Yes	No	No	No	No
Move × Size of Memory	No	No	Yes	Yes	Yes	Yes
Difference in Players' Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	Yes	Yes
Mean of LHS Variable	5	5	4.98	4.98	4.98	4.98
R^2	0.30	0.30	0.37	0.37	0.37	0.38
N	3,340	3,340	3,340	3,340	3,340	3,340

Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (3) by ordinary least squares. The sample consists of all available opening moves in players' first game of Chess960 as White, restricting attention to players who are shown the mirrored standard board (i.e., starting position 534). All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player, and are shown in parentheses.

Appendix Table AT.3: Memory Premium, Boards with Corner Royalty

A. All Moves

	Probability of Choosing Move					
	(1A)	(2A)	(3A)	(4A)	(5A)	(6A)
In-Memory (÷ 100)	11.12	3.91	3.89	4.63	4.63	4.63
	(0.15)	(0.13)	(0.13)	(0.17)	(0.17)	(0.17)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.06	5.06	5.06	5.06	5.06	5.06
R^2	0.04	0.11	0.13	0.20	0.20	0.20
N	219,952	219,952	219,952	219,952	219,952	219,952

B. Centerpawns Only

	Probability of Choosing Move					
	(1B)	(2B)	(3B)	(4B)	(5B)	(6B)
In-Memory (÷ 100)	16.71	6.81	6.78	7.76	7.58	7.20
	(0.26)	(0.26)	(0.26)	(0.32)	(0.34)	(0.35)
Centipawn Rating (÷ 100)		0.04				
		(0.01)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	13.42	13.42	13.42	13.42	13.42	13.42
R^2	0.06	0.14	0.15	0.22	0.25	0.26
N	44,488	44,488	44,488	44,488	44,488	44,488

Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (3) by ordinary least squares. In Panel A, the sample consists of all opening moves in starting postions in which the king and queen sit on squares a1, b1, g1, or h1. In Panel B, the sample is restricted to opening moves ♙d3, ♙d4, ♙e3, and ♙e4 and to positions in which the king and queen sit on squares a1, b1, g1, or h1. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player, and are shown in parentheses.

Appendix Table AT.4: Heterogeneity in the Memory Premium

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory ($\div$ 100)	4.63 (0.07)	4.36 (0.07)	3.13 (0.06)	6.49 (0.09)	6.38 (0.10)	3.56 (0.05)
Sample	Strong Players	Weak Players	High Experience Players	Low Experience Players	Close to Optimal Moves	Suboptimal Moves
Fixed Effects:						
Move $\times$ Position $\times$ Size of Memory	Yes	Yes	Yes	Yes	Yes	Yes
Player $\times$ Opponent Strength	Yes	Yes	Yes	Yes	Yes	Yes
Previous Standard Games as White	Yes	Yes	Yes	Yes	Yes	Yes
Mean of LHS Variable	5.07	5.07	5.07	5.08	9.34	3.34
R^2	0.27	0.30	0.29	0.26	0.23	0.22
N	1,495,058	1,403,397	1,446,770	1,451,685	839,173	2,059,282

Notes: Entries are coefficients and standard errors from estimating the most saturated regression model in Table 2 by ordinary least squares. The sample consists of all available opening moves in players' first game of Chess960 as White. Columns 1 and 2 distinguish between players with above and below median strength ratings in standard chess. Columns 3 and 4 split the sample at the median of the number of previous standard chess games on Lichess, whereas columns 5 and 6 distinguish between moves' whose Stockfish rating does and does not fall within 20 centipawns of that of the best available alternative in the relevant starting position. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player, and are shown in parentheses.

Appendix Table AT.5: Memory Premium, by Quality of Move in a Particular Starting Position

		Estimated Memory Premium		
Move	Choice Frequency	All Starting Positions	Close to Optimal	Suboptimal
♙a3	0.3	1.00	1.03	1.01
		(0.14)	(0.53)	(0.14)
♙a4	0.5	1.59	2.03	1.44
		(0.15)	(0.34)	(0.17)
♘a3	0.1	1.13	0.37	1.07
		(0.43)	(0.63)	(0.45)
♘c3	5.7	3.50	3.74	3.27
		(0.42)	(0.62)	(0.56)
♙b3	5.5	5.01	7.60	3.68
		(0.21)	(0.46)	(0.21)
♙b4	2.0	1.26	1.43	1.17
		(0.18)	(0.45)	(0.19)
♙c3	1.9	1.14	1.60	1.01
		(0.15)	(0.42)	(0.16)
♙c4	4.5	3.40	3.92	2.89
		(0.15)	(0.25)	(0.18)
♙d3	2.5	2.14	1.97	2.19
		(0.12)	(0.27)	(0.14)
♙d4	17.4	11.60	12.44	10.43
		(0.23)	(0.31)	(0.34)
♙e3	3.3	2.74	2.69	2.78
		(0.11)	(0.22)	(0.13)
♙e4	35.1	29.42	28.28	31.10
		(0.38)	(0.54)	(0.54)
♙f3	1.9	1.07	1.07	1.07
		(0.15)	(0.43)	(0.15)
♙f4	4.3	1.88	1.65	2.05
		(0.16)	(0.25)	(0.20)
♘f3	8.8	7.78	8.13	7.43
		(0.39)	(0.58)	(0.52)
♘h3	0.1	2.05	–0.17	2.09
		(0.41)	(0.35)	(0.44)
♙g3	6.3	5.24	7.00	4.41
		(0.20)	(0.44)	(0.21)
♙g4	2.4	1.51	1.29	1.55
		(0.17)	(0.47)	(0.18)
♙h3	0.4	1.42	2.28	1.35
		(0.14)	(0.72)	(0.14)
♙h4	0.7	2.87	3.26	2.76
		(0.16)	(0.46)	(0.17)

Notes: Entries are average choice frequencies and estimated memoria premia for all standard opening moves, either in all starting positions or in starting positions in which the move is and is not close to optimal (i.e., within 20 centipawns of the optimal move in a given position). All estimates are based on the regression model and sample in col. (6) of Table 2, allowing for heterogeneity in the memory premium across opening moves. Estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player and shown in parentheses.

Appendix Table AT.6: Memory Premium, 1-Hour Donut-Definition of Memory

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	12.09	3.82	3.82	4.51	4.51	4.51
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.05	0.13	0.16	0.23	0.23	0.23
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

Notes: See Table 2 in the text. The only difference between this table and Table 2 is that we apply a donut definition of memory, according to which a move is said to be in-memory if, and only if, it had been played in at least one standard chess game on Lichess that occured at least one hour prior to a player's first game of Chess960.

Appendix Table AT.7: Memory Premium, 1-Day Donut-Definition of Memory

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	12.15	3.83	3.82	4.75	4.75	4.74
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.05	0.13	0.16	0.23	0.23	0.23
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

Notes: See Table 2 in the text. The only difference between this table and Table 2 is that we apply a donut definition of memory, according to which a move is said to be in-memory if, and only if, it had been played in at least one standard chess game on Lichess that occured at least one day prior to a player's first game of Chess960.

Appendix Table AT.8: Memory Premium, 1-Week Donut-Definition of Memory

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	11.94	3.54	3.53	5.16	5.16	5.15
	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.05	0.13	0.16	0.23	0.23	0.23
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

Notes: See Table 2 in the text. The only difference between this table and Table 2 is that we apply a donut definition of memory, according to which a move is said to be in-memory if, and only if, it had been played in at least one standard chess game on Lichess that occured at least one week prior to a player's first game of Chess960.

Appendix Table AT.9: Memory Premium, 1-Month Donut-Definition of Memory

	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	11.29	2.87	2.87	5.16	5.16	5.16
	(0.05)	(0.04)	(0.04)	(0.07)	(0.07)	(0.07)
Centipawn Rating (÷ 100)		0.03				
		(0.00)				
Fixed Effects:						
Move	No	Yes	No	No	No	No
Move × Position	No	No	Yes	No	No	No
Move × Position × Size of Memory	No	No	No	Yes	Yes	Yes
Player × Opponent Strength	No	No	No	No	Yes	Yes
Previous Standard Games as White	No	No	No	No	No	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	5.08	5.08
R^2	0.03	0.13	0.16	0.22	0.22	0.22
N	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455	2,898,455

Notes: See Table 2 in the text. The only difference between this table and Table 2 is that we apply a donut definition of memory, according to which a move is said to be in-memory if, and only if, it had been played in at least one standard chess game on Lichess that occured at least one month prior to a player's first game of Chess960.

Appendix Table AT.10: Memory Premia for Second and Third Moves

A. Second Move

	Probability of Choosing Move				
	(1A)	(2A)	(3A)	(4A)	(5A)
In-Memory (÷ 100)	15.54	14.06	14.06	9.93	9.94
	(0.11)	(0.14)	(0.14)	(0.19)	(0.19)
Fixed Effects:					
Starting Position × Move × Size of Memory	No	Yes	Yes	No	No
Board Position × Move × Size of Memory	No	No	No	Yes	Yes
Own × Opponent Strength	No	No	Yes	No	Yes
Previous Standard Games as White	No	No	Yes	No	Yes
Mean of LHS Variable	4.21	4.21	4.21	4.21	4.21
R^2	0.02	0.14	0.14	0.41	0.41
N	3,361,249	3,361,249	3,361,249	3,361,249	3,361,249

B. Third Move

	Probability of Choosing Move				
	(1B)	(2B)	(3B)	(4B)	(5B)
In-Memory (÷ 100)	28.26	26.83	26.82	17.69	17.83
	(0.46)	(0.58)	(0.59)	(1.26)	(1.30)
Fixed Effects:					
Starting Position × Move × Size of Memory	No	Yes	Yes	No	No
Board Position × Move × Size of Memory	No	No	No	Yes	Yes
Own × Opponent Strength	No	No	Yes	No	Yes
Previous Standard Games as White	No	No	Yes	No	Yes
Mean of LHS Variable	3.80	3.80	3.80	3.80	3.80
R^2	0.01	0.11	0.11	0.54	0.54
N	3,639,681	3,639,681	3,639,681	3,639,681	3,639,681

Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (3) for White's second (upper panel) and third (lower panel) moves. The sample consists of all available second and third moves in players' first game of Chess960 as White. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player, and are shown in parentheses.

Appendix Table AT.11: Memory Speedup Mechanisms

	Log Deliberation Time							
	Repetition		Decay		Similarity		Interference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
In-Memory	–0.078	–0.108	–0.117	–0.160	–0.083	–0.123	–0.113	–0.153
	(0.005)	(0.007)	(0.006)	(0.009)	(0.009)	(0.014)	(0.007)	(0.012)
In-Memory × Choice Frequency in Standard Chess	–0.364	–1.440						
	(0.043)	(0.162)						
In-Memory × Log Days Since Prior Use			0.010	0.010				
			(0.001)	(0.002)				
In-Memory × Board Similarity					–0.005	–0.002		
					(0.002)	(0.003)		
In-Memory × Played Chess960 Back-to-Back							0.023	0.026
							(0.007)	(0.013)
Move	2nd	3rd	2nd	3rd	2nd	3rd	2nd	3rd
Sample	All Games	All Games	All Games	All Games	All Games	All Games	All Games	All Games
Fixed Effects:								
Board Position	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Player	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Own × Opponent Strength	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Standard Chess Experience	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Chess960 Experience	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Previous Board Similarity	No	No	No	No	No	No	Yes	Yes
Log Days Since Previous Chess960	No	No	No	No	No	No	Yes	Yes
Mean Delib. Time (seconds)	3.468	3.424	3.468	3.424	3.468	3.424	3.454	3.398
R^2	0.506	0.607	0.506	0.607	0.506	0.607	0.508	0.608
N	3,353,446	3,695,051	3,353,446	3,695,051	3,353,446	3,695,051	3,325,374	3,651,128

Notes: Entries are coefficients and standard errors from estimating variants of the regression model in eq. (4) by ordinary least squares. Columns (1) and (2) explore how the memory speedup varies according to the selected move's choice frequency in previous standard chess games. Columns (3) and (4) show comparative statics with respect to receny of the move's prior use in standard chess. Columns (5) and (6) offer comparative statics with respect to board similarity, defined as the number of pieces that are on the same squares as in standard chess. Columns (7) and (8) present evidence of interference based on whether a player engaged in a standard game of chess between the current and previous game of Chess960. Odd columns focus on players' second move, while even colums focus on third moves. The respective samples consist of second and third moves in all games of Chess960 as White for players who have played at least fifty such games. Deliberation time is measured from Black's preceding move to White's response. Choices with a deliberation time of zero are excluded. The corresponding moves were preselected by White, i.e., before Black even chose their move. Standard errors are clustered by player, and reported in parentheses.

Appendix Table AT.12: Summary Statistics for Experimental Data

Variable	Mean	SD	Percentile 5%	25%	50%	75%	95%	N
A. Subject Characteristics								
Age	41.0	13.4	22.0	31.0	38.0	50.0	66.0	501
Female (%)	48.9							501
Instructions Clear (%)	98.4							501
B. Choice Variables across Players								
Alternatives Evaluated (%)	73.0	29.3	16.7	48.3	85.0	98.2	100.0	501
Share In-Memory among Chosen Alternatives (%)	59.9	17.3	35.0	47.4	57.9	70.0	94.7	501
Share In-Memory among Evaluated Alternatives (%)	54.6	11.5	46.2	50.0	50.4	54.8	85.4	501
Share In-Memory among Alternatives Evaluated First (%)	65.7	20.9	36.8	50.0	63.2	84.2	100.0	501
C. Outcomes across Players								
Average Value of Chosen Alternatives	8.6	1.2	6.6	7.9	8.7	9.5	10.4	501
Average Value of Evaluated Alternatives	6.4	0.8	5.6	5.9	6.1	6.5	8.2	501
Average Value of Alternatives Evaluated First	6.9	1.2	5.1	6.0	6.8	7.7	8.9	501

Notes: This table displays summary statistics for selected variables in the data from our experiment. All observations are at the participant level. Each observation in Panel B and C corresponds to a participant-level average across all 20 rounds of the choice task.

Appendix Table AT.13: Experimental Results, Including Choice Failures

A. Main Sample

	Probability Chosen				Probability Evaluated				Probability Evaluated First			
	(1A)	(2A)	(3A)	(4A)	(5A)	(6A)	(7A)	(8A)	(9A)	(10A)	(11A)	(12A)
In-Memory (÷ 100)	6.60	6.82	6.81	6.81	7.06	7.10	7.10	7.10	10.58	10.73	10.74	10.74
	(0.51)	(0.51)	(0.51)	(0.51)	(0.64)	(0.63)	(0.63)	(0.63)	(0.63)	(0.63)	(0.61)	(0.61)
Fixed Effects:												
Own × Other Profit	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Display Position	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Time Limit	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Mean of LHS Variable	16.67	16.67	16.67	16.67	73.25	73.25	73.25	73.25	16.67	16.67	16.67	16.67
R^2	0.01	0.21	0.21	0.21	0.01	0.22	0.22	0.23	0.02	0.18	0.37	0.37
N	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792

B. Incl. Choice Failures

	Probability Chosen				Probability Evaluated				Probability Evaluated First			
	(1B)	(2B)	(3B)	(4B)	(5B)	(6B)	(7B)	(8B)	(9B)	(10B)	(11B)	(12B)
In-Memory (÷ 100)	6.34	6.55	6.55	6.55	6.94	7.02	7.01	7.01	10.42	10.59	10.54	10.54
	(0.49)	(0.50)	(0.50)	(0.50)	(0.63)	(0.62)	(0.62)	(0.62)	(0.61)	(0.61)	(0.60)	(0.60)
Fixed Effects:												
Own × Other Profit	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Display Position	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Time Limit	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Mean of LHS Variable	16.02	16.02	16.02	16.02	72.48	72.48	72.48	72.48	16.52	16.52	16.52	16.52
R^2	0.01	0.20	0.20	0.20	0.01	0.22	0.22	0.23	0.02	0.18	0.37	0.37
N	60,120	60,120	60,120	60,120	60,120	60,120	60,120	60,120	60,120	60,120	60,120	60,120

Notes: Entries are coefficients and standard errors from ordinary least squares regressions. In columns 1–4, the outcome is an indicator for whether an alternative is chosen. The outcome in columns 5–8 is an indicator for whether an alternative is evaluated, while that in columns 9–12 is an indicator for whether the alternative is evaluated first. Following the pre-registration, panel A includes all alternatives from choice tasks in which participants submitted a timely choice. Panel B includes all observations, including observations from choice tasks in which participants ran out of time. In that case, none of the available alternatives are coded as having been chosen. All estimates are scaled to correspond to the percentage-point change in the probability of the corresponding outcome associated with a one-unit increase in the respective regressor. Standard errors are clustered at the participant level, and are shown in parentheses.

Appendix Table AT.14: Experimental Results, by Ranking of Alternatives in the Research Stage

	Probability Chosen				Probability Evaluated				Probability Evaluated First			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Best In-Memory (÷ 100)	22.04	16.37	16.37	16.37	17.21	14.67	14.67	14.71	25.68	24.54	24.74	24.74
	(1.01)	(1.04)	(1.04)	(1.04)	(1.15)	(1.08)	(1.08)	(1.08)	(1.34)	(1.33)	(1.30)	(1.30)
Second Best In-Memory (÷ 100)	–1.16	1.78	1.79	1.79	2.54	3.70	3.74	3.71	1.49	2.54	2.79	2.79
	(0.49)	(0.51)	(0.51)	(0.51)	(0.67)	(0.68)	(0.68)	(0.68)	(0.47)	(0.53)	(0.48)	(0.48)
Third Best In-Memory (÷ 100)	–6.07	–0.46	–0.48	–0.48	–1.96	0.63	0.58	0.58	–0.04	1.28	0.75	0.75
	(0.43)	(0.45)	(0.45)	(0.45)	(0.56)	(0.55)	(0.55)	(0.55)	(0.48)	(0.54)	(0.45)	(0.45)
Fixed Effects:												
Own × Other Profit	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Display Position	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Time Limit	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Mean of LHS Variable	16.67	16.67	16.67	16.67	73.25	73.25	73.25	73.25	16.67	16.67	16.67	16.67
R^2	0.06	0.22	0.22	0.22	0.02	0.23	0.23	0.24	0.07	0.21	0.40	0.40
N	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792	57,792

Notes: Entries are coefficients and standard errors from ordinary least squares regressions. In columns 1–4, the outcome is an indicator for whether an alternative is chosen. The outcome in columns 5–8 is an indicator for whether an alternative is evaluated, while that in columns 9–12 is an indicator for whether the alternative is evaluated first. The regressors correspond to indicators for whether the alternative has the highest, second highest, or third highest expected value among all in-memory options. The sample includes all alternatives from choice tasks in which participants submitted a timely choice. All estimates are scaled to correspond to the percentage-point change in the probability of the corresponding outcome associated with a one-unit increase in the respective regressor. Standard errors are clustered at the participant level, and are shown in parentheses.

Appendix Table AT.15: Results from Robustness Experiment

	Outcome		
	Chosen	Evaluated	Evaluated First
A. All In-Memory Alternatives			
In-Memory	2.55	2.71	5.67
	(0.75)	(0.79)	(0.82)
B. Most Promising In-Memory Alternative			
In-Memory × Expected Value = 10	6.41	5.19	13.28
	(2.35)	(1.66)	(2.35)
In-Memory × Expected Value = 9	6.23	5.34	13.84
	(2.37)	(1.56)	(2.09)
In-Memory × Expected Value = 8	10.31	5.31	16.11
	(2.45)	(1.62)	(2.30)
In-Memory × Expected Value = 7	9.73	7.44	11.57
	(2.33)	(1.73)	(2.20)
In-Memory × Expected Value = 6	8.02	1.79	5.85
	(2.10)	(1.80)	(2.41)
In-Memory × Expected Value = 5	16.05	6.30	7.55
	(2.83)	(2.70)	(2.69)
In-Memory × Expected Value = 4	12.77	8.73	4.37
	(3.19)	(3.02)	(3.39)
In-Memory × Expected Value = 3	5.55	5.90	6.45
	(3.32)	(5.12)	(4.21)
In-Memory × Expected Value = 2	0.92	–3.29	2.10
	(4.59)	(14.01)	(7.47)
C. Other In-Memory Alternatives			
In-Memory × Expected Value = 9	–7.92	2.54	4.61
	(3.01)	(1.89)	(2.83)
In-Memory × Expected Value = 8	–3.47	0.54	3.24
	(2.22)	(1.72)	(1.76)
In-Memory × Expected Value = 7	–1.77	0.11	1.56
	(1.68)	(1.48)	(1.62)
In-Memory × Expected Value = 6	–1.23	3.38	0.23
	(1.54)	(1.52)	(1.63)
In-Memory × Expected Value = 5	0.23	1.69	2.48
	(1.07)	(1.42)	(1.51)
In-Memory × Expected Value = 4	–0.18	–1.69	0.01
	(1.01)	(1.63)	(1.50)
In-Memory × Expected Value = 3	–0.59	1.00	1.14
	(0.85)	(1.73)	(1.46)
In-Memory × Expected Value = 2	0.80	2.01	1.98
	(0.85)	(1.64)	(1.38)

Notes: Entries are coefficients and standard errors from ordinary least squares regressions replicating the results in panels (a) and (b) of Figure 15, based on the data from the robustness experiment described in Appendix F. Panel A compares mean outcomes for all in-memory alternatives with all out-of-memory alternatives. Panels B and C compare the most promising in-memory alternative (Panel B) and the remaining in-memory alternatives (Panel C) with out-of-memory alternatives of the same expected value. The sample corresponds to all alternatives from choice tasks in which participants submitted a timely choice. Standard errors are clustered at the participant level, and are shown in parentheses.

Appendix Table AT.16: Ancillary Experimental Results

	Log Deliberation Time						Number of Evaluations			Optimization Error		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Choose In-Memory Option	–0.13	–0.09	–0.08				–0.53	–0.24	–0.22			
	(0.02)	(0.01)	(0.01)				(0.07)	(0.03)	(0.03)			
Number of Evaluations				0.11	0.15	0.12				9.33	9.28	9.40
				(0.01)	(0.01)	(0.01)				(0.27)	(0.40)	(0.40)
Log Deliberation Time										–8.04	–3.51	1.32
										(1.08)	(1.14)	(1.23)
Fixed Effects:												
Subject	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Time Limit	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Mean of LHS Variable	2.56	2.56	2.56	2.56	2.56	2.56	4.39	4.39	4.39	32.40	32.40	32.40
R^2	0.01	0.46	0.57	0.18	0.55	0.62	0.02	0.69	0.71	0.15	0.27	0.28
N	9,632	9,632	9,632	9,632	9,632	9,632	9,632	9,632	9,632	9,632	9,632	9,632

Notes: Entries are coefficients and standard errors from ordinary least squares regressions. In columns 1–6, the outcome of interest is the logarithm of the time that a subject spent in the choice stage of a given round. The outcome in columns 7–9 is the number of evaluated alternatives in a given round, while that in columns 10–12 is an indicator for choosing an alternative that is suboptimal relative to the best option that the subject evaluated. The regressors shown on the left correspond to an indicator for whether a subject chose an in-memory alternative in that round, the number of alternatives they evaluated, and the logarithm of the deliberation time. The estimates in columns 10–12 are scaled to correspond to the percentage-point change in the probability of choosing a suboptimal alternative associated with a one-unit increase in the respective regressor. The sample includes all choice tasks in which participants submitted a timely choice. Standard errors are clustered at the participant level, and are shown in parentheses.

Appendix Table AT.17: Estimates from Causal Research Designs

A. RD Estimates

	Probability of Choosing Move					
	(1A)	(2A)	(3A)	(4A)	(5A)	(6A)
In-Memory (÷ 100)	0.25 (0.09)	0.22 (0.10)	0.29 (0.11)	0.26 (0.12)	0.37 (0.14)	0.36 (0.16)
Sample	All Games	Rated Games	All Games	Rated Games	All Games	Rated Games
Estimator	OLS	OLS	OLS	OLS	Calonico et al. (2014)	Calonico et al. (2014)
Polynomial Bandwidth	Linear 20	Linear 20	Linear 10	Linear 10	Linear Optimal	Linear Optimal
Mean of LHS Variable	5.26	5.28	5.24	5.27	5.25	5.29
N	1,219,263	1,080,203	687,013	602,314	565,062	493,419

B. Difference-in-Differences Estimates

	Probability of Choosing Move					
	(1B)	(2B)	(3B)	(4B)	(5B)	(6B)
In-Memory (÷ 100)	0.43 (0.05)	0.36 (0.05)	0.43 (0.07)	0.42 (0.07)	0.26 (0.06)	0.21 (0.06)
Sample	All Games	Rated Games	All Games	Rated Games	All Games	Rated Games
Estimator	TWFE	TWFE	Borusyak et al. (2024)	Borusyak et al. (2024)	Cengiz et al. (2019)	Cengiz et al. (2019)
Fixed Effects:						
Player × Move	Yes	Yes	Yes	Yes	No	No
Chess960 Games as White	Yes	Yes	Yes	Yes	No	No
Player × Move × Stack	No	No	No	No	Yes	Yes
Chess960 Games as White × Stack	No	No	No	No	Yes	Yes
Mean of LHS Variable	5.08	5.09	3.47	3.45	3.27	3.27
N	16,453,129	14,704,058	12,137,434	10,734,051	15,501,782	13,833,500

Notes: Entries in Panel A are estimates of the memory premium from a regression discontinuity design around the first time a player uses a particular opening move in standard chess. Columns 1A–4A report estimates based on a standard linear regression model, restricting attention to observations within the respectively indicated bandwidths around the threshold and allowing for different slopes in the running variable on either side of the discontinuity. Columns 5A and 6A report bias-corrected regression discontinuity estimates based on the procedure of Calonico et al. (2014), using a uniform kernel and the associated MSE-optimal bandwidth. The sample in Panel A consists of all available opening moves in players' first fifty games of Chess960 as White, provided that they enter players' memory after the first game of Chess960. A move is said to enter a player's memory when she first chooses it in a standard chess game on Lichess. Entries in Panel B are difference-in-differences estimates of the memory premium. Columns 1B and 2B report results from a standard two-way fixed effects estimator with player-by-move and number-of-Chess960-games fixed effects, applied to the sample of all available moves in players' first fifty games of Chess960 as White. Columns 3B and 4B implement the estimator of Borusyak et al. (2024). The sample in these columns corresponds to all player-move combinations in which the move enters the player's memory either after her first game of Chess960 or not at all. Columns 5B and 6B implement the stacked difference-in-differences estimator of Cengiz et al. (2019). Each stack consists of moves that enter a given player's memory at a particular point in time as well as moves that remain out of the same player's memory over the entire sample period. In these specifications, all identifying variation comes from within-player comparisons of treated and never-treated moves over time. All estimates are scaled to correspond to the percentage-point change in choice probability associated with a one-unit increase in the respective regressor. Standard errors are clustered by player and shown in parentheses.