The Memory Premium*

Yuval Salant,ⓡ *Northwestern University*
Jörg L. Spenkuch,ⓡ *Northwestern University*
David Almog, *Northwestern University*

March 2025

Abstract

We explore the role of memory for choice behavior in unfamiliar environments. Using a unique data set, we document that decision makers exhibit a "memory premium." They tend to choose in-memory alternatives over out-of-memory ones, even when the latter are objectively better. Consistent with well-established regularities regarding the inner workings of human memory, the memory premium is associative, subject to interference and repetition effects, and decays over time. Even as decision makers gain familiarity with the environment, the memory premium remains economically large. Our results imply that the ease with which past experiences come to mind plays an important role in shaping choice behavior.

*The first two authors appear in random order. We are grateful to Thibault Duplessis and Theo Wait for providing us with data from Lichess, and for patiently answering our questions. 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. Correspondence can be addressed to the authors at: MEDS Department, Kellogg School of Management, 2211 Campus Dr, Evanston, IL 60208; or by email to `y-salant@kellogg.northwestern.edu`, `j-spenkuch@kellogg.northwestern.edu`, and `david.almog@kellogg.northwestern.edu`

1. Introduction

Memory provides context for human decision making, allowing us to use past experiences to guide future choices. The mechanisms through which memory affects choice behavior, however, remain only incompletely understood. Cognitive scientists generally conceptualize memory as a three-stage process: encoding, storage, and retrieval (Melton 1963; Baddeley et al. 2009). In the encoding stage, the brain transforms sensory inputs into neural representations. These representations are consolidated and stored over time before being accessed and utilized during the retrieval stage. In this paper, we explore the implications of memory retrieval for real-world choice behavior.

We document that, in unfamiliar environments, decision makers are more likely to choose alternatives that are retrievable from memory than those that are not, even when the latter are objectively better. This phenomenon, which we call the "memory premium," can be understood as an extension of the availability heuristic of Tversky and Kahneman (1973) to the domain of choice. Whereas the availability heuristic concerns judgments and probability assessments, the memory premium relates recall to choice behavior.[1]

Our second contribution is to develop a set of testable predictions that relate the cognitive mechanisms governing recall to choice behavior and the memory premium. Previous research in psychology and neuroscience establishes four stylized facts about the retrieval process (see, e.g., Kahana 2012): (i) Memory is associative. Mental representations are interconnected, and past experiences are more easily recalled when they are related to current environmental cues. (ii) Memory decays over time. Recent experiences are more likely to be retrievable than temporally distant ones. (iii) Memory is strengthened by repetition. Experiences that are encountered more frequently are more likely to be recalled than single-instance encounters. (iv) Retrieval is subject to interference. Newer memories disrupt the recall of older ones, and vice versa. Taking these mechanisms as given, we document that the memory premium exhibits the same comparative statics. It is associative, subject to repetition and interference effects, and it decays over time. As a result, the cognitive mechanisms behind recall also predict choice behavior.

To establish these empirical regularities, we draw on a unique 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

[1] In the words of Tversky and Kahneman (1973, p. 208), "a person is said to employ the availability heuristic whenever he estimates frequency and probability by the ease with which instances or associations could be brought to mind."

the pieces on players' back ranks are randomized, resulting in 960 different possible starting positions.

Chess960 provides an almost ideal setting to study the influence of memory on choice behavior 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 pieces can have profound implications for optimal strategy, opening moves that tend to perform well in classical chess need not be a good choice in a particular starting position in Chess960. As a consequence, players should rely on their analytical skills rather than memory when selecting an opening move. In our analysis, we ask to which extent they nonetheless draw on memory and how this reliance depends on features of the new environment.

Our data come from `lichess.org`, one of the most popular internet chess servers. Every day, Lichess hosts several million games of classical chess and thousands of games of Chess960.[2] We have information on all moves in the universe of games on the platform from January 2013 through June 2021.

There are two key challenges to empirically documenting the memory premium 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 chess 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 and the corresponding outcomes 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

[2] Lichess also allows users to play other variants of chess, most of which are more obscure than Chess960.

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. 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 only about 3.2 p.p. The average memory premium is, therefore, large enough to overcome even significant differences in the objective quality of moves.

The average memory premium, however, masks considerable heterogeneity. The memory premium of inexperienced chess players is about twice as large as that of experienced ones, and "good" opening moves exhibit larger premia than "bad" ones. Moreover, opening moves that are common in standard chess exhibit much higher premia than moves that are only rarely used. Importantly though, we detect an economically meaningful memory premium for all opening moves, even in starting positions in which they are not close to being optimal.

In order to link the observed differences in choice frequencies to the inner workings of human memory, we draw on extant work in neuroscience and psychology to develop comparative statics that the memory premium should satisfy. These comparative statics bridge the gap between well-established cognitive mechanisms and choice behavior. For example, based on associative recall, we predict that the memory premium increases in the similarity between the starting position in Chess960 and the standard chess board. Measuring similarity as the number of chess pieces that are placed on the same square, this is indeed what we find. Overall, we provide strong evidence that choice behavior is moderated by the same cognitive forces that have previously been shown to influence recall.

Our main results focus on players' first game of Chess960. We, therefore, document the existence of the memory premium in an unfamiliar choice environment. Does the memory premium decrease as individuals become more familiar with the setting? In order to speak to this question, we expand the scope of our analysis to all opening moves in players' first fifty games of Chess960 as White. As expected, the memory premium shrinks considerably as individuals gain experience; but it does not vanish. 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 addition, we provide evidence to suggest that a move entering memory *causes* higher choice frequencies for this move going forward. Specifically, we analyze the usage of a given opening move in Chess960 before and after a player chose that move for the first time in a regular chess game. We then compare the change in usage with the corresponding change for moves that remained out of memory. Our difference-in-differences estimates imply an increase in the choice frequency between 0.2 an 0.4 p.p., and a regression discontinuity approach shows that there is a discrete jump of about 0.3 p.p. immediately after the move's first use in standard chess.

While these dynamic estimates support a causal link between memory and choice behavior, they are significantly smaller than the estimated memory premium in the first part of our analysis. There are two reasons for why this is the case. First, on average, the moves that identify our causal estimates enter players' memories (according to our workhorse definition) between the 16th and 17th game of Chess960. At this point, the average memory premium has already declined considerably. Second, an opening move being used once in regular chess constitutes a very weak treatment. By contrast, the moves that identify the memory premium in our main analysis had, on average, been played more than a hundred times before. Since repetition improves recall, the estimates from our causal research design *should* be much smaller than the average memory premium.

Broadly summarizing, our empirical analysis shows that choice behavior varies in predictable ways depending on how easily prior experiences come to mind. While alternative explanations might account for some of our findings in isolation, memory retrieval offers a simple, unified mechanism to explain *all* of them. Memory—and its natural counterpart, forgetting—provide a particularly straightforward explanation for why the influence of prior experiences decays over time. Taken together, our results imply that memory retrieval plays a fundamental role in shaping choice behavior, especially in unfamiliar situations where analytical reasoning might otherwise be expected to dominate.

We conclude by proposing that relying on memory to guide decision-making in unfamiliar environments may be boundedly rational; it can economize on cognitive costs while yielding favorable outcomes—at least if the new environment is not too dissimilar from the usual one. We illustrate this point by drawing directly on our data. Since we observe the objective value of all available moves, we can ask how the quality of players' choices would vary under different choice procedures. Our counterfactuals establish that a naïve decision maker could, in expectation, improve upon random choice by randomly selecting an in-memory alternative; and the magnitude of the improvement would increase in the degree of similarity between the starting position in Chess960 and that in standard chess. Our counterfactuals also consider decision makers who maximize from consideration sets (Manzini and Mariotti 2014). We show that, except in highly dissimilar board positions, restricting the consideration set to in-memory moves would cost such a decision maker very little. Analyzing players' *actual* choices, we observe that, on average, they do better when they choose in- rather than out-of-memory moves. Only in the most dissimilar starting positions do in-memory choices tend to be worse.

Related Literature Our findings contribute to a growing literature on the role of memory in economic decision-making. An early decision-theoretic contribution to this literature is Gilboa and Schmeidler (1995, 2001). In their theory of case-based decisions, the value of an alternative is a weighted sum of its past performance, with larger weights attached to outcomes in decision problems that are more similar to the current one. In another early contribution, Mullainathan (2002) develops a model in which associative recall and rehearsal affect beliefs about a state variable, such as income. More recently, Bordalo et al. (2020, 2022, 2024) incorporate associative memory and interference into models of belief formation. For example, Bordalo et al. (2020) present a theory of consumer choice with selective recall. In their model, the decision maker relies on memories of past purchases to form a norm, or expectation, about the quality and price of the good in the current environment, which in turn influences purchase decisions. Other important contributions include Azeredo da Silveira et al. (2019, 2024), Nagel and Xu (2022); and Wachter and Kahana (2024).

Unlike extant theoretical work, our investigation does not micro-found how decision makers use memory to evaluate an object or form expectations about a payoff-relevant parameter. Instead, we examine the effect of memory on the mapping from choice situations to outcomes. As we explain in Section 2, there are several natural choice procedures that can give rise to the memory premium, including maximization with memory-dependent perceptions of utility and satisficing.

There is also an empirical literature on the effect of memory on belief formation and decision-making (see, e.g., Afrouzi et al. 2023; Charles 2022, 2024; D'Acunto et al. 2021; Enke et al. 2024; Jiang et al. 2024). In influential work, Malmendier and Nagel (2011, 2016) show that investors who lived through the Great Depression are less likely to invest in the stock market later in life, and that experiencing high inflation has lasting effects on inflation expectations.[3] Hartzmark et al. (2021) draw on associative memory to explain why individuals over-extrapolate from signals about goods they already own. We contribute to this empirical literature by identifying a novel feature of memory-based choice behavior. More broadly, our findings help to shed light on the cognitive mechanisms underlying real-world decision-making.

2. Conceptual Framework

To guide our empirical analysis, we develop a simple framework that maps choice situations into choice probabilities, conditional on the memory of the decision maker (DM). We do not attempt to micro-found how memory affects the evaluation of the available alternatives. Rather, we propose a new feature of memory-based choice behavior, which we call the memory premium. Drawing on previous research on human memory, we also formulate comparative

[3]See also Malmendier and Wachter (2024) for a survey of the literature on how past experiences shape expectations and behavior.

statics that the memory premium should satisfy. These comparative statics provide additional testable predictions about how memory influences choice behavior in unfamiliar environments.

2.1. *Preliminaries*

Fix a grand set of alternatives $\mathcal{X}$ and a collection of environments $\mathcal{E}$. Alternatives have objective value to the DM, which may depend on the environment. We use $U(x, E)$ to denote the value of alternative $x \in \mathcal{X}$ in environment $E \in \mathcal{E}$. Environments may be more or less similar to one another, with $\sigma(E_1, E_2)$ denoting the degree of similarity between environments E_1 and E_2. A choice situation is a pair (A, E), where $A \subset \mathcal{X}$ includes all the alternatives available for choice, and E is the environment in which choices are to be made.

For simplicity, assume that the DM encountered a fixed choice situation $\hat{S} = (\hat{A}, \hat{E})$ in the past. Her memory consists of all previous choices and the associated outcomes. The DM's memory is thus a vector of experiences $M = (m(1), m(2), ...)$, where $m(k) = (a(k), o(k))$ includes the alternative $a(k) \in \hat{A}$ that was chosen k periods ago as well as the realized outcome, $o(k) \in \mathcal{R}$. We refer to an alternative a as "in memory" if there exists a k such that $a = a(k)$. Alternatives that are not in memory are said to be "out of memory." Given a memory M and an out-of-memory alternative a, let M^{+a} denote another memory that is obtained by adding to M one or more experiences in which a was chosen.

2.2. *Memory Premium*

We are interested in studying memory-related properties of the random choice function $C\left(S, a \mid M, U\right)$. This function assigns, to every choice situation $S = (A, E)$ and any alternative $a \in A$, the probability that the DM chooses a in S, conditional on memory M and value function U. Randomness in choice may arise from two or more sources. First, retrieval from memory is probabilistic. For example, the likelihood that a particular experience can be retrieved from memory may depend on how often or how recently the corresponding alternative has been chosen in the past. Since retrieval is not deterministic, choice behavior that depends on memory may be non-deterministic as well. Second and irrespective of memory, value assessments in complex choice situations may be noisy, which too can lead to randomness in choice (Woodford 2020; Salant and Spenkuch 2024).

We assume that C satisfies two standard properties with respect to the value function U. First, conditional on memory, C is monotone in U. That is, for any two alternatives a and b with the same memory status, if $U(a, E) > U(b, E)$ then $C\left(S, a \mid M, U\right) \geq C\left(S, b \mid M, U\right)$. Second, C is continuous in U, so that, all else equal, small changes in value lead to small changes in choice probabilities. In what follows, we simplify notation by omitting the dependence of C on U.

To formalize how memory might affect choice behavior, let

$$\Delta(S, x \mid M^{+a}) = C\left(S, x \mid M^{+a}\right) - C\left(S, x \mid M\right)$$

denote the change in the choice probability of alternative x in situation S as a "enters" memory. With this notation in hand, we define the memory premium as follows.

Definition 1 (Memory Premium): *A choice function C gives rise to a* memory premium *if for any choice situation $S = (A, E)$ and any two memories M and M^{+a}:*
(i) $\Delta(S, a \mid M^{+a}) > 0$, and
(ii) $\Delta(S, a \mid M^{+a}) > \Delta(S, b \mid M^{+a})$ for any alternative $b \neq a$.

In words, choice behavior exhibits a memory premium if an alternative is more likely to be chosen on account of being in- rather than out-of-memory; and if the increase in choice probability exceeds any potential increase in the choice probability of any other alternative.[4]

When C satisfies monotonicity and continuity with respect to U, Definition 1 implies that the DM is more likely to choose an in-memory alternative than an out-of-memory one, even when the value of the former is smaller than that of the latter. In particular, an in-memory alternative is more likely to be chosen than an equally good out-of memory alternative.

Observation 1: *Suppose C satisfies Definition 1, and is monotone and continuous in U. For any choice situation $S = (A, E)$, there exists a threshold $\epsilon > 0$ such that if $U(b, E) - U(a, E) \leq \epsilon$ for two alternatives $a, b \in A$ that are out of memory M, then $C\left(S, a \mid M^{+a}\right) > C\left(S, b \mid M^{+a}\right)$.*

Several natural choice procedures can generate a memory premium. One example is a risk-neutral utility maximizer who perceives the value of in-memory alternatives to be systematically higher than it actually is. Another example is a utility maximizer who is risk-averse and holds more precise beliefs about in-memory alternatives compared to out-of-memory ones. A third example is a decision maker who maximizes from consideration sets that disproportionately include in-memory alternatives. A fourth example involves satisficing. A satisficer does not consider all possible alternatives to select the best one; instead, she sequentially examines a small number of alternatives and chooses the first one she finds satisfactory (Simon 1955). Such a DM would also display a memory premium if the order of consideration prioritizes in-memory alternatives. Of course, a hybrid choice procedure that combines satisficing and maximization may generate a memory premium as well. For instance,

[4]If part (i) of Definition 1 holds, then the *average* change in the choice probability of other alternatives must be negative. Part (ii) recognizes that this need not be the case for every single alternative.

a DM who first evaluates whether one or more in-memory alternatives are satisfactory before maximizing over the remaining options would similarly exhibit a memory premium.

2.3. *Comparative Statics*

The availability heuristic suggests that individuals judge the likelihood or frequency of events by the ease with which relevant instances come to mind (Tversky and Kahneman 1973). Extending this idea to choice behavior, we conjecture that the memory premium may also vary with the ease of recall. Drawing on previous findings in cognitive psychology and neuroscience, we now develop testable implications that relate the size of the memory premium to well-established regularities in the retrieval process.

We consider four regularities that stand out as fundamental. First, memory decays over time. That is, recent experiences are more easily retrieved than temporally distant ones. If the memory premium is indeed linked to the ease or probability of recall, then we would expect a larger premium when the experience in question entered the DM's memory more recently.

PROPERTY 1 (Recency): *Consider memories $(M^{+a})_1$ and $(M^{+a})_2$, which were obtained from memory M by adding an experience (a, o) in positions k_1 and $k_2 < k_1$, respectively. Then, for any choice situation S, $\Delta(S, a \mid (M^{+a})_2) > \Delta\left(S, a \mid (M^{+a})_1\right)$.*

A second fundamental property of memory is that repetition strengthens recall. Since more frequent experiences are more likely to be retrievable than less frequent ones, we would expect that repetition translates into a larger memory premium.

PROPERTY 2 (Repetition): *Let $(M^{+a})_1$ be a memory obtained from M by adding experiences with the alternative a in positions $k_1 < \ldots < k_i$. Let $(M^{+a})_2$ be a memory obtained from $(M^{+a})_1$ by adding another experience with a in position $k_{i+1} > k_i$. Then, for any choice situation S, $\Delta(S, a \mid (M^{+a})_2) > \Delta\left(S, a \mid (M^{+a})_1\right)$.*

A third key property of memory is associativity. Past experiences are more readily retrieved when the context in which they occurred shares meaningful similarities with the current one. If choice behavior is, indeed, a function of recall, then the memory premium should be larger when the current choice environment resembles the setting in which the in-memory alternative was previously chosen.

PROPERTY 3 (Similarity): *Consider two choice situations $S_1 = (A, E_1)$ and $S_2 = (A, E_2)$. If E_1 is more similar to the previously encountered environment than E_2, i.e., $\sigma(E_1, \hat{E}) > \sigma(E_2, \hat{E})$, then $\Delta(S_1, a \mid M^{+a}) > \Delta\left(S_2, a \mid M^{+a}\right)$.*

If the memory premium is linked to the ease with which an experience comes to mind, then similarity should have another measurable effect. Recent experiences stand out more when the associated environment is dissimilar to the choice environments of the past. Since salient experiences interfere with the recall of less salient ones, we should observe a smaller memory premium in subsequent choices when the distinctiveness of the most recent experience increases.

PROPERTY 4 (Interference): *Consider memories M_1 and M_2, which are obtained by attaching two different experiences, m_1 and m_2, to the first position of memory M.*[5] *Let $(M_1)^{+a}$ and $(M_2)^{+a}$ denote the memories that result from adding alternative a to M_1 and M_2 in any position other than the first one. If the choice environment associated with m_1 is less similar to the previously encountered environment than that of m_2, i.e., $\sigma(E_1, \hat{E}) < \sigma(E_2, \hat{E})$, then $\Delta(S, a \mid (M_2)^{+a}) > \Delta\left(S, a \mid (M_1)^{+a}\right)$.*

In words, Property 4 states that the memory premium ought to be smaller when a recent experience stands out as different and thereby disrupts the recall of memories from the default choice environment.

In an analogous fashion to Property 4, imagine the following two scenarios. In the first, the DM encounters two unfamiliar environments, one after the other. In the second scenario, the DM encounters the familiar environment $\hat{E}$ in between the two unfamiliar ones. We would expect that encountering the familiar environment has two effects. It should strengthen the DM's recall of other, prior experiences from this environment, and it should interfere with her ability to recall the unfamiliar choice situation. As a consequence, we would expect a larger memory premium in the second scenario.

Our empirical analysis investigates whether choice behavior in unfamiliar environments does, indeed, exhibit a memory premium and, importantly, whether Properties 1–4 above hold. The context in which we study these questions is that of Chess960.

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

[5] In symbols, $M_i = (m_i(0), m(1), m(2), \ldots)$ where $M = (m(1), m(2), \ldots)$.

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.[6] 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 counter-moves, 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 choices 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 relationships and effects.

3.2. *Data and Descriptive Statistics*

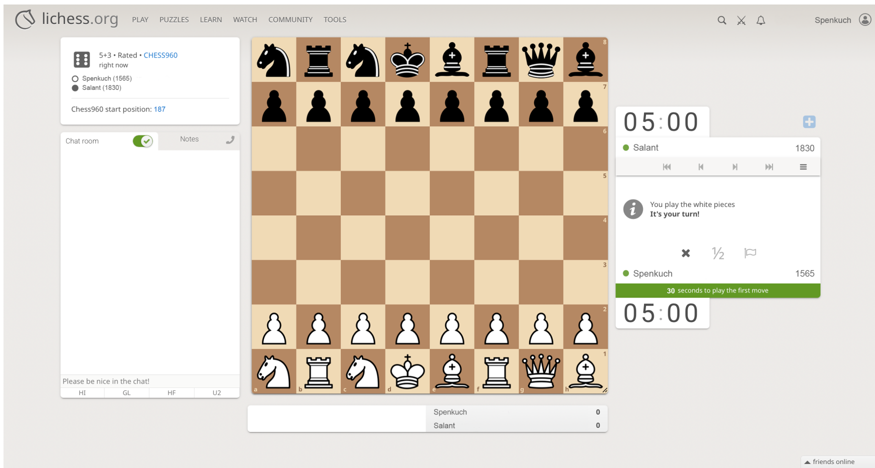
Our data come from `lichess.org`, one of the most popular online chess servers.[7] Funded by donations, Lichess is ad-free and allows anyone to play live chess games at no cost through a high-quality graphical user interface. Every day, Lichess hosts several million games of standard chess and thousands 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 causal 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. Enforcement of this policy relies on a variety of methods, including community reporting of suspected offenders and automatic detection algorithms. Figure 1 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

[6] In Chess960, the position of the king and rook after castling is exactly the same as in standard chess. Unlike in standard chess, this can mean that the king or rook "jump over" other pieces in the process of castling.

[7] Our description of these data borrows from Salant and Spenkuch (2024).

Figure 1: Chess960 on Lichess



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

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.

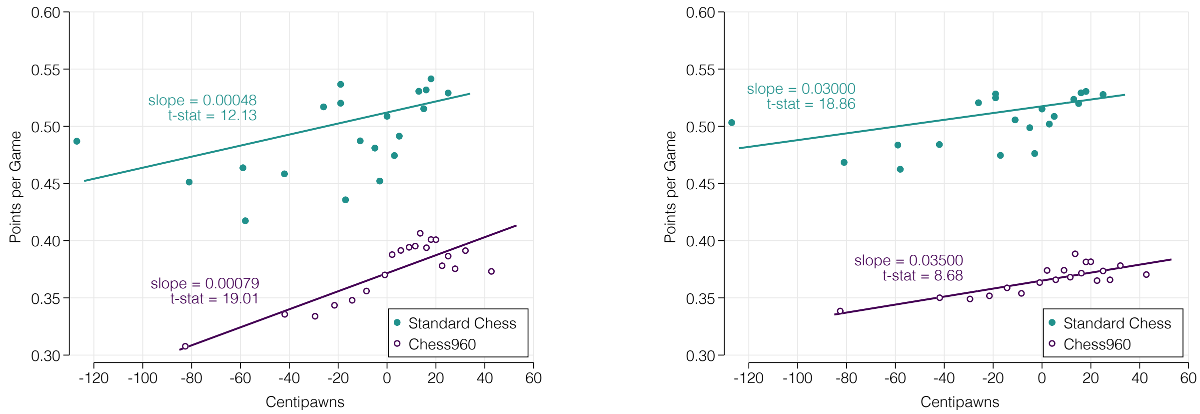
In our main analysis, we restrict attention to the opening move of approximately 147,000 players as they engage in a game of Chess960 for the first time.[8] There are two restrictions that limit the size of this sample. First, users must play their first game of Chess960 as White, and the game must be 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 initial choices in an unfamiliar decision problem with real stakes. Second, prior to their first game of 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 study opening moves in players' first fifty games of Chess960 as White, irrespective of which color they played in the first game and whether that game was rated. We also impose higher thresholds on the number of previous standard games. Reassuringly, alternative sample restrictions 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

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

Figure 2: 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.

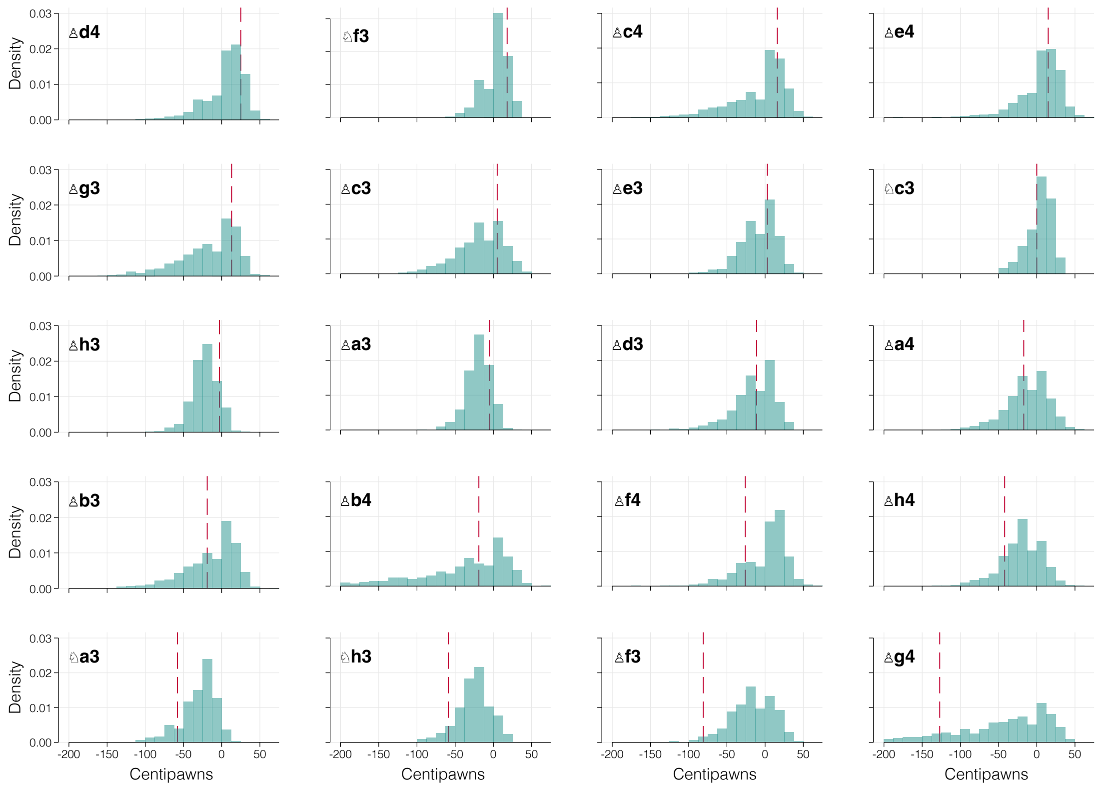
one of the most powerful open-source chess engines. It is far better at analyzing chess positions than even the best grandmasters.[9] Stockfish evaluates board configurations by examining sequences of possible moves (up to a certain depth) using a combination of heuristic evaluation functions and advanced 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.[10] Important for our purposes, Stockfish explicitly supports Chess960.

Figure 2 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-time players of Chess960 win less than half of their games—at least in part because the new environment is unfamiliar and they lack the requisite experience. 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

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

[10]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 3: 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.

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 ($\approx$ 30 centipawns) better increases White's chance of winning by about one percentage point.

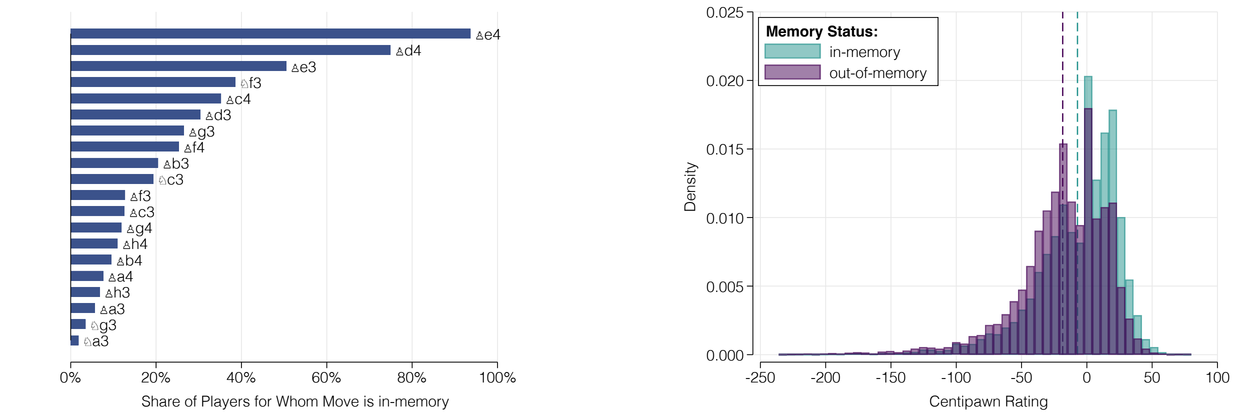
Figure 3 shows that there is a great amount of within-move variation in quality across different starting positions. Each panel in this figure depicts the distribution of Stockfish ratings for a given opening move across starting positions in which the respective move is available. Vertical lines indicate 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

Figure 4: 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.

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-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. This is a conservative approach to estimating the memory premium. It is conservative because it will falsely classify some out-of-memory moves as in-memory, which will lead us to underestimate the true memory premium. As a robustness check, we estimate the memory premium for several thousand alternative definitions. Virtually all of them yield larger estimates (see Section 4).

Drawing on our workhorse definition, the left panel in Figure 4 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.[11]

[11]Opening moves that are only available in Chess960 are never classified as in-memory. They are excluded

Table 1: Summary Statistics

Variable	Mean	SD	Percentile					N
			5%	25%	50%	75%	95%	
A. Player Level								
ELO Rating in Standard Chess	1,555	344	1,030	1,300	1,530	1,790	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	4.51	3.09	1	2	4	6	11	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.

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 4 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.

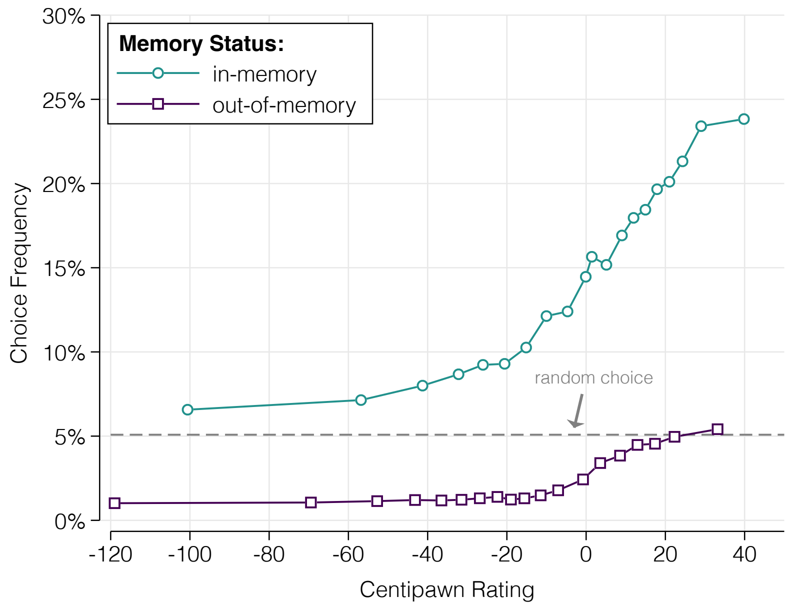
Summary Statistics Descriptive statistics for the most important remaining variables in our data are shown in Table 1. The average (median) player has engaged in more than 1,000 (270) standard chess games prior to her first game of Chess960. Over the course of these games, she has executed 4.5 (4) different 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 Chess960 is, indeed, unfamiliar.

4. The Memory Premium

Figure 5 provides the first piece of evidence that in-memory alternatives are chosen more frequently than equally good out-of-memory ones. The figure depicts binscatter plots relating

from Figure 4. To gauge the popularity of each opening move in standard chess, see Appendix Figure AF.2.

Figure 5: 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.

move quality to choice frequency. It does so for in-memory and out-of-memory moves. Two key patterns emerge. First, conditional on memory status, choice frequencies are monotonically increasing in quality, as assumed in Section 2. 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, suggesting that the ϵ in the statement of Observation 1 is meaningfully large.

To provide more rigorous evidence of the memory premium, we estimate the following linear probability model:

$$(1) \qquad Choose_{p,m,s} = \delta\, In\text{-}Memory_{p,m} + \gamma\, Quality_{m,s} + \mu_m + \epsilon_{p,m,s}.$$

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 the move is and is not in memory, holding the

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	4.53
	(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	0.24	0.24	0.24
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. (1) 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.

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.[12] 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 Section 6, we additionally leverage within-player variation over time. For the remainder of this section, however, we focus on players' first move in their first game of Chess960.

Table 2 presents our main result. 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 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 also more likely to be in memory.

[12] 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 that are less precise. Reassuringly both sets of results are qualitatively equivalent.

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 document that 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.

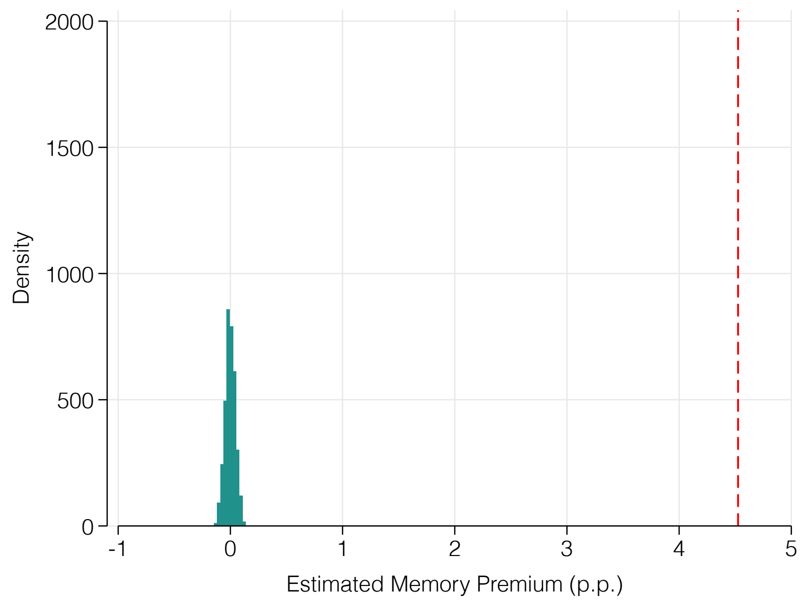
Robustness To assess the robustness of our main result we conduct a placebo test. Specifically, we randomly reassign histories of prior play across the individuals in our data. We then reestimate the regression model in col. (6) of Table 2 based on the newly induced memories. Repeating this process a thousand times, we compare the true point estimate to the distribution of placebo coefficients.

Figure 6 shows the results. As one might expect, the median placebo estimate is approximately zero.[13] 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, indeed, linked.

We proceed to assess the robustness of our main finding to different definitions of memory. To this end, we follow the alternative approach outlined in Section 3. 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.

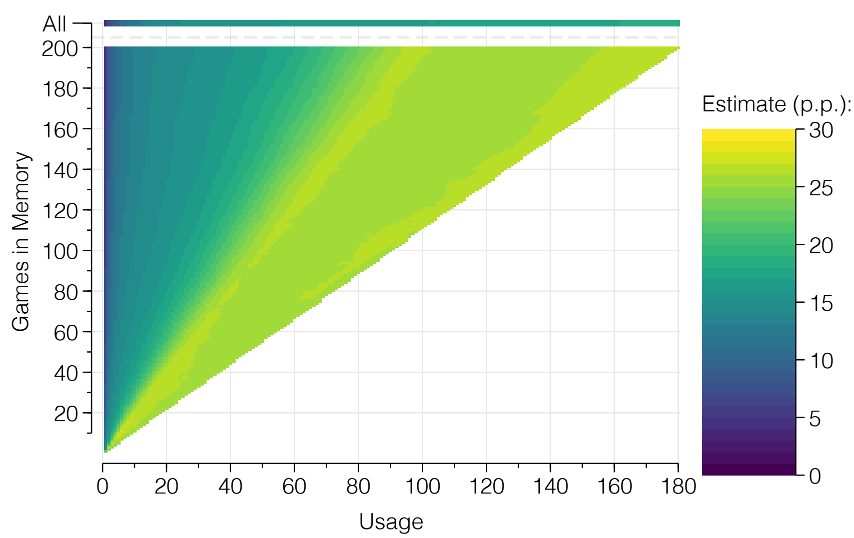
[13]By contrast, the regression model in col. (1) of Table 2 produces placebo estimates that tend to be significantly greater than zero. This observation is consistent with omitted variables bias, which explains why the point estimate decreases upon controlling for centipawn ratings and move fixed effects.

Figure 6: 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: 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 moves to be classified as in-memory.

Table 3: Heterogeneity in the Memory Premium

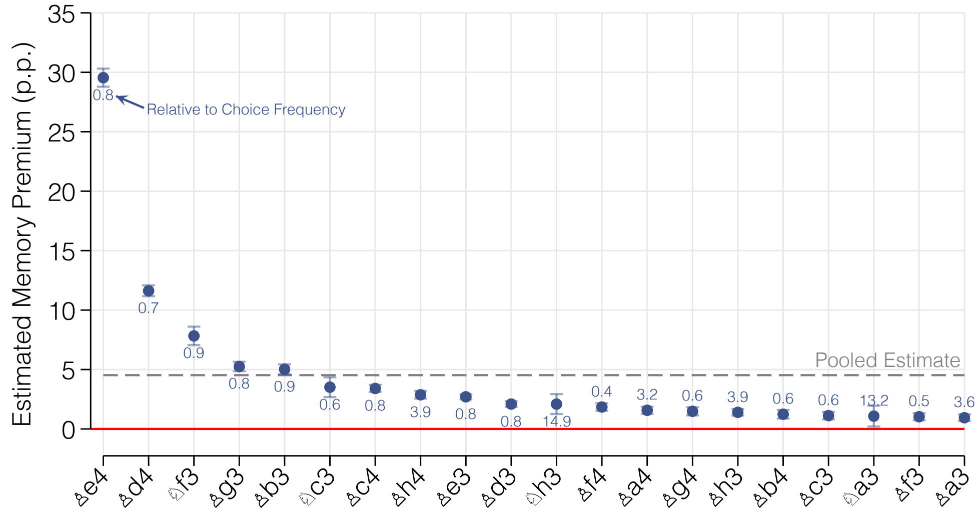
	Probability of Choosing Move					
	(1)	(2)	(3)	(4)	(5)	(6)
In-Memory (÷ 100)	4.66 (0.07)	4.41 (0.07)	3.15 (0.06)	6.49 (0.09)	6.41 (0.10)	3.57 (0.05)
Sample	Strong Players	Weak Players	High Experience Players	Low Experience Players	Close to Optimal Moves	Suboptimal Moves
Fixed Effects:						
Move × Position × Size of Memory	Yes	Yes	Yes	Yes	Yes	Yes
Player × Opponent Strength	Yes	Yes	Yes	Yes	Yes	Yes
Previous Standard Games as White	Yes	Yes	Yes	Yes	Yes	Yes
Mean of LHS Variable	5.08	5.08	5.08	5.08	9.35	3.35
R^2	0.30	0.32	0.31	0.28	0.25	0.23
N	1,456,160	1,442,295	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.

Figure 7 visualizes the resulting point estimates. Anticipating the comparative statics in the next section, we note two regularities. First, holding the length of players' assumed memory fixed, we tend to observe larger coefficients for higher usage thresholds, i.e., when in-memory moves have been used more frequently in the past. Second, for any particular usage threshold, we find that the estimated memory premium decreases if we assume that more distant openings can still be recalled—consistent with memory decay. More importantly, the estimates in Figure 7 range from about 4.5 to 27 p.p. Comparing this range with the coefficients in Table 2 demonstrates that our workhorse definition yields a conservative estimate of the memory premium.

Heterogeneity Returning to our workhorse definition, Table 3 investigates heterogeneity in the memory premium. The first two columns present estimates for players with above- and below-median strength ratings in regular chess. The next two columns differentiate players according to their previous chess experience. Finally, the last two columns distinguish between opening moves that are and are not close to optimal in a given starting position (i.e., within 20 centipawns of the best available move). 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. Interestingly, we also find heterogeneity by relative move quality. The estimated memory premium is significantly higher for moves that are close to optimal. Yet the point

Figure 8: 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.

estimate in col. (6) of Table 3 indicates that, even for moves that are suboptimal, the memory premium is economically large and statistically different from zero.

We also investigate heterogeneity across individual moves. Figure 8 plots the estimated memory premium for each of the twenty opening moves in standard chess. Although we observe much larger premia for moves that are frequently used in regular chess, such as ♙e4, ♙d4 and ♘f3, it is important to note that all estimates in Figure 8 are statistically significant and economically sizeable—especially compared to the mean choice frequencies of the respective moves. Appendix Table AT.2 further distinguishes between starting positions in which a particular move is and is not close to optimal. In the former case, sixteen out of twenty point estimates are positive and statistically significant.[14] In the latter case, all twenty estimates are positive and statistically different from zero. The evidence, therefore, indicates that the memory premium applies to all moves, not just ones that are popular or good.

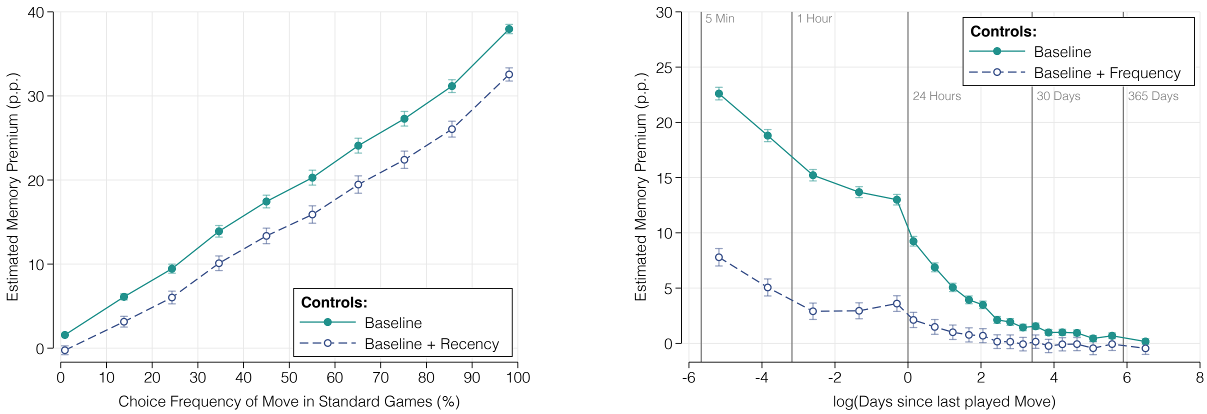
5. Comparative Statics

In order to establish a more-direct connection between the memory premium and the ease with which experiences can be retrieved from memory, we proceed to examine Properties 1–4. That is, we ask whether the memory premium exhibits the same key comparative statics that cognitive scientists have previously documented for the probability of recall.

[14]Of the remaining four estimates, two are positive and two are negative. All four are statistically insignificant.

Figure 9: Reinforcement and Decay

(a) Memory Premium, by Frequency of Prior Use (b) Memory Premium, by Recency of Prior Use



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.

5.1. *Repetition and Decay*

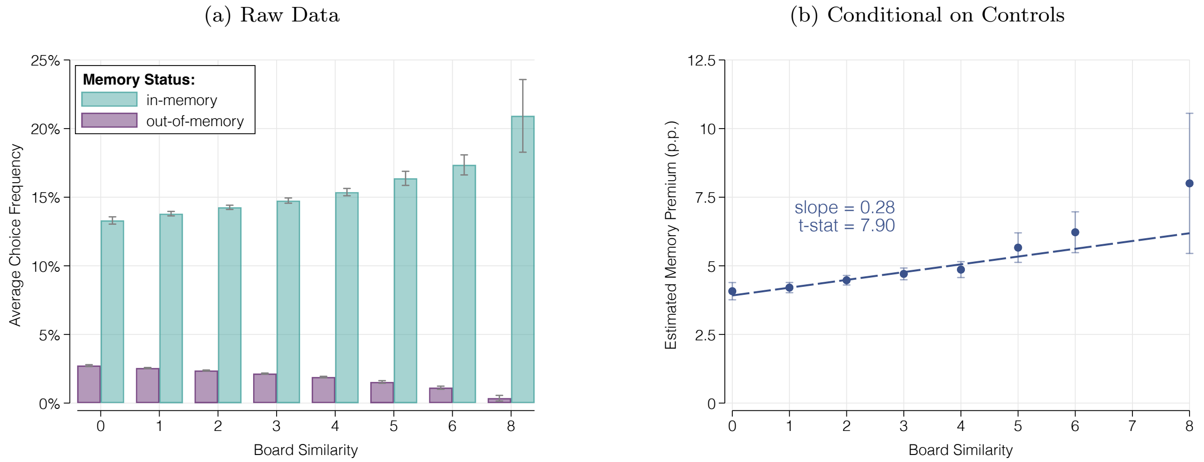
A fundamental property of memory is that it is reinforced by repetition, and that it decays over time. Figure 7 above already provides some evidence that similar forces operate on the memory premium. After all, alternative definitions of memory that impose higher usage thresholds and limit recall to shorter histories of play yield larger point estimates. To investigate the comparative statics of the memory premium more carefully, we apply our workhorse definition and extend our most saturated regression model in Table 2 to allow for heterogeneity in δ, depending on how often and how recently a given in-memory move had previously been used.

Figure 9 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. Property 2 is, therefore, supported by the data.

The right panel of Figure 9 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 Property 1, moves that had been played more recently in standard chess exhibit significantly higher memory premia.[15]

[15] Appendix Figure AF.4 replicates the analysis in the right panel of Figure 9 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, one 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

Figure 10: Memory Premium, by Board Similarity



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.

In Appendix Figure AF.5, we differentiate between moves with above- and below-median payoffs, i.e., moves that a given player used more or less successfully in the past (as measured by the outcomes of the relevant games). Conditional on a move's frequency of use, the memory premium increases in prior success—though perhaps not as much as one might have expected. Although there is a very strong unconditional relationship between choice frequencies in Chess960 and moves' prior payoffs, this relationship appears to be mediated by the frequency of use. Taken together, the evidence implies that the memory premium is strengthened by repetition and that it decays over time.

5.2. *Associativity*

Memory is associative in the sense that past experiences are more likely to be recalled when they are related to current environmental cues. To investigate this property, 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. 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.

This is, indeed, what we find. The left panel of Figure 10 partitions the raw data according to board similarity, and shows raw choice frequencies for in- and out-of-memory moves.[16] The

of our main result in Table 2 based on donut definitions of memory, see Appendix Tables AT.3–AT.6.

[16]Given that the standard chess board is one of the possible starting positions, our measure of similarity

right panel shows, for each level of 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 associative, consistent with Property 3.

5.3. *Interference*

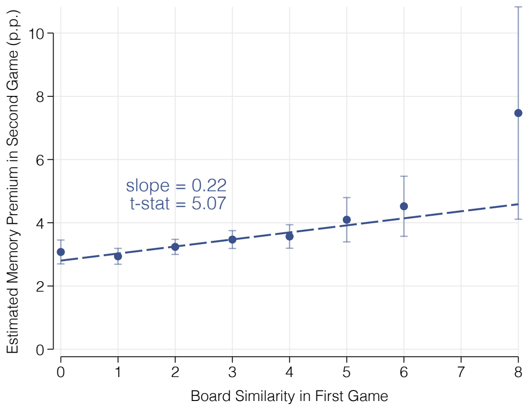
Experiences compete for recall to the point where new experiences interfere with the retrieval of older ones. To test for interference effects in the memory premium, we take a three-pronged approach. First, we test Property 4 by asking whether the memory premium in a player's *second* game of Chess960 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 are more salient when the starting position was very 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

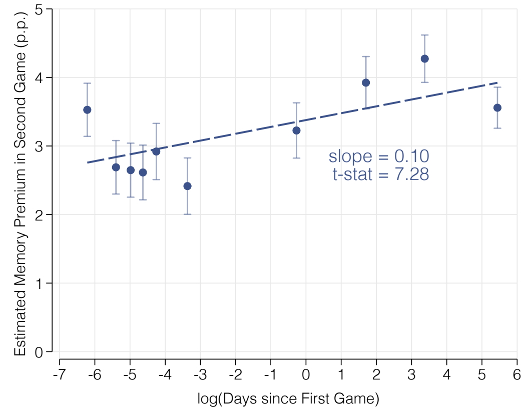
ranges from zero to eight. There are no starting positions in which seven pieces are placed on their usual squares.

Figure 11: Evidence of Interference

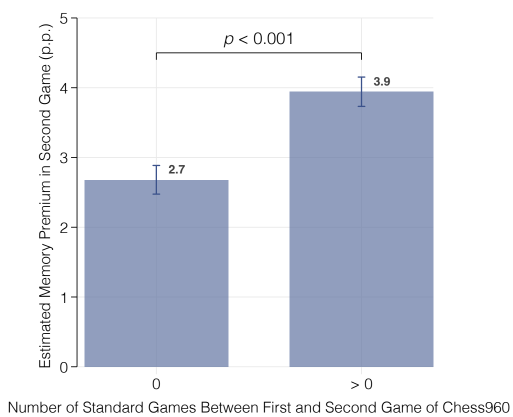
(a) Interference via Similarity



(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.

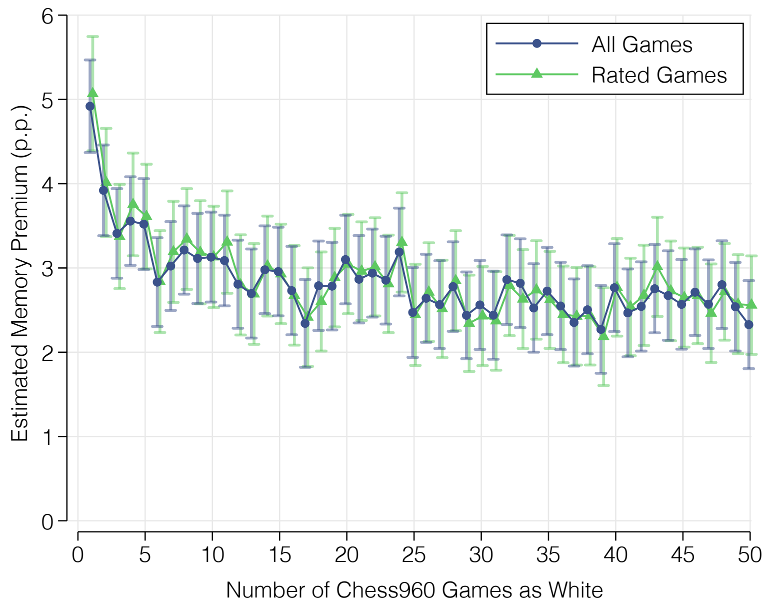
in at least one game of standard chess between their first two games of Chess960.[17] The memory premium is significantly smaller when they have not. The evidence in panels (a)–(c) is, therefore, consistent with the idea that the memory premium is moderated by interference in recall.

6. Dynamics and Causality

We next examine how the memory premium evolves as individuals become more familiar with the new choice environment. We also provide evidence to suggest that the connection between memory and choice behavior is plausibly causal.

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

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.

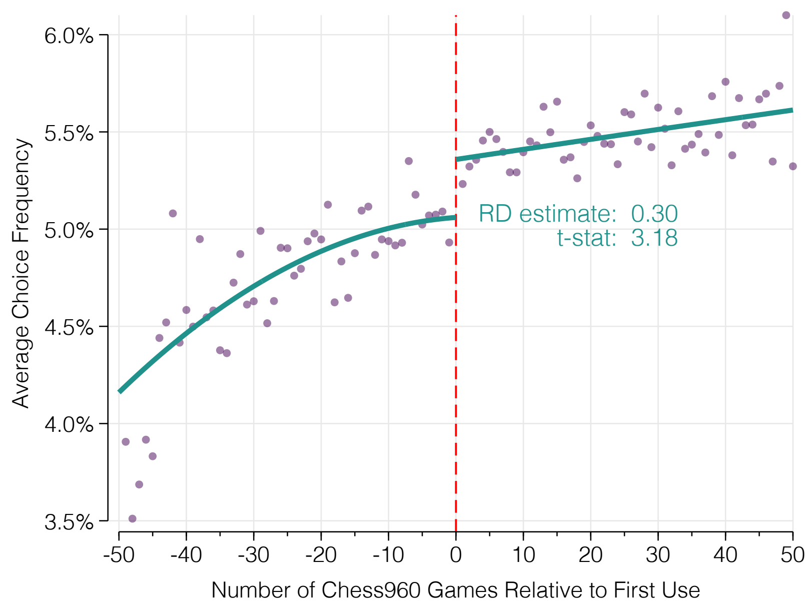
6.1. *Dynamics of the Memory Premium*

In order to investigate how the memory premium evolves as decision-makers gain experience, we change the scope of our analysis to 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. We then estimate the most saturated regression model in Table 2 separately for each of their first fifty games. 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 shows how the memory premium evolves over time. Irrespective of whether we consider all games or only rated ones, two patterns stand out. First, despite limited 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 vanish. 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.[18]

[18]Appendix Figure AF.6 replicates the analysis in Figure 12 based on a definition of memory that is based

Figure 13: 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.

6.2. *Causality*

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.[19] Does the frequency with which the respective move is chosen in Chess960 increase afterwards?

Drawing on the 41,990 instances in which an opening move enters a player's memory after her first game of Chess960, Figure 13 plots the average choice frequency of these moves before and after they were first used in a regular chess game. 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

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. The observation that the memory premium is higher for moves that have previously been used in Chess960 is consistent with associative recall.

[19]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 above).

move is used for the first time in a standard game, the frequency with which it is chosen in Chess960 jumps by approximately 0.3 p.p., or about 6%.[20]

The upper panel of Table 4 shows that the discontinuity in Figure 13 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 4 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. This is not surprising. On average, the moves that identify the estimates in Table 4 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. 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 0.43 p.p. (with a standard error of 0.17). There is, therefore, no contradiction between the causal estimates in Table 4 and our estimates of the average memory premium in Table 2.

[20]Appendix Figure AF.7 shows the density of observations around the threshold. There is no discontinuity.

Table 4: 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	5.08	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.

7. Discussion

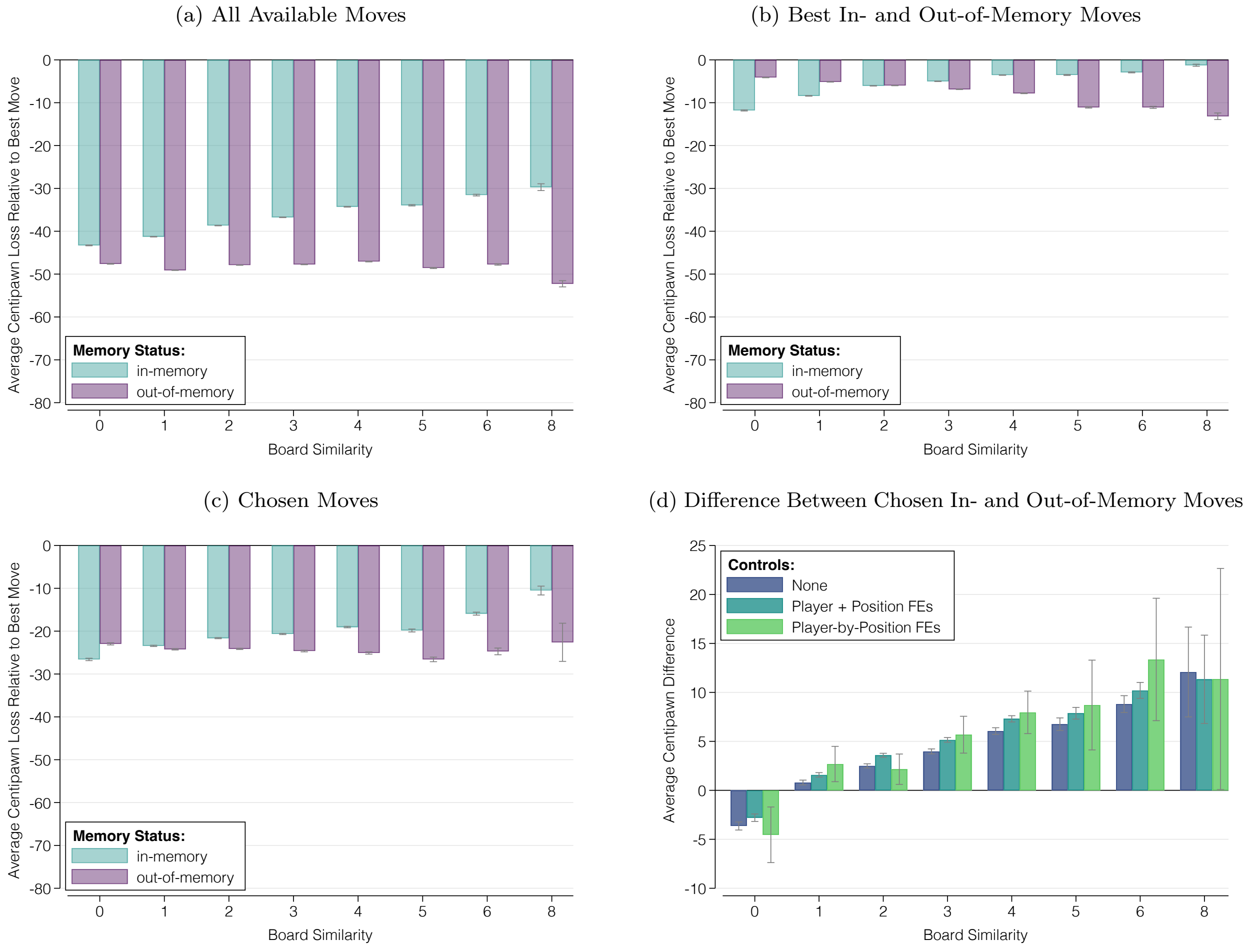
Memory is a key ingredient in human decision-making, allowing us to use past experiences to guide future behavior. This paper documents that, in unfamiliar environments, decision-makers are more likely to choose alternatives that are retrievable from memory than those that are not, even when the latter are objectively better. We dub this phenomenon the "memory premium," and provide evidence that it is associative, subject to interference and repetition effects, and that it decays over time. Our results imply that the ease with which past experiences come to mind plays an important role in shaping choice behavior.

Taking a step back, we can ask whether it makes sense to rely on memory as a guide for decision-making in unfamiliar environments. A key benefit of memory-based heuristics is that they economize on cognitive costs. Suppose that there are N unfamiliar environments and that the cost of analyzing every new environment to find an environment-specific optimal strategy is c. The cost of computing an optimal mapping from environments to strategies is thus cN. An extreme way to economize on cognitive costs would be to use a constant mapping instead, i.e., a strategy that does not change with the environment. Relying entirely on memory rather than careful analysis is an example of such a mapping. Of course, there is a wide range of intermediate cases. For example, a DM may utilize memory when the decision problem is sufficiently similar to the ones that she encountered in the past but engage in deeper analysis when it is not. In this sense, drawing on memory to make decisions in unfamiliar environments can be viewed as a cognitive shortcut.

But does such a shortcut yield good outcomes? While a general answer is beyond the scope of this paper, we note that the performance of any memory-based heuristic will, at least in part, depend on the degree of similarity between familiar and unfamiliar environments. Focusing on the case of Chess960, the top two panels of Figure 14 examine the average quality of opening moves, by memory status. The top-left panel shows the average centipawn rating of in- and out-of-memory moves relative to the best available alternative in a given starting position. In-memory alternatives are, on average, better than out-of-memory ones, with the gap widening as the starting position becomes more similar to the standard board. As a result, a naïve decision maker could, in expectation, improve upon random choice by randomly choosing an in-memory alternative. The magnitude of the improvement would be especially large in starting positions that closely resemble the standard one.

The top-right panel of Figure 14 compares the *best* available in- and out-of-memory moves in a given starting position. For high and intermediate levels of board similarity, the players in our data would, on average, sacrifice very little if they maximized only over the consideration set of in-memory alternatives. Only in the most dissimilar board positions would such a choice procedure result in a large expected loss relative to the first best.

Figure 14: Quality of Moves, by Memory Status



Notes: Panel (a) compares the average Stockfish rating of all available in- and out-of-memory moves relative to the best move in a particular starting position. Negative values thus correspond to losses relative to the optimum. Panel (b) shows the average Stockfish rating of players' highest-rated in- and out-of-memory moves in a particular starting position relative to the best opening move in that position. Panel (c) shows the average Stockfish rating of the moves that players actually choose, by memory status and relative to the best available alternative. Panel (d) presents estimates of the difference in Stockfish ratings between chosen in- and out-of-memory moves, controlling for different sets of fixed effects. All panels differentiate according to the similarity of the starting position in Chess960 with that in standard chess. The sample always consists of moves in players' first fifty Chess960 games as White. Error bars correspond to 95%-confidence intervals, accounting for clustering by player.

The bottom two panels focus on the opening moves that were actually chosen. The panel on the bottom-left presents raw data, while the one on the right shows estimates of the difference in quality between in- and out-of-memory moves after controlling for player and starting position fixed effects. Two findings emerge. First, conditional on memory status, chosen moves are far better than average but significantly worse than the best available alternative, suggesting that choices in this setting are deliberate yet imperfect. Second, unless the starting position is highly dissimilar from the standard one, players do better when they choose in- rather than out-of-memory moves. This pattern is not only borne out in the raw data, but it also holds conditional on player-by-starting-position fixed effects, i.e., when we compare choice situations in which the same player encountered the same starting position more than once. Only in the most dissimilar starting positions do we observe that players do better when they choose moves that are out-of-memory.

To help interpret the previous observation, note that although in-memory moves are, on average, better than out-of-memory ones for all levels of board similarity (see panel (a)), in dissimilar starting positions the best out-of-memory alternatives tend to be better than the best in-memory alternatives (see panel (b)). As a consequence, players can realize nontrivial improvements by carefully considering *all* available options.

In sum, the evidence above is consistent with the idea that relying on memory is boundedly rational. It can reduce cognitive costs while yielding favorable outcomes in environments that are not too dissimilar.

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.

Baddeley, A., M. Eysenck, and M. Anderson (2009). *Memory* (3rd ed.). Hove, UK: Psychology Press.

Bordalo, P., G. Burro, K. Coffman, N. Gennaioli, and A. Shleifer (2024). Imagining the Future: Memory, Simulation, and Beliefs. *Review of Economic Studies*, forthcoming.

Bordalo, P., J. Conlon, N. Gennaioli, S. Kwon, and A. Shleifer (2022). Memory and Probability. *Quarterly Journal of Economics 138*(1), 265–311.

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.

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.

Charles, C. (2024). Memory Moves Markets. mimeographed, London School of Economics.

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.

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

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.

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 (2024). Investor Memory and Biased Beliefs: Evidence from the Field. NBER Working Paper No. 33226.

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

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.

Melton, A. (1963). Implications of Short-Term Memory for a General Theory of Memory. *Journal of Verbal Learning and Verbal Behavior* (1), 1–21.

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.

Salant, Y. and J. Spenkuch (2024). Complexity and Satisficing: Theory with Evidence from Chess. *Review of Economic Studies*, forthcoming.

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

Tversky, A. and D. Kahneman (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology 5*(2), 207–232.

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

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: Proof of Observation 1

By continuity, it suffices to show that $C\left(S,a \mid M^{+a}\right) > C\left(S,b \mid M^{+a}\right)$ when $U(b,E) = U(a,E)$. Given Definition 1, we have that

$$C\left(S,a \mid M^{+a}\right) - C\left(S,a \mid M\right) = \Delta(S,a \mid M^{+a}) > \Delta(S,b \mid M^{+a}) = C\left(S,b \mid M^{+a}\right) - C\left(S,b \mid M\right).$$

The observation follows because monotonicity and $U(b,E) = U(a,E)$ imply that $C\left(S,a \mid M\right) = C\left(S,b \mid M\right)$.

Appendix B: Data Appendix

The 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.[1] 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 147,357 players in their first game of Chess960 on the platform. In order to be included in this sample, users must have played at least 20 standard chess games 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 decisions are subject to real stakes.

In the second part of our analysis, we study the opening 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 opening 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.

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

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 replicates Figure 7 in the text controlling for *which* other moves players hold in memory. The estimates in this figure are based on a regression model that replaces the move-by-starting-position-by-size-of-memory fixed effect in the specification in col. (6) of Table 2 with a fixed effect for the combination of the move, starting position, and the set of other in-memory moves.

Appendix Table AT.2 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.

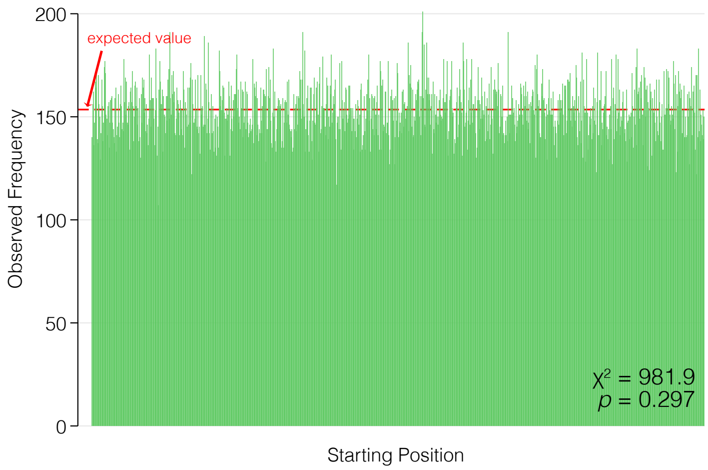
Appendix Figure AF.4 and Appendix Tables AT.3–AT.6 respectively replicate the analyses in Figure 9(b) and Table 2 in the text based on "donut definitions" of memory. According to these definitions, a move is said to be in-memory if, and only if, it had been played at least once sufficiently long prior to a player's first game of Chess960.

Appendix Figure AF.5 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 shows the density of observations around the threshold in the regression discontinuity design in Figure 13 and Table 4 in the text.

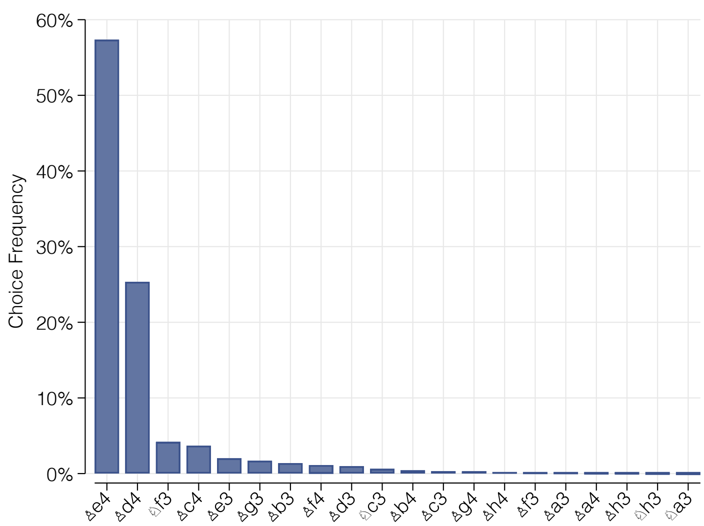
Appendix Figures

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



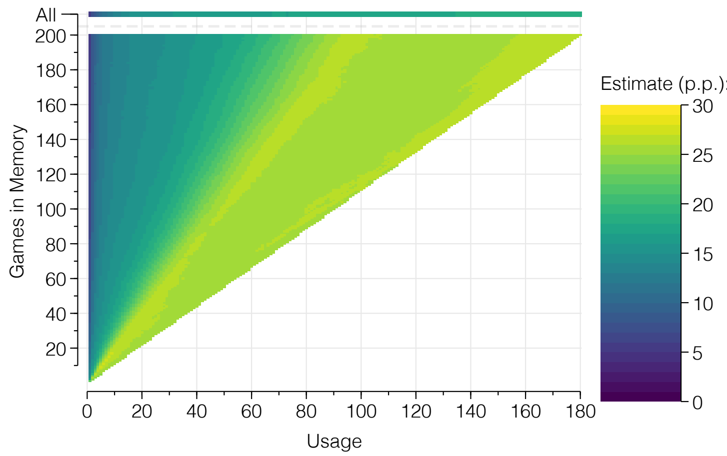
Notes: Figure shows the frequency of each starting position in our main sample of players' first game of Chess960. The reported χ^2 test refers to the null hypothesis of no differences across starting positions.

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, Controlling for the Set of Other In-Memory Moves

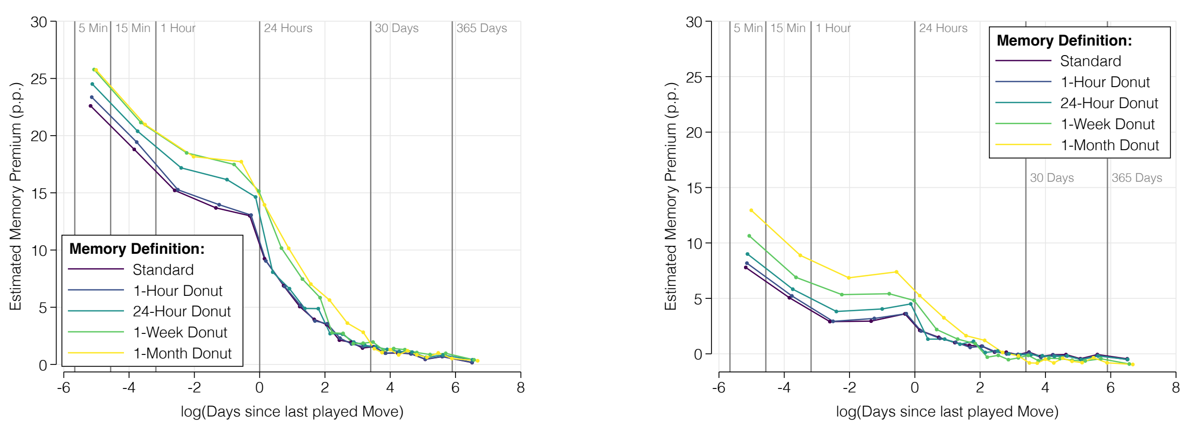


Notes: See Figure 7 in the main text. The only difference between this figure and that in the text is that the results above do not only control for the number of other in-memory moves but for *which* other moves are in-memory. More specifically, the estimates above are based on a regression model that replaces the move-by-starting-position-by-size-of-memory fixed effect in the specification in col. (6) of Table 2 with a move-by-starting-position-by-set-of-other-in-memory-moves fixed effect.

Appendix Figure AF.4: Decay in the Memory Premium, Based on Donut Definitions of Memory

(a) Baseline Set of Controls

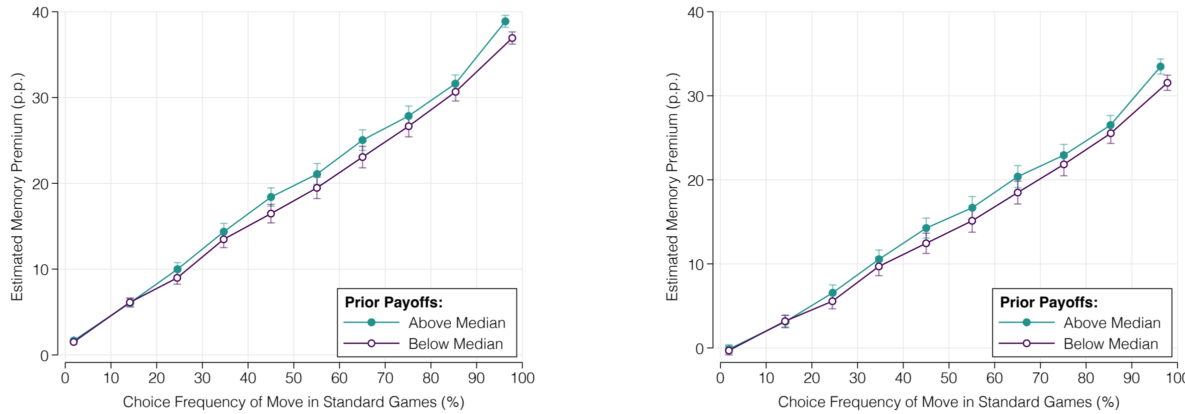
(b) Baseline and Frequency Controls



Notes: Figure replicates the analysis in Figure 9(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.

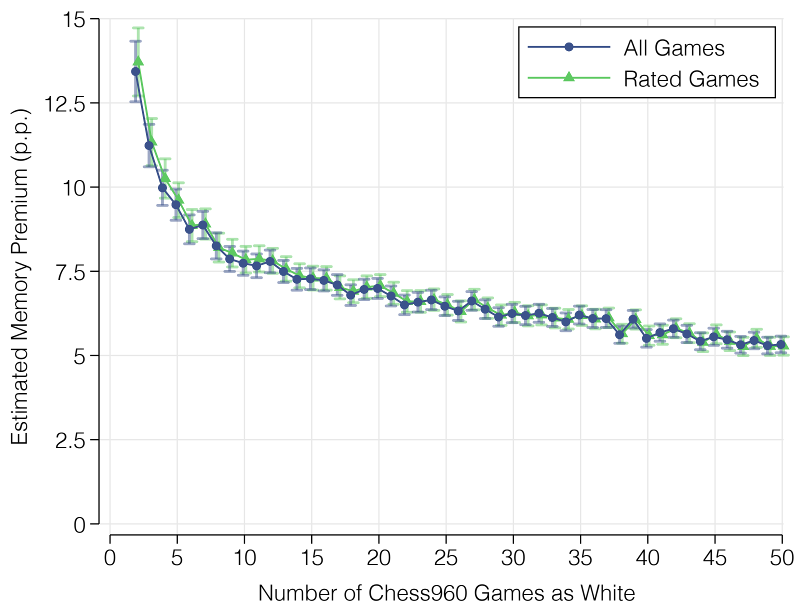
Appendix Figure AF.5: Frequency Effects, By Prior Success

(a) Baseline Set of Controls (b) Baseline and Frequency Controls



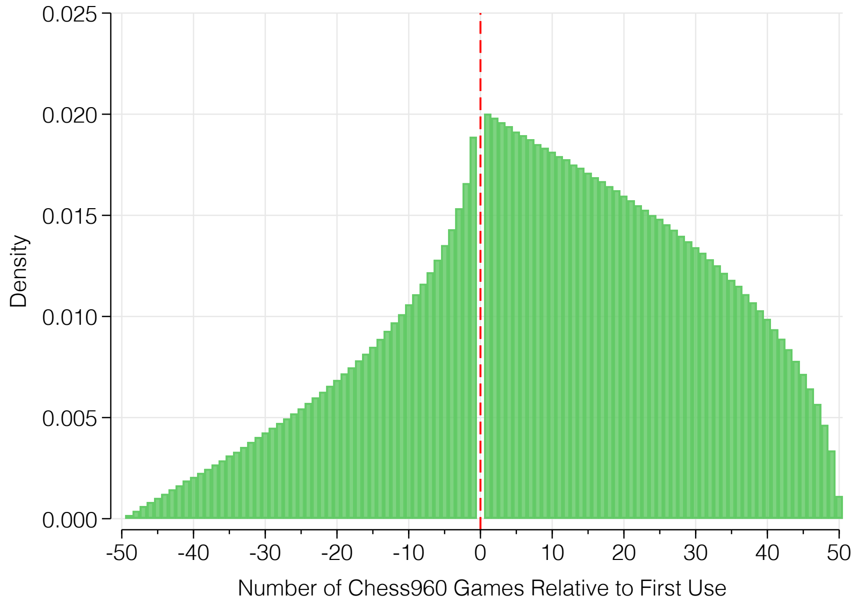
Notes: Figure replicates the analysis in Figure 9(a) in the text, differentiating between moves with above- and below-median player-specific average payoffs in prior use in standard chess games on Lichess. Payoffs are calculated by averaging over the outcomes of all games in which a player previously 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.6: 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.7: Density of Observations around RD Threshold



Notes: Figure shows the density of observations around the threshold in the regression discontinuity design in Section 6.

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	4.96	5.03
	(0.04)	(0.04)	(0.04)	(0.13)	(0.13)	(0.13)
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	0.67	0.67	0.67
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, 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	0.81	1.03
		(0.15)	(0.64)	(0.16)
♙a4	0.5	1.64	2.42	1.39
		(0.16)	(0.40)	(0.18)
♘a3	0.1	1.08	–0.12	1.07
		(0.43)	(1.22)	(0.45)
♘c3	5.7	3.66	3.85	3.45
		(0.46)	(0.71)	(0.62)
♙b3	5.5	5.20	7.81	3.87
		(0.23)	(0.52)	(0.23)
♙b4	2.0	1.52	2.12	1.34
		(0.20)	(0.51)	(0.21)
♙c3	1.9	1.12	1.80	0.91
		(0.17)	(0.48)	(0.17)
♙c4	4.5	3.99	4.66	3.36
		(0.17)	(0.29)	(0.21)
♙d3	2.5	2.18	1.75	2.28
		(0.14)	(0.32)	(0.15)
♙d4	17.4	15.62	16.78	14.15
		(0.28)	(0.38)	(0.42)
♙e3	3.3	2.80	2.82	2.82
		(0.12)	(0.26)	(0.14)
♙e4	35.1	34.70	33.77	36.39
		(0.41)	(0.58)	(0.57)
♙f3	1.9	0.94	0.77	0.97
		(0.16)	(0.50)	(0.17)
♙f4	4.3	2.24	1.99	2.42
		(0.18)	(0.29)	(0.22)
♘f3	8.8	8.48	9.05	7.95
		(0.44)	(0.67)	(0.60)
♘h3	0.1	1.90	–1.25	1.96
		(0.42)	(0.66)	(0.45)
♙g3	6.3	5.50	7.40	4.62
		(0.22)	(0.50)	(0.23)
♙g4	2.4	1.58	1.83	1.50
		(0.19)	(0.52)	(0.19)
♙h3	0.4	1.39	2.10	1.26
		(0.15)	(0.88)	(0.16)
♙h4	0.7	2.96	3.31	2.81
		(0.17)	(0.51)	(0.19)

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.3: 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	4.52
	(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	0.24	0.24	0.24
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.4: 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	4.75
	(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	0.24	0.24	0.24
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.5: 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	5.17	5.17
	(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	0.24	0.24	0.24
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.6: 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	5.17	5.19
	(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	0.23	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 month prior to a player's first game of Chess960.