PHBS WORKING PAPER SERIES

Complexity Beyond Incentives:
The Critical Role of Reporting Language

Rustamdjan Hakimov
University of Lausanne & WZB Berlin Social Science Center

Manshu Khanna
Peking University

January 2026

Working Paper 20260101

Abstract

Many assignment systems require applicants to rank multi-attribute bundles (e.g., programs combining institution, major, and tuition). We study whether this reporting task is inherently difficult and how reporting interfaces affect accuracy and welfare. In laboratory experiments, we induce preferences over programs via utility over attributes, generating lexicographic, separable, or complementary preferences. We compare three reporting interfaces for the direct serial dictatorship mechanism: (i) a full ranking over programs; (ii) a lexicographic-nesting interface; and (iii) a weighted-attributes interface, the latter two eliciting rankings over attributes rather than programs. We also study the sequential serial dictatorship mechanism that is obviously strategy-proof and simplifies reporting by asking for a single choice at each step. Finally, we run a baseline that elicits a full ranking over programs but rewards pure accuracy rather than allocation outcomes. Four main findings emerge. First, substantial misreporting occurs even in the pure-accuracy baseline and increases with preference complexity. Second, serial dictatorship induces additional mistakes consistent with misperceived incentives. Third, simplified interfaces for the direct serial dictatorship fail to improve—and sometimes reduce—accuracy, even when they match the preference structure. Fourth, sequential choice achieves the highest accuracy while improving efficiency and reducing justified envy. These findings caution against restricted reporting languages and favor sequential choice when ranking burdens are salient.

Keywords: matching market, random priority, market design, multiple attributes
JEL Classification: C92, D47

Peking University HSBC Business School
University Town, Nanshan District
Shenzhen 518055, China

Complexity Beyond Incentives: The Critical Role of Reporting Language*

Rustamdjan Hakimov† Manshu Khanna‡

November 28, 2025

Abstract

Many assignment systems require applicants to rank multi-attribute bundles (e.g., programs combining institution, major, and tuition). We study whether this reporting task is inherently difficult and how reporting interfaces affect accuracy and welfare. In laboratory experiments, we induce preferences over programs via utility over attributes, generating lexicographic, separable, or complementary preferences. We compare three reporting interfaces for the direct serial dictatorship mechanism: (i) a full ranking over programs; (ii) a lexicographic-nesting interface; and (iii) a weighted-attributes interface, the latter two eliciting rankings over attributes rather than programs. We also study the sequential serial dictatorship mechanism that is obviously strategy-proof and simplifies reporting by asking for a single choice at each step. Finally, we run a baseline that elicits a full ranking over programs but rewards pure accuracy rather than allocation outcomes. Four main findings emerge. First, substantial misreporting occurs even in the pure-accuracy baseline and increases with preference complexity. Second, serial dictatorship induces additional mistakes consistent with misperceived incentives. Third, simplified interfaces for the direct serial dictatorship fail to improve—and sometimes reduce—accuracy, even when they match the preference structure. Fourth, sequential choice achieves the highest accuracy while improving efficiency and reducing justified envy. These findings caution against restricted reporting languages and favor sequential choice when ranking burdens are salient.

Keywords: Matching Market, Random Priority, Market Design, Multiple Attributes.

JEL: C92, D47.

*We thank Inacio Bó, Dorothea Kübler, Bozhang Xia and Xi Jin for valuable comments and suggestions. We are grateful to Haomin He, Shapeng Jiang, and Wang Jingyi for excellent research assistance. We acknowledges financial support from the National Natural Science Foundation of China (Grant #W2433170) (Khanna) and Swiss National Science Foundation (Project #10003543) (Hakimov). The experiment was pre-registered on the AEA registry, number AEARCTR-0013135.

†University of Lausanne & WZB Berlin Social Science Center, Internef 536, Quartier de Chamberonne, CH-1015, Lausanne, Switzerland. Email: rustamdjan.hakimov@unil.ch

‡Peking University HSBC Business School, Shenzhen 518055, China. Email: manshu@phbs.pku.edu.cn

1 Introduction

Matching mechanisms allocate scarce resources across many domains—medical residencies, school seats, public housing, organ donations, and limited appointment slots (Roth, 2002; Abdulkadiroglu and Sönmez, 2003; Abdulkadiroğlu and Sönmez, 1999; Budish and Cantillon, 2012; Roth *et al.*, 2004; Hakimov *et al.*, 2021). A fundamental requirement of these mechanisms is that participants submit preference rankings over fully specified options.[1] When options are defined by multiple attributes, i.e., levels of various attributes uniquely define an object (e.g., university $\times$ major $\times$ tuition), this seemingly simple task becomes cognitively demanding, taxing attention, memory, and computation (Rees-Jones and Skowronek, 2018; Milgrom, 2009, 2011). The burden is evident in practice: Chinese applicants rank hundreds of institution–major bundles (Hu *et al.*, 2025); U.S. Military Academy cadets evaluate branch–service combinations (Greenberg *et al.*, 2024; Sönmez, 2013); and bidders in combinatorial auctions assess packages with complementarities (Parkes, 2005; Kwasnica *et al.*, 2005; Sandholm and Boutilier, 2005). The question of how to elicit these complex preferences accurately has received surprisingly little attention.

We address two central questions. First, is ranking multi-attribute options inherently difficult, and does this difficulty vary with preference structure? Second, do simplified reporting interfaces—widely adopted in practice—actually help participants communicate their preferences more accurately?

These questions matter because many real-world systems restrict how preferences can be expressed, ostensibly to reduce cognitive burden. China's college admissions system, for example, uses a structured rank-order format that requires ranking universities first, then within each university, ranking the majors offered by that institution. This nesting of majors

[1] Canonical mechanisms—Deferred Acceptance (Gale and Shapley, 1962; Balinski and Sönmez, 1999), Random Serial Dictatorship (Abdulkadiroğlu and Sönmez, 1998; Pycia and Troyan, 2024), Top Trading Cycles, and the Cumulative Offer Mechanism (Hatfield and Milgrom, 2005; Hatfield *et al.*, 2020)—require agents to submit rankings over fully specified options. Preferences in standard models are typically assumed to be known, a few studies relax this assumption by allowing costly preference learning (Chen and He, 2021, 2022; Hakimov *et al.*, 2023), while others allow for indecisiveness (Caspari and Khanna, 2025).

under institutions effectively enforces lexicographic preferences (Hu *et al.*, 2025).[2] At the U.S. Military Academy, Sönmez and Switzer (2013) proposed replacing the USMA-2006 mechanism with a more expressive alternative, but the Academy retained USMA-2006 to avoid excessive reporting burden (Greenberg *et al.*, 2024).[3] Other simplification strategies include bidding with artificial budgets (Budish, 2011; Budish and Kessler, 2022), constrained choice sets (Haeringer and Klijn, 2009; Calsamiglia *et al.*, 2010a; Huang and Zhang, 2025), sequential menu choice (Bó and Hakimov, 2024; Mackenzie and Zhou, 2022), and AI-guided elicitation (Soumalias *et al.*, 2025). While these approaches reduce cognitive load, they may restrict expressiveness and lower efficiency.

We study these questions in laboratory experiments with student subjects. Our experimental design mimics a stylized college admissions setting where 27 participants are assigned to 27 program seats. Each seat is defined by three attributes—university, field of study, and tuition.[4] Each attribute has three levels, yielding $3 \times 3 \times 3 = 27$ distinct programs (bundles). Students know the distribution from which exam scores are drawn, and their exam scores, and allocations are determined by Serial Dictatorship in most treatments. We independently vary (i) the complexity of preferences over attributes and (ii) the interface used to report preferences.

Our design has three innovations. First, unlike most matching experiments, we do not assign ordinal preferences directly.[5] Instead, each participant is given a utility function that maps attribute profiles into utility scores. These scores define ordinal rankings over

[2] As of January 2025, 23 of 31 provinces maintain this format. Appendix B shows sample forms from Fujian and Shanghai.

[3] Greenberg *et al.* (2024) write "While this (Sönmez and Switzer, 2013) proposal had desirable theoretical properties, it required a more complex strategy space in which cadets have to rank branches and contractual terms (also referred to as prices) jointly. Under the USMA-2006 mechanism, cadets only rank branches and separately indicate their willingness to BRADSO for any branch. The Army considered the existing strategy space manageable compared to a more complex alternative and kept the USMA-2006 mechanism in the intervening years."

[4] Tuition is a relevant attribute in systems where the same seat may be offered with or without tuition, such as in Israel and Hungary (Hassidim *et al.*, 2021; Shorrer and Sóvágó, 2023).

[5] An exception is Budish and Kessler (2022), where participants express home grown preferences, and the preference information is elicited through a series of binary choices. Also, Kloosterman and Troyan (2023) who use real objects, while Guillen and Hakimov (2018) run a field experiment with real preferences of students over topics.

programs, and payoffs depend on the rank of the assigned program. This allows us to manipulate complexity of preferences over attributes exogenously[6]—ranging from lexicographic (LEX), to additive separable (SEP), to non-separable with complementarities (COMP).[7] This treatment dimension captures cognitive challenges in forming and articulating preferences.

Second, to separate cognitive barriers to truthful reporting from incentive-driven misreporting in a strategy-proof mechanism, we include a baseline treatment (ACCURACY). Here, participants are rewarded for reporting true rankings of 27 bundles, but no allocation takes place. This provides a measure of pure reporting frictions, abstracting from strategic incentives.

Third, we vary reporting interfaces, i.e., how preferences are communicated to the Serial Dictatorship mechanism. In the direct interface (SD-DIRECT), participants submit a full ranking over all 27 bundles. The lexicographic interface (SD-LEX) asks participants to rank each attribute separately[8] and constructs full bundle rankings for Serial Dictatorship assuming lexicographic preferences (university $\succ$ field $\succ$ tuition). This mirrors systems like China's, where the lexicographic structure over attributes is exogenously imposed: it fits lexicographic preferences exactly but distorts rankings when preferences are more complex. A third interface (SD-WEIGHT) also asks participants to rank attributes separately, but allows weights across attributes.[9] Expressiveness differs across interfaces.[10]

[6]It also reduces experimenter demand effect, since participants cannot simply copy pre-filled preference tables to be truthful.

[7]In LEX, university is the most important attribute, dominating field of study, which in turn dominates tuition; that is, any program in the top-ranked university is preferred to any program in the second-ranked university, and so on. In SEP, no attribute dominates: each attribute enters with a weight, and the program score increases linearly in the values of the three attributes. In COMP, in addition to attribute weights, there is a complementarity term between field of study and tuition, so the relationship between program scores and attributes is no longer linear.

[8]Under SD-LEX, participants rank the three universities, the three fields, and the three tuition levels separately, instead of submitting a single 27-item ranking required in SD-DIRECT.

[9]Participants rank three rankings of three items as in SD-LEX, but additionally specify a weight for each of the attributes.

[10]In this context, *expressiveness* refers to the ability of a reporting interface to accurately represent all possible preference orderings over the available options. A *fully expressive* interface allows participants to communicate any ranking of the 27 programs. An interface has limited expressiveness when certain preference orderings cannot be accurately represented through that format.

Under LEX preferences, all three interfaces (SD-LEX, SD-WEIGHT, and SD-DIRECT) are fully expressive; under COMP and SEP, neither SD-LEX nor SD-WEIGHT is fully expressive.[11] Thus, while SD-LEX and SD-WEIGHT reduce the burden of submitting a full 27-item ranking, they necessarily introduce misspecification under more complex preference structures, where attribute trade-offs cannot be represented by lexicographic or weighted-attribute aggregation.[12]

Additionally, we study a mechanism that changes both interface and strategy space relative to SD-DIRECT: sequential serial dictatorship (SD-CHOICE). Here, each participant chooses her most preferred available option at the stage of allocation when she is on top of priority. SD-CHOICE requires no full ranking and is *obviously strategy-proof* (Li, 2017; Pycia and Troyan, 2023). It thus serves dual roles—as a simpler interface and as a mechanism with stronger incentive properties. Comparing SD-CHOICE to SD-DIRECT isolates gains from obvious strategy-proofness, while comparing SD-CHOICE to ACCURACY tests whether it eliminates all strategic misreporting and whether its simplified interface can even outperform the benchmark of pure accuracy.

We conducted the experiment with 810 university students, randomly assigned across treatments. Preference complexity was varied within subjects, while reporting interfaces and mechanisms were varied between subjects. Each session included 12 repeated rounds covering the different complexities in random order. Depending on the treatment, participants' final payoffs were determined either by the outcomes of the resulting allocations of SD or by the accuracy of their reported preferences (ACCURACY).

Our results yield five main findings. First, in ACCURACY, with no strategic incentives,

[11] Initially, we hypothesized that SD-WEIGHT would be fully expressive under SEP preferences, since SEP involves additive utilities across attributes and SD-WEIGHT allows participants to assign importance weights to different attributes. However, SD-WEIGHT cannot represent all possible SEP preferences because the attribute values in our setting are not uniformly spaced. The SD-WEIGHT interface implicitly assumes uniform spacing between ranked items within each attribute, but when actual attribute values have non-uniform gaps, the weighted sum of ranks cannot capture all possible additive utility functions that SEP preferences allow.

[12] Thus, practical relevance of SD-LEX and SD-WEIGHT depends not only on how they match underlying preferences, but also on how high are the behavioral benefits of simplified reporting relative to SD-DIRECT.

participants frequently misreport their preferences, and misreporting increases with preference complexity.[13] Reporting accuracy is significantly higher in ACCURACY-LEX than in ACCURACY-SEP or ACCURACY-COMP. This highlights that ordinal reporting over many alternatives is cognitively demanding, especially when attribute-based preferences are complex. The negative effect of complexity on truth-telling is replicated across all other treatments (SD-DIRECT, SD-LEX, SD-WEIGHT, SD-CHOICE).

Second, in the direct Serial Dictatorship treatments (SD-DIRECT-LEX, SD-DIRECT-SEP, SD-DIRECT-COMP), misreporting is widespread and often consistent with misperceived strategic incentives. Error rates are significantly higher than in the corresponding ACCURACY treatments, where no mechanism incentives are present. This complements earlier evidence on misreporting in the serial dictatorship mechanism (Li, 2017; Bó and Hakimov, 2024), showing that it is likely driven by misperceived incentives, and not solely by noise or by complexity of reporting the long rank-order list.

Third, simplifications in the reporting language (SD-LEX and SD-WEIGHT) do not improve accuracy of reporting. These reporting interfaces require three separate rankings (over universities, fields, and tuition levels) instead of a full ranking of 27 bundles. Yet overall the accuracy of reporting is significantly lower than in SD-DIRECT. Even under lexicographic preferences, SD-LEX does not outperform SD-DIRECT (accuracy rates are statistically indistinguishable between SD-LEX-LEX and SD-DIRECT-LEX, and significantly lower in SD-WEIGHT-LEX). Simplification thus fails to improve accuracy, even when the reporting interface matches the underlying preference structure.

Fourth, the sequential version of Serial Dictatorship (SD-CHOICE) performs best across all metrics. SD-CHOICE yields higher reporting accuracy than even the ACCURACY benchmark, particularly under complex preferences. This striking result suggests that SD-CHOICE improves outcomes not only because of stronger incentive properties, but also

[13] We use two measures of accuracy. One is choice accuracy that is a dummy for choosing/ranking the best program among the programs available to them when it is their turn to choose or be allocated an item. Second is the the Kendall–Tau distance between the submitted and true rankings of 27 programs.

because it changes the reporting interface. By eliminating full rankings and reducing participants' tasks to a single choice among remaining options, SD-CHOICE lowers cognitive demands. Its success therefore reflects both strategic simplicity (obvious strategy-proofness) and reporting interface simplification.

Finally, reporting accuracy translates into economically meaningful outcomes. Treatments with higher truthfulness generate more efficient and fair allocations, as reflected in lower justified envy and smaller efficiency losses. Misreporting is thus not confined to irrelevant parts of preference lists, but arises in consequential parts that directly affect allocations.

Overall, our findings highlight the challenges of accurately reporting multi-attribute preferences and show how complexity interacts with cognitive and strategic factors to generate misreporting. Sequential Serial Dictatorship (SD-CHOICE) delivers the strongest performance, combining strategic simplicity with a cognitively simple interface. When SD-CHOICE is not feasible, our evidence favors standard reporting over bundles rather than simplified formats—such as those used in practice in China—which fail to improve accuracy even when aligned with underlying preferences.

Literature. This paper relates to four strands of research. First, work on the costs of eliciting multi-attribute preferences shows that rich reporting is hard both technologically and cognitively. Compte and Jehiel (2004) formalize limits to communication and elicitation when preferences are high-dimensional. In combinatorial environments, Parkes (2005) shows how package structure and complementarities strain elicitation and computation. From a behavioral angle, Milgrom (2009, 2011) argue that multi-attribute trade-offs tax attention and working memory, predicting systematic error and heuristic reporting. In large-scale course allocation, Budish (2011) introduces an approximate competitive equilibrium mechanism (A-CEEI) that requires bidding for courses with fake money instead of ranking course bundles, and Budish and Kessler (2022) document that participants can communicate preferences accurately enough to realize the efficiency and fairness benefits of A-CEEI. But preference-reporting mistakes are common and meaningfully harm mechanism performance.

Sönmez and Ünver (2010) provide further evidence that course-assignment environments induce nontrivial reporting frictions. We build on this strand by separating the complexity of underlying preferences from the complexity of the reporting task and by benchmarking reports against a known ground truth.

Second, a growing literature documents misreporting under strategy-proof mechanisms and analyzes its sources both in the lab and in the field (Chen and Sönmez, 2006; Hakimov and Kübler, 2021; Rees-Jones and Shorrer, 2023). Using residency data, Rees-Jones (2018) show sizable deviations from truth-telling in dominant-strategy environments; Hassidim *et al.* (2021); Chen and Pereyra (2019); Shorrer and Sóvágó (2024) demonstrate that even well-incentivized centralized procedures exhibit systematic non-truthful play. Another line of papers tries to rationalize reasons behind the deviations, such as reference dependence, rank utility, disappointment aversion, and others (Dreyfuss *et al.*, 2022; Chen *et al.*, 2024; Meisner and Von Wangenheim, 2023; Meisner, 2023; Kloosterman and Troyan, 2023). Gonczarowski *et al.* (2023, 2024) demonstrate that the way mechanisms are explained—particularly whether strategy-proofness is made salient—can significantly affect behavior. Similarly, Katuščák and Kittsteiner (2024) show that simplifying the explanation of strategy-proof mechanisms increases truthful reporting. Finally, Guillen and Veszteg (2021) caution that lab estimates of truth-telling may be inflated due to experimenter demand effects, where participants conform to perceived expectations. Our contribution is to show that these deviations are not mere noise: holding the choice set fixed, misreporting rises with preference complexity; holding complexity fixed, misreporting is higher under direct serial dictatorship than in a non-allocative accuracy benchmark, consistent with misperceived incentives rather than list length alone.

Third, studies compare sequential (iterative) to direct implementations and quantify the role of menu size and feedback. In the lab, Klijn *et al.* (2019); Bó and Hakimov (2020) find that iterative deferred acceptance improves stability and welfare relative to direct submissions; Mackenzie and Zhou (2022); Bó and Hakimov (2024) show theoretically and in experiments that staging choices raises truthful play and assignment quality; and Haeringer and Iehlé (2021) analyze iterative designs that simplify participation. Stephenson (2022) shows that real-time assignment feedback during reporting increases equilibrium behavior in school choice. Dur *et al.* (2021) analyze sequential Boston mechanisms, finding similar improvements in participant outcomes. In higher-education admissions, Hakimov *et al.* (2023) show that sequential serial dictatorship facilitates information acquisition about preferences, and Grenet *et al.* (2022) in the field show importance of a dynamic multi-offer mechanism used in Germany for preference formation. Gong and Liang (2024) provide theory and experiment for a dynamic admissions mechanism that achieves efficient and stable outcomes under mild conditions. On incentives, Li (2017) establish that a sequential version of Random Priority is obviously strategy-proof, and Pycia and Troyan (2023) develop a broader simplicity framework clarifying why shrinking menus reduce incentive misperception. We contribute to this literature by again documenting superior performance of the sequential serial dictatorship mechanism, but showing that part of the performance is driven by a simpler interface—menus—rather than just stronger incentives.

Finally, platforms often restrict the message space to ease reporting, trading expressiveness for lower burden. One example is constrained lists, which are prevalent in practice worldwide, despite clear theoretical and experimental evidence that they harm incentives (Haeringer and Klijn, 2009; Calsamiglia *et al.*, 2010b). In practice, lexicographic formats are used to structure reporting: Hu *et al.* (2025) document institution–major submission in China, while Greenberg *et al.* (2024) describe a minimalist redesign for Army branching that limits what must be reported to meet policy goals. Alternative simplifications include budget-based inputs for course allocation (Budish, 2011; Budish and Kessler, 2022). Our experiments quantify the trade-off: attribute-based simplifications (lexicographic and weighted attributes) do not improve accuracy even when they match the preference domain.

The remainder of the paper proceeds as follows. Section 2 details our experimental design and procedures. Section 3 presents the key findings across reporting languages and preference structures. Section 4 concludes and offers implications for the design of multi-attribute matching mechanisms.

2 Experiment Design and Behavioral Hypotheses

2.1 Environment

2.1.1 Admissions Scenario

In each round, participants acted as applicants in a simulated college admissions market. Each market consisted of 27 students and 27 programs, with one seat per program. Programs were defined by three attributes: university prestige, field of study, and tuition. Attribute values were fixed as follows: prestige points were 500 (University A), 200 (University B), and 600 (University C); tuition points were 0 (no tuition), 250 (half tuition), and 500 (full tuition); field-of-study points were $\{300, 500, 700\}$, with Economics, Finance, and Law randomly assigned to these values each market.

Each participant received a *preference-score formula* (utility function) mapping attribute values into scores for every program. Higher scores implied higher positions in the participant's induced ordinal ranking. An on-screen calculator allowed participants to compute scores for any program. Our design replaces per-seat monetary payoffs with these formulas: scores induce the participant's true ordinal ranking, and payoffs depend only on the rank of the assigned program, not on cardinal scores.[14] This ensures comparable incentives across treatments and rounds and allows us to vary preference complexity while abstracting from strategic motives.

[14] We avoid usual assignment of payoffs to each program, as the reporting even long rankings become trivial and does not allow us to change complexity of preferences. These reasons are similar to discussion of Budish and Kessler (2022) on the unsuitability of inducing preferences in a format immediately reportable to the mechanism. However, we deliberately adopt the approach that is criticized by Budish and Kessler (2022): one "language" assigns incentives (utility formulas), and participants must translate it into the reporting language of the mechanism. While we agree with the general criticism, this serves our purpose: it allows us to separate the complexity of *inducing* preferences from the complexity of *reporting* them, while holding preferences fixed.

In the SD-DIRECT, SD-WEIGHT, and SD-LEX treatments, priorities were determined by exam marks (uniformly drawn without replacement from $\{0, \ldots, 100\}$ to avoid ties). Participants knew their own mark but not others'. In SD-CHOICE, priorities were revealed sequentially as participants were called to act in order. In ACCURACY, no priority order was used as there is no assignment. Each session consisted of 12 rounds, with new preference formulas and marks each round.

2.1.2 Mechanisms

In SD-DIRECT, SD-WEIGHT, and SD-LEX, programs were allocated via standard serial dictatorship. The mechanism uses the ranked ordered lists of programs by participants as an input. The participant with the highest mark received their top ranked program. The participants with the second highest mark received the top ranked program among available, and so on.

In SD-CHOICE, we implemented sequential serial dictatorship. Participants did not submit rankings; instead, the student with the highest priority selected their most preferred program among all programs, then the students with the second-highest mark selected a program from the remaining ones, and so forth.[15]

In ACCURACY, no allocation took place. Participants were rewarded purely for reporting rankings that matched their true preferences. Accuracy was measured using the normalized Kendall distance, the fraction of discordant pairs between the reported and true rankings:

$$\text{Kendall distance} = \frac{\text{number of discordant pairs}}{\text{total number of pairs}}.$$

The measure ranges from 0 (identical rankings) to 1 (complete reversal).[16]

[15] To keep 12 rounds while minimizing time, we ran them in parallel. Conceptually, there were 12 independent admissions processes. Each of 27 "turns" advanced all 12 processes by one pick: we randomly selected 12 participants (one per process) to choose in the current turn while the other 15 waited. After 27 turns, all 12 processes were complete and each participant had made exactly one choice in each process.

[16] Example: Ranking 1 is A, B, C, D; Ranking 2 is B, A, D, C. The discordant pairs are (A, B) and (C, D), so there are 2 discordances out of $\binom{4}{2} = 6$ possible pairs, giving a normalized Kendall distance of $2/6 = 0.333$.

2.1.3 Payoffs

In all treatments, one of the 12 rounds was randomly selected for payment. In all treatments except ACCURACY, participants received CNY 160 if assigned their truly most-preferred program (based on their induced preference formula), CNY 155 for their true second-ranked program, and so on, decreasing by CNY 5 per rank to CNY 30 for the least-preferred program.[17] In ACCURACY, the payoff was

$$\text{Payoff} \;=\; 160 \times \big(1 - \text{Kendall distance}\big),$$

so closer agreement with the true ranking yielded higher earnings.

2.2 Treatment Variations

We implement a 3×5 design combining three levels of preference complexity with five reporting interfaces/mechanisms. Preference complexity varies *within* subjects across rounds; the reporting interface/mechanism varies *between* subjects. We now describe both dimensions.

2.2.1 Complexity of Preferences

Participants faced three preference domains induced by preference-score formulas. Let $U \in \{200, 500, 600\}$ denote university prestige points, $F \in \{300, 500, 700\}$ field-of-study points (major-to-points mapping randomized each market), and $T \in \{0, 250, 500\}$ tuition points (higher T means higher tuition). For each subject and round, coefficients are drawn independently and uniformly from the indicated intervals.

1. **Lexicographic (LEX).** Scores are

$$s = a\,U + b\,F - c\,T, \quad a \in [90, 110],\; b \in [9, 11],\; c \in [0.9, 1.1].$$

[17] Participants received a payoff table in the instructions (Appendix C).

The order-of-magnitude separation of coefficients enforces the lexicographic priority $U \succ F \succ -T$ on the feasible attribute ranges, so differences in U dominate any differences in F or T, and differences in F dominate T.

2. **Additively separable (SEP).** Scores are

$$s = a\,U + b\,F - c\,T, \quad a, b, c \in [30, 40],$$

so attributes carry comparable weights and trade-offs must be computed across all three dimensions.

3. **Non-separable with complementarities (COMP).** Scores are

$$s = a\,U + b\,F - c\,T + d\,U \cdot T, \quad a, b, c \in [30, 40],\ d \in [-5, 5],$$

introducing an interaction between prestige and tuition. Positive d weakens the tuition penalty at high-prestige universities (willingness to pay for prestige); negative d strengthens it.

The 12 rounds are organized into four blocks of three rounds (except in SD-CHOICE due to the parallel implementation described above). In the first three rounds, participants face all three preference types (LEX, SEP, COMP), one per round in random order. In the next three rounds, they again face all three preference types in a different random order, and this pattern continues for the subsequent two blocks of three rounds. The block design allows for a balance of experience across preference domains.

2.2.2 Preference Reporting Languages

1. **SD-DIRECT.** After observing their exam mark, participants submit a full ranking of the 27 programs.

2. **SD-WEIGHT.** Participants rank each attribute (university, field of study, tuition) separately and report attribute weights in $[0, 100]$. A composite "points" index is computed and minimized to produce the implied ranking of 27 programs.[18]

3. **SD-LEX.** Participants rank each attribute separately; the ranking of 27 programs is constructed by imposing a lexicographic structure with university $\succ$ field $\succ$ tuition.[19]

4. **SD-CHOICE.** Participants act in order of priority. When called, a participant observes the remaining programs and selects one.[20]

5. **ACCURACY.** No allocation occurs. Participants submit a ranking of the 27 programs and are paid based on the (normalized) Kendall distance between their submitted and true rankings.

2.3 Procedures

We ran the experiments at Wuhan University's Research Center for Behavioral Science during 2024–2025 with 810 Wuhan University students. Table 1 summarizes the sessions by treatment. For each treatment, we conducted six sessions with 27 participants each (one independent matching group per session), yielding 162 participants per treatment. Participants were randomly assigned to one treatment and took part in only that treatment. Payments were made privately by transfer at the end of the session.

[18]Displayed to participants:

$$\text{Points} = \text{rank}(U) \times w_U + \text{rank}(F) \times w_F + \text{rank}(T) \times w_T.$$

Lower points imply a higher bundle rank. Example: if U=A, F=Economics, T=No tuition are ranked first, and Finance is ranked second for F, with $(w_U, w_F, w_T) = (50, 20, 70)$, then $[A, \text{Economics}, \text{No tuition}]$ has $1 \times 50 + 1 \times 20 + 1 \times 70 = 140$, whereas $[A, \text{Finance}, \text{No tuition}]$ has $1 \times 50 + 2 \times 20 + 1 \times 70 = 160$.

[19]Displayed to participants:

$$\text{Points} = 100 \times \text{rank}(U) + 10 \times \text{rank}(F) + 1 \times \text{rank}(T).$$

Lower points imply a higher bundle rank. Example: if U=A, F=Economics, T=No tuition are ranked first, and Finance is ranked second for F, then $[A, \text{Economics}, \text{No tuition}]$ has $1 \times 100 + 1 \times 10 + 1 \times 1 = 111$, while $[A, \text{Finance}, \text{No tuition}]$ has $1 \times 100 + 2 \times 10 + 1 \times 1 = 121$.

[20]Participants were allowed to submit an empty choice; doing so yielded zero payoff for that round.

Table 1: Experimental Design

Treatment	Participants	Sessions	Reporting Interface	Allocation Mechanism
ACCURACY	162	6	Full ranking	None; incentivized truthful reporting
SD-DIRECT	162	6	Full ranking	Serial Dictatorship
SD-WEIGHT	162	6	Attribute rankings + weights	Serial Dictatorship
SD-LEX	162	6	Attribute rankings	Serial Dictatorship
SD-CHOICE	162	6	Pick when called	Sequential Serial Dictatorship

At the beginning of each session, the 27 participants were given the written experimental instructions, allowing them to follow along as the experimenters read the instructions aloud. The instructions described the environment, allocation procedures, and payoffs. For each round a countdown timer limited reporting time: 3 minutes per round in SD-CHOICE and 8 minutes per round in all other treatments. Sessions lasted on average 112 minutes. Average earnings were CNY 113.65 with no show-up fee, above typical payments in China.

2.4 Behavioral Hypotheses

We test how preference complexity and the reporting interface/mechanism affect reporting accuracy and allocation outcomes.

To simplify exposition, we first introduce some notations.

Let $\mathcal{T} = \{\text{ACCURACY}, \text{SD-DIRECT}, \text{SD-LEX}, \text{SD-WEIGHT}, \text{SD-CHOICE}\}$ be the between-subjects treatments and $\mathcal{D} = \{\text{LEX}, \text{SEP}, \text{COMP}\}$ the preference domains. For $t \in \mathcal{T}$ and $d \in \mathcal{D}$, write $\text{Acc}(t, d)$ for expected reporting accuracy.

First, we hypothesize that the complexity of preferences hurts outcomes. Prior work shows that higher task complexity raises errors (Kahneman, 2003; Gigerenzer and Kahneman, 2008); non-separabilities increase comparison difficulty (Milgrom, 2009, 2011); and preference reporting can be costly (Parkes, 2005; Budish and Kessler, 2022; Sönmez and Ünver, 2010).

Behavioral Hypothesis 1 (Complexity decreases reporting accuracy across all treatments)**.** *Fix any between-subjects treatment $t \in \mathcal{T}$. Reporting accuracy declines with preference com-*

plexity:

$$\text{Acc}(t, \textit{LEX}) \ > \ \text{Acc}(t, \textit{SEP}) \ > \ \text{Acc}(t, \textit{COMP}).$$

Second, we hypothesize that rates of misreporting in the direct serial dictatorship mechanism are higher than in the ACCURACY treatment. This would point to the fact that previously documented misreporting in serial dictatorship is nonrandom and consistent with misperceived incentives, because ACCURACY removes strategic considerations.

Behavioral Hypothesis 2 (Incentive misperception under direct serial dictatorship)**.** *For each preference domain $d \in \mathcal{D}$,*

$$\text{Acc}(\textit{ACCURACY}, d) \ > \ \text{Acc}(\textit{SD-DIRECT}, d).$$

Third, we hypothesize that simplified reporting interfaces increase accuracy when they allow underlying preferences to be reported correctly (for example, SD-LEX for lexicographic preferences), and lower it otherwise, as approximation error offsets reduced input burden. This is in line with Milgrom (2009, 2011), who emphasize that carefully simplifying the message space can relieve cognitive stress, but only if it does not overly constrain one's ability to convey true preferences.

Behavioral Hypothesis 3 (Attribute-based simplifications improve accuracy in LEX but not in SEP and COMP)**.** *For the LEX domain and treatment $t \in \{\textit{SD-LEX}, \textit{SD-WEIGHT}\}$,*

$$\text{Acc}(t, \textit{LEX}) \ \geq \ \text{Acc}(\textit{SD-DIRECT}, \ \textit{LEX}),$$

for domain $d \in \{\textit{SEP}, \textit{COMP}\}$ and treatment $t \in \{\textit{SD-LEX}, \textit{SD-WEIGHT}\}$,

$$\text{Acc}(t, d) \ \leq \ \text{Acc}(\textit{SD-DIRECT}, d).$$

Fourth, SD-CHOICE simplifies both incentives—making the mechanism obviously strategy-proof—and the reporting interface by narrowing the active choice set (Li, 2017; Pycia and Troyan, 2023; Chernev *et al.*, 2015). Thus, we expect improvements not only over the treatments with direct serial dictatorship (consistent with the effect of obvious strategy-proofness) but also over ACCURACY, which would be driven by a simpler reporting language.

Behavioral Hypothesis 4 (SD-CHOICE improves accuracy over all other treatments)**.** *For each domain $d \in \mathcal{D}$,*

$$\mathrm{Acc}(\textit{SD-CHOICE}, d) > \mathrm{Acc}(\textit{SD-DIRECT}, d); \mathrm{Acc}(\textit{SD-CHOICE}, d) \geq \mathrm{Acc}(\textit{ACCURACY}, d).$$

Finally, while all hypotheses are formulated about accuracy, we also analyze efficiency and justified envy of the resulting allocations in Section 3.5. We expect deviations from accurate reporting to occur in parts of the preference ranking that matter for allocation given priorities. Accordingly, treatment differences in reporting accuracy should lead to analogous differences in efficiency and justified envy.

3 Results

Unless stated otherwise, statistical significance is assessed at the 5% level. For pairwise treatment comparisons, we report p-values on treatment coefficients from regressions run on the relevant two-treatment subsample, with standard errors clustered at the matching-group level.

3.1 Preference complexity and reporting accuracy

We use two measures of reporting accuracy. The first, *choice accuracy*, evaluates whether a participant ultimately receives her most-preferred program (by the *true* ranking) among the seats available at her turn. This measure enables a fair comparison between direct mechanisms and SD-CHOICE. In ACCURACY, there is no allocation; to create comparable choice sets, we simulate Random Priority markets using participants' reported preferences: in each market we draw a random priority order without replacement and allocate by serial dictatorship, repeating this 100 times and averaging at the market level.[21] Formally, choice accuracy equals one if the participant's top-ranked reported program among the available set coincides with her top-ranked true program among that set.

A potential concern is that choice accuracy may downweight errors that occur by low-priority participants. We therefore also report a cardinal measure of deviation from truthful reporting that ignores priority: the normalized Kendall distance between the submitted and true rankings over all 27 items, taking values in $[0, 1]$ ($0 =$ identical; $1 =$ complete reversal).[22] Note that Kendall Distance cannot be calculated in SD-CHOICE, as we observe only one choice per participant.

Table 2: Summary Table: Reporting Accuracy

Treatments	Choice Accuracy (%)				Kendall Distance			
	Overall	LEX	SEP	COMP	Overall	LEX	SEP	COMP
ACCURACY	76.72	92.00	67.77	70.38	0.11	0.04	0.13	0.17
SD-DIRECT	55.76	63.89	49.54	53.86	0.37	0.23	0.46	0.41
SD-WEIGHT	49.85	57.56	46.76	45.22	0.42	0.29	0.49	0.48
SD-LEX	47.84	65.43	40.28	37.81	0.39	0.20	0.50	0.47
SD-CHOICE	84.26	91.67	85.65	75.46	$\times$	$\times$	$\times$	$\times$

Notes: Choice Accuracy is the market-level average share of participants who, at their turn, select the highest-ranked available program by their reported ranking; 27 participants per market; 72 markets per treatment. Kendall Distance is the average normalized Kendall (tau) distance between a participant's reported and true rankings. Kendall distance is not defined for SD-CHOICE because no full ranking is elicited.

[21] For each market of 27 participants, we run 100 independent simulations. In each, we draw a unique priority order (1–27) without replacement and allocate via Random Priority under reported preferences. For a given participant and simulation, choice accuracy equals 1 if her top reported program among the available seats at her turn matches her top true program among those same available seats. We then average within market.

[22] An alternative and more typical measure would be a dummy for submitting the exact truthful ranked list. In typical experiments with short lists (e.g., up to eight items; see Hakimov and Kübler (2021)) this is informative. In our setting with 27 items, exact truth-telling is rare, so a dummy would discard too much information; moreover, it is by design that it can only be zero in SD-WEIGHT and SD-LEX under SEP and COMP, where interfaces do not allow for exactly truthful full rankings.

As Table 2 shows (columns 2–5 under "Choice Accuracy"), LEX preferences yield the highest accuracy across all treatments, with differences relative to SEP and COMP significant at $p < 0.01$. For example, in ACCURACY, choice accuracy is 92.00% in LEX but only 67.77% in SEP and 70.38% in COMP. The same pattern appears in Kendall distance: in every treatment where Kendall distance is defined, LEX has the lowest distance, significantly below SEP and COMP ($p < 0.01$). Nonparametric tests at the market level (Kruskal–Wallis) also reject equality across preference domains ($p < 0.01$).

By contrast, the ranking between SEP and COMP is not systematic and depends on the reporting interface. Choice accuracy and Kendall distance occasionally disagree on which of SEP or COMP performs better. Overall, we do not reject equality between SEP and COMP.

Table 3 confirms these patterns. In columns (2) and (4), the COMP dummy is sizeable and highly significant ($p < 0.01$), indicating larger Kendall distances and a lower probability of choosing the correct program. LEX, the simplest domain, consistently yields the highest accuracy (and lowest Kendall distance). There is no significant difference between SEP and COMP coefficients.

Thus, HYPOTHESIS 1 is partially supported: LEX generates the most accurate reporting, but there is no robust additional decrease in accuracy from SEP to COMP. One interpretation is that SEP already imposes sufficient computational burden to trigger substantial errors, so further nonseparabilities in COMP do not produce a clear, incremental drop in accuracy.[23]

Result 1 (Complexity-induced misreporting)**.** *As preference complexity increases—from lexicographic to separable or nonseparable—choice accuracy decreases and Kendall distance increases significantly. We find no robust statistical difference between SEP and COMP.*

[23] LEX is the only domain that can often be resolved without extensive calculator use or cross-attribute computations. Under SEP and COMP, subjects must rely on on-screen calculations and documentation of scores for many programs; this likely depresses accuracy in both domains. While COMP is theoretically more complex, the calculator may attenuate the incremental difficulty relative to SEP, yielding similar accuracy.

Table 3: Choice Accuracy and Kendall Distance Regressions

	Choice Accuracy		Kendall Distance	
	(1)	(2)	(3)	(4)
Treatment:				
SD-DIRECT	-0.207***	-0.207***	0.254***	0.254***
	(0.023)	(0.023)	(0.007)	(0.008)
SD-WEIGHT	-0.267***	-0.267***	0.306***	0.306***
	(0.013)	(0.013)	(0.007)	(0.007)
SD-LEX	-0.288***	-0.288***	0.275***	0.275***
	(0.012)	(0.012)	(0.008)	(0.008)
SD-CHOICE	0.075***	0.075***		
	(0.014)	(0.014)		
Pref. Type:				
SEP		-0.169***		0.206***
		(0.017)		(0.016)
COMP		-0.176***		0.194***
		(0.017)		(0.011)
Constant			0.114***	-0.019*
			(0.006)	(0.010)
Observations	9720	9720	7774	7774

Notes: Mixed effects regressions. Choice Accuracy is a dummy, which equals to 1 if a participant selected the highest-ranked available program (out of programs remaining after higher-ranked participants have made their selections) based on their reported ranking, and 0 otherwise. Kendall Distance represents the proportion of discordant pairs in a participant's reported ranking list compared to her true preferences, ranging from 0 to 1. Choice accuracy effects expressed in marginal effects from logistic mixed effects regressions. The baseline preference type is LEX. The baseline treatment is ACCURACY. For ACCURACY treatment, data from only the first simulation is used (this does not affect the results). Standard errors in parentheses cluster by matching group, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

3.2 Incentive misperception under direct serial dictatorship

We compare ACCURACY to SD-DIRECT to isolate the role of incentive misperception in the serial dictatorship mechanism. The two treatments use the same reporting task (a full rank-order list) and the same induced preferences; ACCURACY removes mechanism incentives incentivizing only truthful reporting. We refer to the gap in accuracy between these treatments as strategic misreporting—a composite of incentive-related forces (e.g., misperceived incentives, reference dependence, rank utility, disappointment aversion). Importantly, this gap is not attributable to list-length burden, preference complexity or demand effect, which are held fixed across the two treatments.

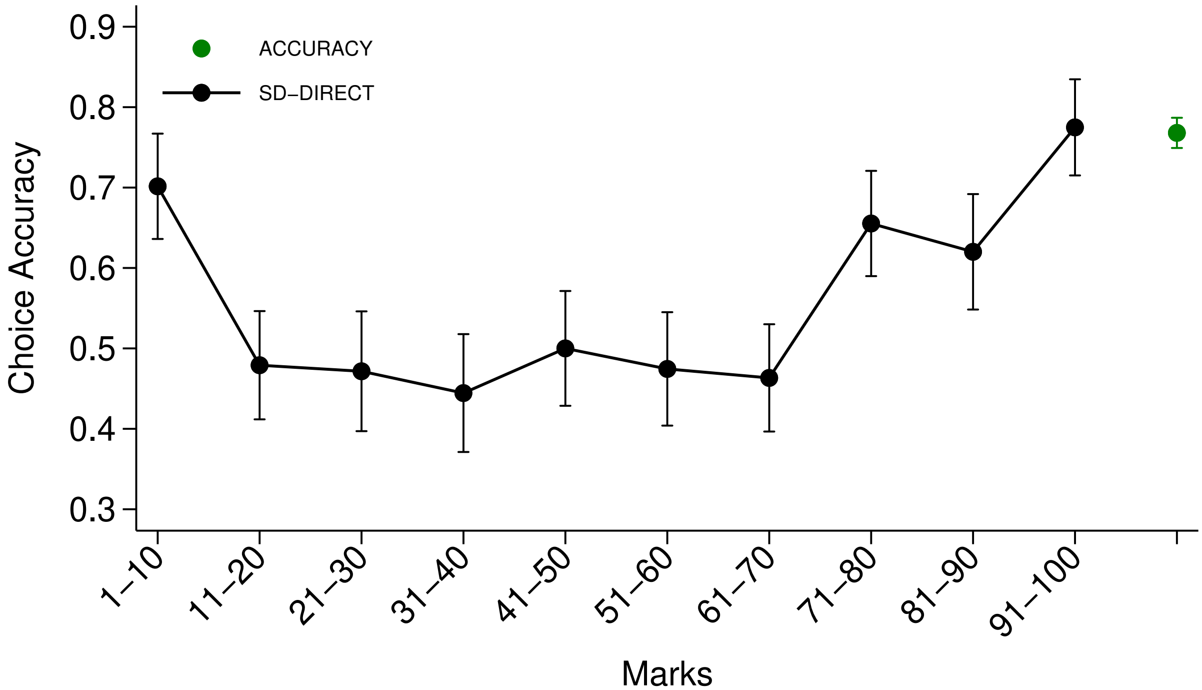
As shown in Table 2, overall choice accuracy in SD-DIRECT is 55.76% versus 76.72% in ACCURACY, a 21 percentage point difference. The gap is significant overall and within each preference domain ($p < 0.01$). Thus, while complexity of reporting contributes to errors in both treatments, SD-DIRECT exhibits substantial additional misreporting consistent with strategic misreporting under Serial Dictatorship.

Further evidence comes from the gradient of accuracy by marks in SD-DIRECT presented in Figure 1. Participants with very high marks (91-100) achieve high choice accuracy (around 77%), as they need only ensure their top-ranked program is correctly placed first in their submitted ranking. Those with very low marks (1-10) also maintain relatively high accuracy (around 70%), likely because with most programs already taken by higher-priority participants, they need only correctly identify their preferred option among the few that will remain available. In contrast, participants with intermediate marks (31-70) exhibit the lowest accuracy (around 44-50%), suggesting these participants face the most strategic confusion—they may overthink their rankings, attempting to game the system when they should simply report truthfully.[24]

These findings support Hypothesis 2. Mixed-effects regressions in Table 3 corroborate

[24] This pattern aligns with Hassidim *et al.* (2021), who show that applicants with poor grades are more likely to submit dominated rank-order lists than those with better grades.

Figure 1: The Role of Marks



Notes: The groups (e.g., "1–10", "91–100") refer to the marks assigned to participants. CIs at 95%.

the raw differences: the SD-DIRECT indicator is associated with significantly lower choice accuracy and higher Kendall distance relative to ACCURACY (all $p < 0.01$).

Result 2 (Incentive misperception under direct serial dictatorship)**.** *Within each domain (LEX, SEP, COMP), choice accuracy is significantly higher and Kendall distance significantly lower in ACCURACY than in SD-DIRECT.*

3.3 Reporting interfaces and accuracy

We next compare the three direct reporting interfaces: SD-DIRECT (a full ranking over 27 programs), SD-WEIGHT (separate rankings over the three attributes plus attribute weights), and SD-LEX (separate attribute rankings aggregated lexicographically). Behavioral Hypothesis 3 predicts that simpler formats should reduce misreporting provided they do not impose distortions.

In the LEX domain, SD-LEX and SD-DIRECT are fully expressive (SD-WEIGHT can approximate lexicography when weights are sufficiently separated). In the data (Table 2, LEX columns), SD-LEX attains the highest choice accuracy and the lowest Kendall distance among the three, but its advantage over SD-DIRECT is not statistically significant ($p > 0.10$); by contrast, SD-WEIGHT performs significantly worse than SD-LEX ($p < 0.01$) and worse than SD-DIRECT ($p < 0.05$). Hence we do not find a clear accuracy gain from simplified reporting even when the interface matches the preference domain.

When preferences require trade-offs (SEP) or include complementarities (COMP), SD-DIRECT yields the highest choice accuracy and the lowest Kendall distance; the advantage is significant relative to SD-LEX in SEP and significant relative to both SD-LEX and SD-WEIGHT in COMP ($p < 0.01$). As complexity rises, SD-WEIGHT delivers higher choice accuracy than SD-LEX ($p < 0.01$), consistent with weights partially offsetting lexicographic misspecification, although this ordering is not mirrored by Kendall distance.

Aggregating across domains, neither SD-LEX nor SD-WEIGHT improves accuracy relative to SD-DIRECT; SD-DIRECT achieves the highest overall choice accuracy and the lowest Kendall distance. This results goes against the practice of simplifying reporting thorough this interfaces (SD-LEX used in China) as they fail do outperform unconstrained list reporting even in the respective preferences domain. Since designers typically do not know the true preference domain and some participants likely have complex preferences, SD-DIRECT appears the best reporting interface for the direct mechanisms.

Result 3 (Reporting interfaces)**.** *Relative to SD-DIRECT, neither SD-WEIGHT nor SD-LEX significantly increases accuracy—even in LEX. Aggregating across preferences domains, SD-DIRECT yields the highest choice accuracy and the lowest Kendall distance, with significant differences relative to SD-LEX and SD-WEIGHT.*

3.4 Sources of benefits of SD-CHOICE

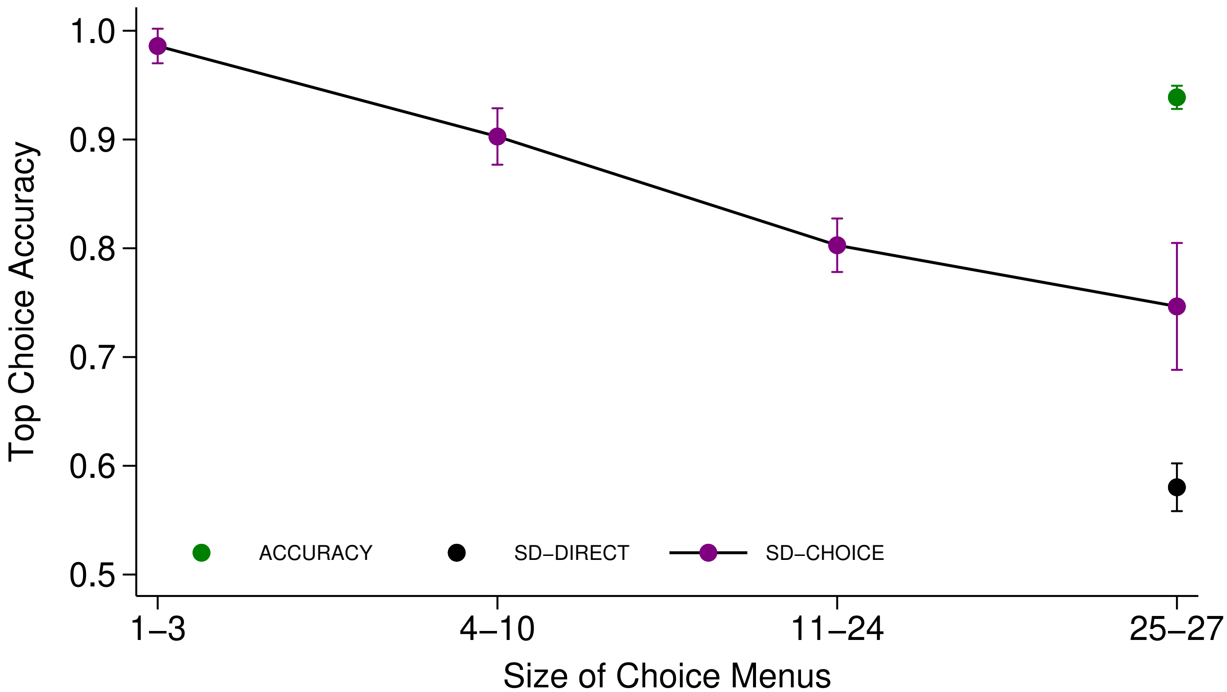
Tables 2 and 3 show that SD-CHOICE delivers substantially higher choice accuracy than SD-DIRECT (84.26% vs. 55.76%), a 28.5 percentage point gain ($p < 0.01$). This replicates Li (2017); Bó and Hakimov (2024) and is consistent with incentive strength: SD-CHOICE is obviously strategy-proof and "1-step simple" (Pycia and Troyan, 2023), whereas SD-DIRECT is not. However, SD-CHOICE also simplifies the reporting interface: participants make a single pick from the current menu rather than constructing a full ranking over 27 items. Hence the improvement over SD-DIRECT can arise from two channels—stronger incentives and lower reporting complexity. Our design permits a bound on the incentive channel by comparing SD-CHOICE to ACCURACY, which retains the same induced preferences but removes strategic incentives and asks for a full ranking. If SD-CHOICE outperforms ACCURACY, the difference cannot be due to incentives alone and points to an interface advantage. Indeed, SD-CHOICE exceeds ACCURACY overall (84.26% vs. 76.72%, $p < 0.01$).[25] These findings support Hypothesis 4: SD-CHOICE improves accuracy relative to SD-DIRECT, and part of the improvement reflects the simpler, one-at-a-time reporting interface rather than incentives alone.

Additional evidence on the benefits of the simplified interface comes from Figure 2. It shows that choices in SD-CHOICE become more accurate with smaller menus. Importantly, for very large menus (more than 25 options), the accuracy of the top choice in ACCURACY is significantly higher than the accuracy of the chosen option in SD-CHOICE, indicating that the better performance of SD-CHOICE over ACCURACY arises from smaller menus—reemphasizing the interface channel rather than only stronger incentives.

Result 4 (Strategic vs. interface channels)**.** *SD-CHOICE significantly increases choice accuracy relative to SD-DIRECT. Moreover, SD-CHOICE outperforms ACCURACY overall, indicating that gains arise not only from obvious strategy-proofness but also from a simpler*

[25] The gap is largest in SEP ($p < 0.01$), while differences in LEX and COMP are not statistically significant. The latter depends on the simulated priority orders for ACCURACY.

Figure 2: Choice Accuracy and Menu Size



Notes: The figure reports a scatter plot of top-choice accuracy by preference domain. "Top-choice accuracy" indicates whether a subject correctly selected or ranked her truly most-preferred item first when making a choice or submitting a full ranking.

reporting interface.

3.5 Efficiency Loss and Justified Envy

Reporting failures affect not only individual payoffs but also the aggregate outcome. To quantify these, we measure: (i) *efficiency loss*, the shortfall of realized welfare from the welfare achievable under truthful reporting (computed at the market level), and (ii) *justified envy*, the fraction of participants who end up preferring someone else's allocated seat despite having higher priority.

We now show that the treatment differences documented above for choice accuracy and Kendall distance translate into economically meaningful differences in welfare and fairness. Table 4 reports efficiency loss (shortfall from the welfare achievable under truthful reporting) and the share of participants with justified envy (a higher-priority participant preferring

Table 4: Summary Table: Efficiency Loss and Justified Envy

Treatments	Efficiency Loss (%)				Justified Envy (%)			
	Overall	LEX	SEP	COMP	Overall	LEX	SEP	COMP
ACCURACY	5.87	1.93	6.80	8.87	23.28	8.00	32.23	29.62
SD-DIRECT	8.79	3.98	9.32	13.07	43.88	35.34	50.15	46.14
SD-WEIGHT	8.64	4.19	8.53	13.18	50.15	42.44	53.24	54.78
SD-LEX	9.13	3.21	7.99	16.19	52.16	34.57	59.72	62.19
SD-CHOICE	1.85	0.32	1.10	4.14	15.64	8.18	14.20	24.54

Notes: Efficiency Loss is the average percentage of welfare loss relative to the optimal allocation at the market level, calculated by

$$E = \frac{M - R}{R} \times 100,$$

where M is the ideal payoff achievable under truthful preferences, and R is the realized payoff. Justified Envy reports the average proportion of participants exhibiting justified envy. A participant exhibits justified envy if another participant with a lower priority is assigned a seat that the envious participant prefers over her own assigned seat.

a lower-priority participant's assignment). Across treatments and within each preference domain, lower accuracy coincides with higher efficiency loss and more envy. Moving from LEX to SEP to COMP, efficiency loss rises in every treatment (e.g., under SD-DIRECT: 3.98% $\rightarrow$ 9.32% $\rightarrow$ 13.07%; under ACCURACY: 1.93% $\rightarrow$ 6.80% $\rightarrow$ 8.87%), indicating that misreporting in more complex domains is costlier.

Aggregating across domains, SD-CHOICE attains the lowest efficiency loss (1.85% overall; 4.14% in COMP) and the lowest justified envy (15.64% overall), while SD-LEX perform worst on both metrics (overall justified envy above 50%). Mixed-effects regressions in Table 5 confirm these patterns: relative to ACCURACY, SD-DIRECT, SD-LEX, and SD-WEIGHT significantly increase efficiency loss and justified envy (all $p < 0.01$), whereas SD-CHOICE significantly reduces both (efficiency loss coefficient -0.038, $p < 0.01$; justified envy coefficient -0.076, $p < 0.01$). Preference complexity independently worsens outcomes: SEP and COMP enter positively and significantly in both specifications. Taken together, these results show that reporting errors occur in consequential parts of the preference lists and map directly into lower realized welfare and more violations of priority fairness; sequential

Table 5: Efficiency Loss and Justified Envy Regressions

	Efficiency Loss		Justified Envy	
	(1)	(2)	(3)	(4)
Treatment:				
SD-DIRECT	0.031***	0.031***	0.203***	0.203***
	(0.009)	(0.009)	(0.023)	(0.023)
SD-WEIGHT	0.030***	0.030***	0.267***	0.266***
	(0.007)	(0.007)	(0.013)	(0.013)
SD-LEX	0.035***	0.035***	0.288***	0.288***
	(0.006)	(0.006)	(0.012)	(0.012)
SD-CHOICE	-0.038***	-0.038***	-0.076***	-0.076***
	(0.005)	(0.005)	(0.014)	(0.014)
Pref. Type:				
SEP		0.041***		0.170***
		(0.004)		(0.017)
COMP		0.083***		0.178***
		(0.008)		(0.017)
Constant	0.057***	0.015***		
	(0.005)	(0.005)		
Observations	360	360	9720	9720

Notes: Mixed effects regressions. Efficiency Loss represents the percentage of welfare loss relative to the optimal allocation at the market level. Justified Envy is a dummy, which equals to 1 if a participant exhibited justified envy, and 0 otherwise. Justified envy effects expressed in marginal effects from logistic mixed effects regressions. The baseline preference type is LEX. The baseline treatment is ACCURACY. For ACCURACY treatment, data from only the first simulation is used (this does not affect the results). Standard errors in parentheses cluster by matching group, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

choice mitigates both by improving reporting.

Result 5 (Efficiency and envy). *Treatment differences in reporting accuracy translate into meaningful welfare losses and priority violations. SD-CHOICE achieves significantly lower efficiency loss and justified envy than the other mechanisms, consistent with its accuracy advantage.*

4 Conclusion

This paper studies how preference complexity and the reporting interface shape reporting accuracy and, in turn, welfare and justified envy in multi-attribute matching. Three facts emerge. First, complexity matters: LEX yields the highest accuracy; SEP and COMP substantially reduce accuracy with no robust difference between them. Second, strategic considerations matter: holding the reporting task fixed, ACCURACY outperforms SD-DIRECT, consistent with strategic misreporting under Serial Dictatorship. Third, the interface matters: attribute-based simplifications (SD-LEX, SD-WEIGHT) do not improve accuracy relative to a full ranking—even in LEX—while a sequential implementation (SD-CHOICE) dominates on accuracy and, consequently, on efficiency and justified envy. A higher choice accuracy than in ACCURACY and menu-size analysis indicates that part of SD-CHOICE's advantage arises from reducing the cognitive burden of full-list construction, not only from stronger incentive properties.

Two design lessons follow. First, *domain-assuming simplifications* are risky. Interfaces that force lexicographic or weighted-attribute aggregation can reduce input burden, but they restrict expressiveness and, in our data, do not deliver better accuracy even when the true domain is lexicographic. Because designers rarely know the distribution of preference domains ex ante and at least some participants plausibly have complex preferences, favoring a full list over constrained, domain-assuming inputs is safer among direct mechanisms. Second, *sequential implementations* are powerful. SD-CHOICE combines obvious strategy-proofness with a cognitively light interface (one pick from a shrinking set), and the gains appear both in accuracy and in downstream outcomes.

We view fully sequential (one-by-one) implementations as a normative benchmark but recognize operational frictions at scale. Centralized platforms may face heavy communication, timing, and coordination costs if every participant must act in serial order. In practice, dynamic implementations can approximate small menus without literal one-by-one moves. Multi-offer or staged procedures progressively resolve uncertainty about high-priority candidates' choices, shrinking remaining menus for others. This de facto "menu pruning" is central to recent deployments and proposals in higher education: dynamic or hybrid designs that batch early offers and elicit updated preferences have been used in France (e.g., Hakimov *et al.*, 2023), Inner Mongolia (Gong and Liang, 2024), and Tunisia (Luflade, 2017), and related mechanisms show preference discovery benefits in university admissions (Grenet *et al.*, 2022). Our results provide microfoundations for these policies: smaller, more informative menus reduce errors and improve both efficiency and priority fairness.

Finally, for direct mechanisms alternative reporting interfaces could be studied and developed. Our deviations from full lists, motivated by practices such as the lexicographic format in China, did not yield behavioral benefits. However, this does not imply that other simplified interfaces cannot be designed—especially with the development of AI tools (see, e.g., Soumalias *et al.* (2025)). The cost of direct reporting remains substantial and is likely even higher in real-world applications with far larger choice sets. We believe that developing improved reporting interfaces is crucial and leave this for future research.

References

Abdulkadiroğlu, A. and Sönmez, T. (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, **66** (3), 689–701.

— and Sönmez, T. (1999). House allocation with existing tenants. *Journal of Economic Theory*, **88** (2), 233–260.

Abdulkadiroglu, A. and Sönmez, T. (2003). School choice: A mechanism design approach. *The American Economic Review*, **93** (3), 729–747.

Balinski, M. and Sönmez, T. (1999). A tale of two mechanisms: student placement. *Journal of Economic Theory*, **84** (1), 73–94.

Bó, I. and Hakimov, R. (2020). Iterative versus standard deferred acceptance: Experimental evidence. *The Economic Journal*, **130** (626), 356–392.

— and Hakimov, R. (2024). Pick-an-object mechanisms. *Management Science*, **70** (7), 4693–4721.

Budish, E. (2011). The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes. *Journal of Political Economy*, **119** (6), 1061–1103.

— and Cantillon, E. (2012). The multi-unit assignment problem: Theory and evidence from course allocation at harvard. *American Economic Review*, **102** (5), 2237–2271.

— and Kessler, J. B. (2022). Can market participants report their preferences accurately (enough)? *Management Science*, **68** (2), 1107–1130.

Calsamiglia, C., Haeringer, G. and Klijn, F. (2010a). Constrained school choice: An experimental study. *American Economic Review*, **100** (4), 1860–1874.

—, — and — (2010b). Constrained school choice: An experimental study. *American Economic Review*, **100** (4), 1860–1874.

Caspari, G. and Khanna, M. (2025). Nonstandard choice in matching markets. *International Economic Review*, **66** (2), 757–786.

Chen, L. and Pereyra, J. S. (2019). Self-selection in school choice. *Games and Economic Behavior*, **117**, 59–81.

Chen, R., Katuščák, P., Kittsteiner, T. and Kütter, K. (2024). Does disappointment aversion explain non-truthful reporting in strategy-proof mechanisms? *Experimental Economics*, **27** (5), 1184–1210.

Chen, Y. and He, Y. (2021). Information acquisition and provision in school choice: an experimental study. *Journal of Economic Theory*, **197**, 105345.

— and — (2022). Information acquisition and provision in school choice: a theoretical investigation. *Economic Theory*, **74** (1), 293–327.

— and Sönmez, T. (2006). School choice: an experimental study. *Journal of Economic theory*, **127** (1), 202–231.

Chernev, A., Böckenholt, U. and Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, **25** (2), 333–358.

Compte, O. and Jehiel, P. (2004). The wait-and-see option in ascending price auctions. *Journal of the European Economic Association*, **2** (2-3), 494–503.

Dreyfuss, B., Glicksohn, O., Heffetz, O. and Romm, A. (2022). *Deferred acceptance with news utility*. Tech. rep., National Bureau of Economic Research.

Dur, U., Hammond, R. G. and Kesten, O. (2021). Sequential school choice: Theory and evidence from the field and lab. *Journal of Economic Theory*, **198**, 105344.

Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, **69** (1), 9–15.

Gigerenzer, G. and Kahneman, D. (2008). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press.

Gonczarowski, Y. A., Heffetz, O., Ishai, G. and Thomas, C. (2024). *Describing Deferred Acceptance and Strategyproofness to Participants: Experimental Analysis*. Tech. rep., National Bureau of Economic Research.

—, — and Thomas, C. (2023). *Strategyproofness-exposing mechanism descriptions*. Tech. rep., National Bureau of Economic Research.

Gong, B. and Liang, Y. (2024). A dynamic matching mechanism for college admissions: Theory and experiment. *Management Science*.

Greenberg, K., Pathak, P. A. and Sönmez, T. (2024). Redesigning the us army's branching process: A case study in minimalist market design. *American Economic Review*, **114** (4), 1070–1106.

Grenet, J., He, Y. and Kübler, D. (2022). Preference discovery in university admissions: The case for dynamic multioffer mechanisms. *Journal of Political Economy*, **130** (6), 1427–1476.

Guillen, P. and Hakimov, R. (2018). The effectiveness of top-down advice in strategy-proof mechanisms: A field experiment. *European Economic Review*, **101**, 505–511.

— and Veszteg, R. F. (2021). Strategy-proofness in experimental matching markets. *Experimental Economics*, **24**, 650–668.

Haeringer, G. and Iehlé, V. (2021). Gradual college admission. *Journal of Economic Theory*, **198**, 105378.

— and Klijn, F. (2009). Constrained school choice. *Journal of Economic Theory*, **144** (5), 1921–1947.

Hakimov, R., Heller, C.-P., Kübler, D. and Kurino, M. (2021). How to avoid black markets for appointments with online booking systems. *American Economic Review*, **111** (7), 2127–2151.

— and Kübler, D. (2021). Experiments on centralized school choice and college admissions: a survey. *Experimental Economics*, **24** (2), 434–488.

—, Kübler, D. and Pan, S. (2023). Costly information acquisition in centralized matching markets. *Quantitative Economics*, **14** (4), 1447–1490.

Hassidim, A., Romm, A. and Shorrer, R. I. (2021). The limits of incentives in economic matching procedures. *Management Science*, **67** (2), 951–963.

Hatfield, J. W., Kominers, S. D. and Westkamp, A. (2020). Stability, strategy-proofness, and cumulative offer mechanisms. *The Review of Economic Studies*, **88** (3), 1457–1502.

— and Milgrom, P. R. (2005). Matching with contracts. *American Economic Review*, **95** (4), 913–935.

Hu, X., Yao, L. and Zhang, J. (2025). Reforming china's two-stage college admission: An experimental study. *Available at SSRN 5207493.*

Huang, L. and Zhang, J. (2025). Bundled school choice. *arXiv preprint arXiv:2501.04241*.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, **93** (5), 1449–1475.

Katuščák, P. and Kittsteiner, T. (2024). Strategy-proofness made simpler. *Management Science.*

Klijn, F., Pais, J. and Vorsatz, M. (2019). Static versus dynamic deferred acceptance in school choice: Theory and experiment. *Games and Economic Behavior*, **113**, 147–163.

Kloosterman, A. and Troyan, P. (2023). Rankings-dependent preferences: A real goods matching experiment. *arXiv preprint arXiv:2305.03644*.

Kwasnica, A. M., Ledyard, J. O., Porter, D. and DeMartini, C. (2005). A new and improved design for multiobject iterative auctions. *Management Science*, **51** (3), 419–434.

Li, S. (2017). Obviously strategy-proof mechanisms. *American Economic Review*, **107** (11), 3257–3287.

Luflade, M. (2017). The value of information in centralized school choice systems (job market paper). *Job Market Paper, Duke University*.

Mackenzie, A. and Zhou, Y. (2022). Menu mechanisms. *Journal of Economic Theory*, **204**, 105511.

Meisner, V. (2023). Report-dependent utility and strategy-proofness. *Management Science*, **69** (5), 2733–2745.

— and Von Wangenheim, J. (2023). Loss aversion in strategy-proof school-choice mechanisms. *Journal of Economic Theory*, **207**, 105588.

Milgrom, P. (2009). Assignment messages and exchanges. *American Economic Journal: Microeconomics*, **1** (2), 95–113.

— (2011). Critical issues in the practice of market design. *Economic Inquiry*, **49** (2), 311–320.

Parkes, D. C. (2005). Auction design with costly preference elicitation. *Annals of Mathematics and Artificial Intelligence*, **44**, 269–302.

Pycia, M. and Troyan, P. (2023). A theory of simplicity in games and mechanism design. *Econometrica*, **91** (4), 1495–1526.

— and — (2024). The random priority mechanism is uniquely simple, efficient, and fair. *CEPR Discussion Papers*, (19689).

Rees-Jones, A. (2018). Suboptimal behavior in strategy-proof mechanisms: Evidence from the residency match. *Games and Economic Behavior*, **108**, 317–330.

— and Shorrer, R. (2023). Behavioral economics in education market design: A forward-looking review. *Journal of Political Economy Microeconomics*, **1** (3), 557–613.

— and Skowronek, S. (2018). An experimental investigation of preference misrepresentation in the residency match. *Proceedings of the National Academy of Sciences*, **115** (45), 11471–11476.

Roth, A. E. (2002). The economist as engineer: Game theory, experimentation, and computation as tools for design economics. *Econometrica*, **70** (4), 1341–1378.

—, Sönmez, T. and Ünver, M. U. (2004). Kidney exchange. *The Quarterly journal of economics*, **119** (2), 457–488.

Sandholm, T. and Boutilier, C. (2005). Preference elicitation in combinatorial auctions. In *Combinatorial Auctions*, The MIT Press.

Shorrer, R. I. and Sóvágó, S. (2023). Dominated choices in a strategically simple college admissions environment. *Journal of Political Economy Microeconomics*, **1** (4), 781–807.

— and Sóvágó, S. (2024). Dominated choices under deferred acceptance mechanism: The effect of admission selectivity. *Games and Economic Behavior*, **144**, 167–182.

Sönmez, T. (2013). Bidding for army career specialties: Improving the rotc branching mechanism. *Journal of Political Economy*, **121** (1), 186–219.

Sönmez, T. and Switzer, T. B. (2013). Matching with (branch-of-choice) contracts at the united states military academy. *Econometrica*, **81** (2), 451–488.

— and Ünver, M. U. (2010). Course bidding at business schools. *International Economic Review*, **51** (1), 99–123.

Soumalias, E., Jiang, Y., Zhu, K., Curry, M., Seuken, S. and Parkes, D. C. (2025). LLM-powered preference elicitation in combinatorial assignment. *arXiv preprint arXiv:2502.10308*.

Stephenson, D. (2022). Assignment feedback in school choice mechanisms. *Experimental Economics*, **25** (5), 1467–1491.

A Appendix: Supplementary Analysis

Table 6: Pairwise Comparison of Treatment Effects on Choice Accuracy (Preference Type: Overall)

	Comparison Treatment				
Base Treatment	SD-WEIGHT	SD-LEX	SD-DIRECT	SD-CHOICE	ACCURACY
SD-WEIGHT	—	-0.080**	0.238***	1.684***	1.203***
SD-LEX	0.080**	—	0.318***	1.764***	1.284***
SD-DIRECT	-0.238***	-0.318***	—	1.446***	0.966***
SD-CHOICE	-1.684***	-1.764***	-1.446***	—	-0.481***
ACCURACY	-1.203***	-1.284***	-0.966***	0.481***	—

Notes: Each cell reports the coefficient from a logistic regression where the row treatment is the base group and the column treatment is the comparison group. The upper triangle of the table is the inverse of the lower triangle (i.e., coefficient signs are flipped).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Pairwise Comparison of Treatment Effects on Choice Accuracy (Preference Type: LEX)

	Comparison Treatment				
Base Treatment	SD-WEIGHT	SD-LEX	SD-DIRECT	SD-CHOICE	ACCURACY
SD-WEIGHT	—	0.333***	0.266**	2.093***	2.315***
SD-LEX	-0.333***	—	-0.068	1.760***	1.981***
SD-DIRECT	-0.266**	0.068	—	1.827***	2.049***
SD-CHOICE	-2.093***	-1.760***	-1.827***	—	0.221
ACCURACY	-2.315***	-1.981***	-2.049***	-0.221	—

Notes: Each cell reports the coefficient from a logistic regression where the row treatment is the base group and the column treatment is the comparison group. The upper triangle of the table is the inverse of the lower triangle (i.e., coefficient signs are flipped).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Pairwise Comparison of Treatment Effects on Choice Accuracy (Preference Type: SEP)

	Comparison Treatment				
Base Treatment	SD-WEIGHT	SD-LEX	SD-DIRECT	SD-CHOICE	ACCURACY
SD-WEIGHT	—	-0.264***	0.111	1.916***	0.768***
SD-LEX	0.264***	—	0.375***	2.180***	1.032***
SD-DIRECT	-0.111	-0.375***	—	1.805***	0.657***
SD-CHOICE	-1.916***	-2.180***	-1.805***	—	-1.148***
ACCURACY	-0.768***	-1.032***	-0.657***	1.148***	—

Notes: Each cell reports the coefficient from a logistic regression where the row treatment is the base group and the column treatment is the comparison group. The upper triangle of the table is the inverse of the lower triangle (i.e., coefficient signs are flipped).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

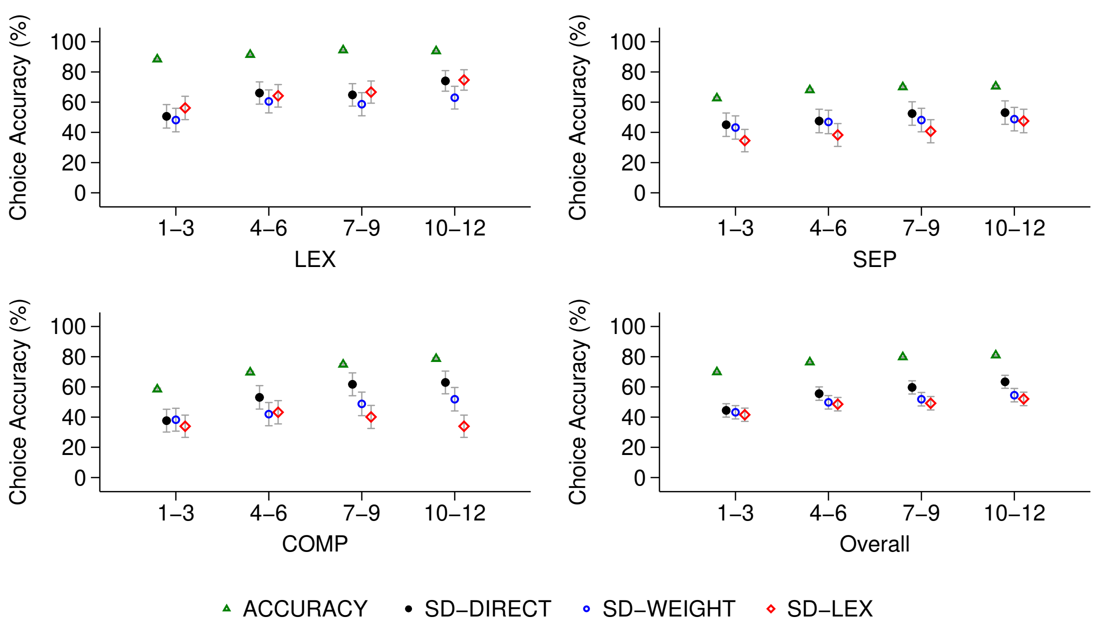
Table 9: Pairwise Comparison of Treatment Effects on Choice Accuracy (Preference Type: COMP)

	Comparison Treatment				
Base Treatment	SD-WEIGHT	SD-LEX	SD-DIRECT	SD-CHOICE	ACCURACY
SD-WEIGHT	—	-0.306***	0.347***	1.315***	1.124***
SD-LEX	0.306***	—	0.652***	1.621***	1.430***
SD-DIRECT	-0.347***	-0.652***	—	0.969***	0.778***
SD-CHOICE	-1.315***	-1.621***	-0.969***	—	-0.191
ACCURACY	-1.124***	-1.430***	-0.778***	0.191	—

Notes: Each cell reports the coefficient from a logistic regression where the row treatment is the base group and the column treatment is the comparison group. The upper triangle of the table is the inverse of the lower triangle (i.e., coefficient signs are flipped).

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure 3: Choice Accuracy (CI: 95%)



Notes: X-axis reports rounds. Recall that the experiment design is such that each preference domain is realized exactly once in every three rounds.

Figure 4: Kendall Distance (CI: 95%)

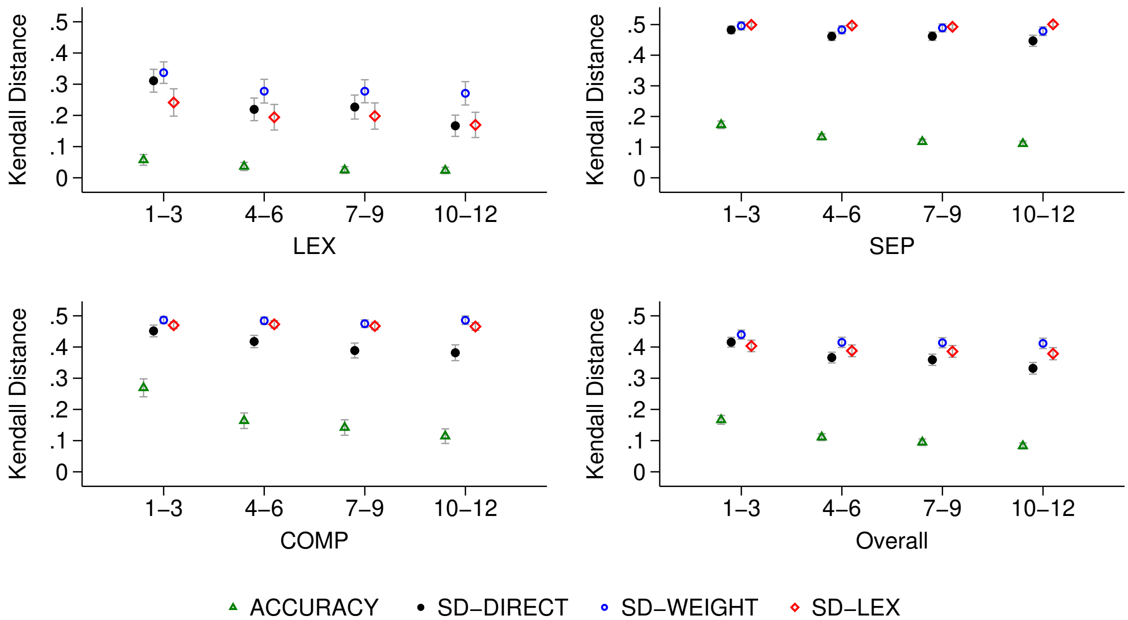


Figure 5: Efficiency Loss (CI: 95%)

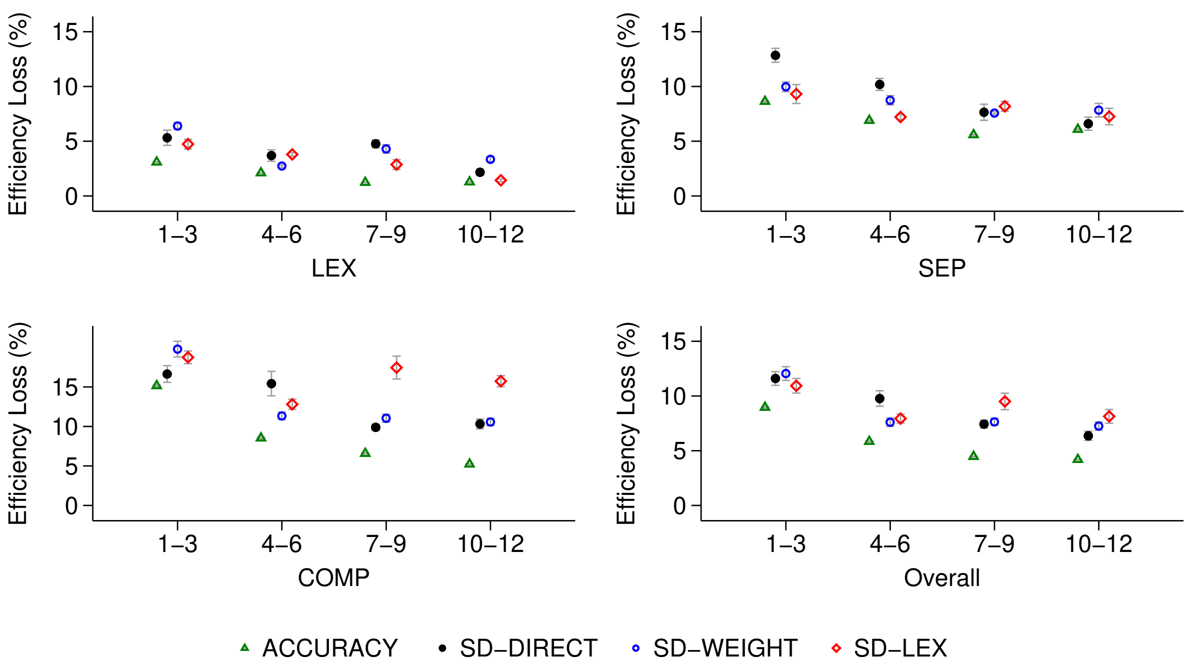
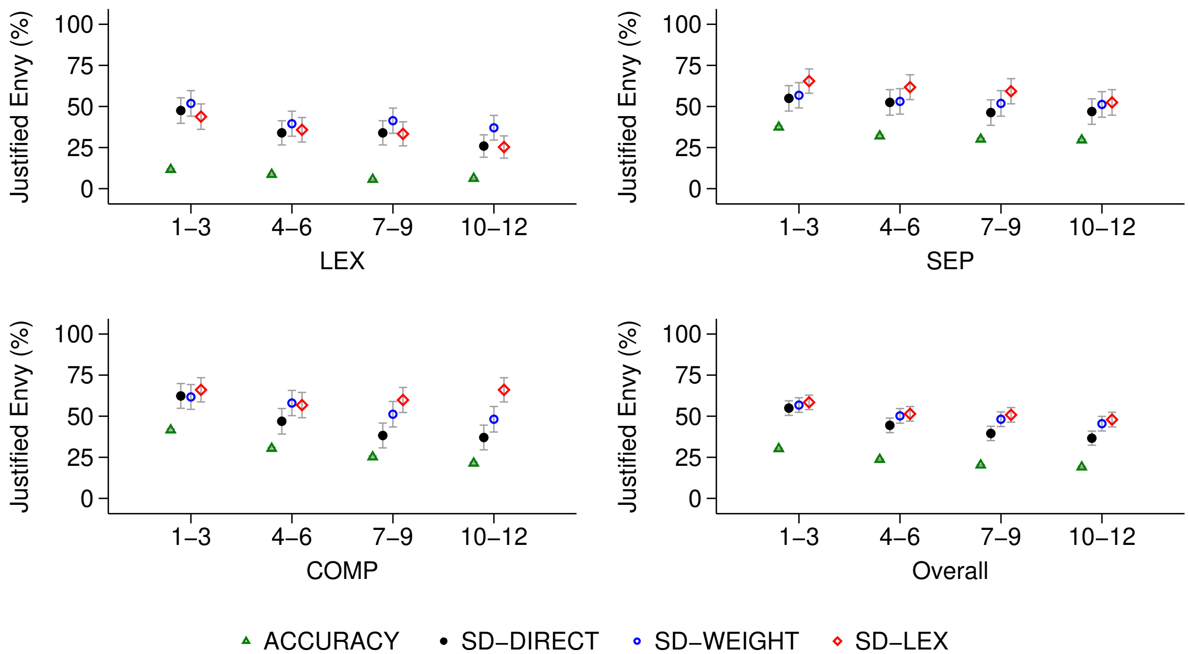


Figure 6: Justified Envy (CI: 95%)



B Appendix: Indirect Message Spaces

B.1 Rank-order lists in College Admissions

23 out of 31 provinces in China implement the structured rank-order list system, in which majors are effectively nested under colleges, as noted by Hu *et al.* (2025). These provinces include: Shanghai, Beijing, Tianjin, Hainan, Jiangsu, Fujian, Hubei, Hunan, Guangdong, Heilongjiang, Gansu, Jilin, Anhui, Jiangxi, Guangxi, Shanxi, Henan, Shaanxi, Ningxia, Sichuan, Yunnan, Tibet, and Xinjiang. For illustrative purposes, we include screenshots of the official college-major preference form from Fujian and Shanghai.

2025 年福建省普通高校招生考生志愿填报

计划科类：普通类（物理科目组）　　　　批次名称：普通类本科批常规志愿

开始时间：2025-07-03 08:00:00 —— 结束时间：2025-07-07 18:00:00

考生基本信息			
姓　名		**考 生 号**	
性　别		**身份证号**	
选考科目			

01 志愿				02 志愿			
院校代号		**院校名称**		**院校代号**		**院校名称**	
专业组代号		**专业组名称**		**专业组代号**		**专业组名称**	
专业代号 1		**专业名称 1**		**专业代号 1**		**专业名称 1**	
专业代号 2		**专业名称 2**		**专业代号 2**		**专业名称 2**	
专业代号 3		**专业名称 3**		**专业代号 3**		**专业名称 3**	
专业代号 4		**专业名称 4**		**专业代号 4**		**专业名称 4**	
专业代号 5		**专业名称 5**		**专业代号 5**		**专业名称 5**	
专业代号 6		**专业名称 6**		**专业代号 6**		**专业名称 6**	
是否服从专业调剂	是　　否			**是否服从专业调剂**	是　　否		

03 志愿				04 志愿			
院校代号		**院校名称**		**院校代号**		**院校名称**	
专业组代号		**专业组名称**		**专业组代号**		**专业组名称**	
专业代号 1		**专业名称 1**		**专业代号 1**		**专业名称 1**	
专业代号 2		**专业名称 2**		**专业代号 2**		**专业名称 2**	
专业代号 3		**专业名称 3**		**专业代号 3**		**专业名称 3**	
专业代号 4		**专业名称 4**		**专业代号 4**		**专业名称 4**	
专业代号 5		**专业名称 5**		**专业代号 5**		**专业名称 5**	
专业代号 6		**专业名称 6**		**专业代号 6**		**专业名称 6**	
是否服从专业调剂	是　　否			**是否服从专业调剂**	是　　否		

Source: `https://www.eeafj.cn/gkptgkgsgg/20250626/14073.html`

2023 年上海市普通高等学校招生考生志愿表
表 4 – 本科普通批次（样表）

高考报名号				姓名		性别		所在报名区		选考科目		应试语种	
批　次		院校专业组代码	院校专业组名称	专业代码	专业名称	专业代码	专业名称	专业代码	专业名称	专业代码	专业名称	愿否专业调剂	
	1												
	2												
	3												
	4												
	5												
	6												
	7												
	8												
	9												
	10												
	11												
	12												
本科普通批次	13												
	14												
	15												
	16												
	17												
	18												
	19												
	20												
	21												
	22												
	23												
	24												

注：1. 考生在本科普通批次最多可填报 24 个院校专业组志愿，每个院校专业组内最多可填报 4 个专业志愿。
2. 本科普通批次实行平行志愿。投档时，按考生高考成绩总分从高到低排序，逐分、逐个地按其填报的院校专业组志愿顺序依次检索，一旦进档则不再继续检索，实行一次投档。
3. 考生在每个院校专业组志愿中均须填写“愿否专业调剂”代码。其中“1”表示全愿意，“2”表示全不愿意，“3”表示除高收费专业外其他愿意，“4”表示除医科专业外其他愿意。考生在该栏必须填写“1”到“4”之间的数字，其中“3”“4”可以并列填写。专业调剂只能在所投档的院校专业组内开展。
4. 考生须在志愿填报系统中按规定输入志愿表信息。志愿表须一式三份打印，并由考生本人签字(其他人签字无效)，作为投档录取依据。

考生签名：　　　　　日期：　　年　月　日

Source: `https://www.shmeea.edu.cn/page/08000/20230407/17353.html`

B.2 Reserve Officer Training Corps (ROTC) Mechanism

Sönmez (2013)'s model of cadet-branch matching problem consists of

1. a finite set of cadets $I = \{i_1, i_2, \ldots, i_n\}$,
2. a finite set of branches $B = \{b_1, b_2, \ldots, b_m\}$,
3. a vector of branch capacities $q = (q_b)_{b \in B}$,
4. a set of "terms" $T = \{t_1, \ldots, t_k\}$,
5. a list of cadet preferences $P = (P_i)_{i \in I}$ over $(B \times T) \cup \{\emptyset\}$, and
6. a list of base priority rankings $\pi = (\pi_b)_{b \in B}$.

The ROTC mechanism is not direct. Instead, each cadet submits a ranking of branches $\succ'_i$, and he can sign a branch-of-choice contract for any of his top three choices under $\succ'_i$.

Online Appendix C

Complexity Beyond Incentives:

The Critical Role of Reporting Language

By Rustamdjan Hakimov and Manshu Khanna

This online appendix contains the experiment instructions for three of the five treatments discussed in the article. The instructions for the remaining two treatments follow directly from them. All original instructions were in Chinese. Contact the authors for the full set of instructions.

[Treatment: SD-DIRECT]

Welcome! This is an experiment in the economics of decision making. If you follow the instructions carefully, you may earn a considerable amount of money. These instructions are identical for every participant. Please turn off your electronic devices. Do not communicate with each other or ask questions aloud during the experiment. If you have questions at any point, raise your hand and we will come to you and answer them.

Overview

In this experiment, we simulate an environment where students are allocated to seats in university programs.

- All of you will be making decisions as students.
- All participants in this experiment are divided into groups of 27. Your group stays the same throughout the entire experiment. You will be competing with the other 26 students in your group for seats in 27 university programs.
- The 27 university seats are all different, and each university seat is characterized by three features: university, field of study and tuition. There are three universities, three fields of study, and three levels of tuition. Each university offers one seat for each field of study at each tuition level. Thus, the three scenarios each for university, field and tuition make up the $3^3 = 27$ university seats available.
- The process of allocating the 27 students in the same group to each university seat is based on their exam marks and submission decisions, which is detailed below.
- The experiment consists of twelve independent rounds. Each round represents a new admission process. Your final payoff is determined by the allocation outcome of a randomly selected round at the end of the experiment.

Exam Marks

- All universities admit students based on their marks in an admission exam.
- For each round, the mark of each student is drawn independently and randomly from the set $\{1, 2, 3, \ldots, 100\}$. Each number is equally likely to be drawn, and higher numbers mean better marks.
- The computer will avoid ties when drawing marks. That is, each of the 27 students in a group will have a different mark.
- You will learn your own mark but not the marks of other students.

<u>Your Preferences over University Seats</u>

- You can obtain a higher payoff if you are assigned a seat at a university you prefer more.
- As shown in the table below, your payoff equals CNY160, CNY155, CNY150…CNY35, CNY30 if you hold a university seat ranked 1st, 2nd, 3rd….26th, 27th in your preferences respectively.

Your Allocation Outcome		**Your Payoff**	
If you hold a university seat ranked:	1st Preference	**Then you will get:**	CNY160
	2nd Preference		CNY155
	3rd Preference		CNY150
	4th Preference		CNY145
	5th Preference		CNY140
	6th Preference		CNY135
	7th Preference		CNY130
	8th Preference		CNY125
	9th Preference		CNY120
	10th Preference		CNY115
	11th Preference		CNY110
	12th Preference		CNY105
	13th Preference		CNY110
	14th Preference		CNY95
	15th Preference		CNY90
	16th Preference		CNY85
	17th Preference		CNY80
	18th Preference		CNY75
	19th Preference		CNY70
	20th Preference		CNY65
	21th Preference		CNY60
	22th Preference		CNY55
	23th Preference		CNY50
	24th Preference		CNY45
	25th Preference		CNY40
	26th Preference		CNY35
	27th Preference		CNY30

- Your preferences over university seats are determined by three elements: university prestige, field fit, and tuition. Your **preference score** for each university seat can be obtained based on your personal preference score formula. For each round, you will learn a new personal preference score formula. The seat with the highest score is your 1st preference, the seat with the second highest score is your 2nd preference, and so on.

- For instance, your preference score formula could look like that:

$$Score = 25 * University\ prestige + 40 * Field\ fit - 10 * Tuition$$

- For each round, every student has a **different** formula that translates university prestige, field fit and tuition into preference scores.
  - Every student has the **same** university prestige and tuition.
  - Students may have **different** field fits.
  - Thus, students' preference ranks over each university seats might be different.
- You will know university prestige, your field fit, and tuition for each university seat.
- You will have a calculator during each round to help you obtain your own preference scores for each university seat.

Your Submission Decisions

Before the allocation procedure, you will make a submission decision which is a ranking list. Other participants cannot observe your decision.

University Listing

- You will be asked to reveal your preferences over the university seats. You will be asked to rank all 27 combinations of the university, field and the tuition by listing them as your 1st, 2nd, 3rd, 4th, 5th, …. 27th choices. You can rank all 27 options or stop at any point.

- These submission rankings will be directly used in the allocation procedure.

Allocation Procedure

In each round, an allocation procedure will be used to allocate students to university seats. The outcome of an allocation procedure depends on:

1) the rankings of university seats submitted by you and the other 26 students in your group;
2) the admission exam marks of you and the other 26 students.

Specifically, the allocation procedure follows the steps below (all the steps take place without any interactions with the other students):

Step 1. The student with the **highest mark** in the exam is assigned a university seat at

- the 1st choice in her rankings submitted.

Step 2. Consider the student with the **second-highest mark**. She is assigned a university seat at

- her 1st choice if her 1st choice is vacant;

- her <u>2nd choice</u> if her 1st choice has been filled.

Step 3. Consider the student with the **third-highest mark**. She is assigned a university seat at

- her 1st choice if her 1st choice is vacant;
- her 2nd choice if her 1st choice has been filled, but 2nd choice is vacant;
- her 3rd choice if her1st and 2nd choices have been filled.

…

Step 4-27. Consider the student with the **next-highest mark**. She is assigned a university seat at her best choice in her rankings submitted among the vacant options. **If all choices from her rankings submitted are full, the student remains unassigned.**

To summarize, a student with a higher exam mark is considered earlier in this allocation procedure. In each step, the student being considered will be admitted to the university seat at her best choice that has not been occupied by other students with higher exam marks. Note that at any step, if **all choices from the rankings submitted are full, the student remains unassigned, and the corresponding payoff is CNY0.**

For example, suppose your exam mark is 96, second only to the highest score of 99, which means you have the second priority for the implementation of your preference submission. The first university seat in your preference submission is [University A, Economics, No Tuition]. The second university seat is [University A, Finance, No Tuition]. In this case, if the student with 99 exam marks also ranks [University A, Economics, No Tuition] as the first university seat in her preference submission, then your final allocation outcome will be [University A, Finance, No Tuition]. If the student with 99 exam marks ranks another university seat as the first choice in her preference submission, then your final allocation outcome will be [University A, Economics, No Tuition].

Summary

- The allocation procedure allocates 27 students to 27 university seats. You are one of the students.
- At the beginning of every round, you learn the formula that translated university prestige, field fit and tuition into the preference scores. The higher the score of the university seat the higher is the ranking in your preferences, and thus the higher the payoff holding the seat can bring to you.
- The rankings submitted by you and the other 26 students, together with your exam marks, will determine the outcome of the allocation procedure.

Explanation for the screenshot

Below is a screenshot of this experiment where you can learn your preferences for each university seat and submit your preferences to the allocation system. First you can learn the information about university prestige, your field fit and tuition from the box in the top left corner. Note that only the value of field fit may vary among the 27 students, the parameters of university prestige and tuition is the same for every student. Second, the top right corner shows your personal preference score formula, and you can use the calculator in the middle to obtain your preference scores for each university seat.

Finally, you can submit your preference-revealing decision in the box at the bottom which is a ranking list of 27 university seats.

[Treatment: SD-CHOICE]

Welcome! This is an experiment in the economics of decision making. If you follow the instructions carefully, you may earn a considerable amount of money. These instructions are identical for every participant. Please turn off your electronic devices. Do not communicate with each other or ask questions aloud during the experiment. If you have questions at any point, raise your hand and we will come to you and answer them.

Overview

In this experiment, we simulate an environment where students are allocated seats in university programs.

- All of you will be making decisions as students.
- The experiment consists of 12 independent admission processes. You will make one choice for each admission process, resulting in 12 choices throughout the experiment. There will be 27 rounds in total. In each round, 12 students will make their choices, each corresponding to a distinct admission process, while the remaining 15 students will wait until the choices are made. In other words, if you are selected in a given round, you will make your choice in that round for a particular admission process. If you are not selected, you will simply wait.
- There are 27 university seats for allocation in each admission process. The 27 university seats are all different, and each university seat is characterized by three features: university, field of study and tuition. There are three universities, three fields of study, and three levels of tuition. Each university offers one seat for each field of study at each tuition level. Thus, the three scenarios each for university, field and tuition make up the $3^3 = 27$ university seats available.
- After the final round (i.e., the 27th round), the computer will allocate university seats based on students' choices in each independent admission process.
- The process of allocating the 27 students to each university seat in each admission process is based on their priority order and submitted choices.
- Your final payoff will be determined by the allocation outcome of a randomly selected admission process at the end of the experiment.
- In each admission process, you will have 3 minutes to make your choice, and the remaining time will be displayed on the decision screen. Please note that if you do not submit your decision within 3 minutes, the university seat you selected on the decision page will be automatically submitted. If you do not select any university seat, you will not be assigned to any university seat.

Priority Order

- All universities admit students based on their priority order.
- The priority order will be represented by priority numbers. For example, if you are selected to make your choice in the first round, your priority order will be represented by **priority 1**, meaning you rank the first to make your choice in a specific admission process. If you are not selected, you will simply wait. Similarly, if you are selected in the second round, you will have **priority 2**, also making your choice in a specific admission process. This process continues for all 27 rounds. If you are selected in the 27^{th} round, you will have **priority 27**, meaning you will make your choice in the final round.
- Each student will randomly receive a unique priority number in each admission process. You will know your own priority numbers but not those of other students.

Your Preferences over University Seats

- You can obtain a higher payoff if you are assigned a seat at a university you prefer more.
- As shown in the table below, your payoff equals CNY160, CNY155, CNY150…CNY35, CNY30 if you hold a university seat ranked 1^{st}, 2^{nd}, 3^{rd}….26^{th}, 27^{th} in your preferences respectively.

Your Allocation Outcome		**Your Payoff**	
If you hold a university seat ranked:	1^{st} Preference	**Then you will get:**	CNY160
	2^{nd} Preference		CNY155
	3^{rd} Preference		CNY150
	4^{th} Preference		CNY145
	5^{th} Preference		CNY140
	6^{th} Preference		CNY135
	7^{th} Preference		CNY130
	8^{th} Preference		CNY125
	9^{th} Preference		CNY120
	10^{th} Preference		CNY115
	11^{th} Preference		CNY110
	12^{th} Preference		CNY105
	13^{th} Preference		CNY110
	14^{th} Preference		CNY95
	15^{th} Preference		CNY90
	16^{th} Preference		CNY85
	17^{th} Preference		CNY80
	18^{th} Preference		CNY75
	19^{th} Preference		CNY70

	20th Preference		CNY65
	21th Preference		CNY60
	22th Preference		CNY55
	23th Preference		CNY50
	24th Preference		CNY45
	25th Preference		CNY40
	26th Preference		CNY35
	27th Preference		CNY30

- Your preferences over university seats in each admission process are determined by three elements: university prestige, field fit, and tuition. Your **preference score** for each university seat can be obtained based on your personal preference score formula. For each admission process, you will learn a new personal preference score formula. The seat with the highest score is your 1st preference, the seat with the second highest score is your 2nd preference, and so on.
- For instance, your preference score formula could look like that:

$$Score = 25 * University\ prestige + 40 * Field\ fit - 10 * Tuition$$

- For each admission process, every student has a **different** formula that translates university prestige, field fit and tuition into preference scores.
  - Every student has the **same** university prestige and tuition.
  - Students may have **different** field fits.
  - Thus, students' preference ranks over each university seats might be different.
- You will know university prestige, your field fit, and tuition for each university seat during each admission process.
- You will have a calculator during each admission process to help you obtain your own preference scores for each university seat.

Your Submission Decisions

- You will be asked to choose your most preferred university seat **among the available university seats** in each admission process.

Allocation Procedure

At the end of the experiment, a sequential allocation procedure will be used to allocate students to university seats in each admission process. The outcome of an allocation procedure for each admission process depends on:

3) the choices submitted by you and the other 26 students.

4) the sequence in which students make their decisions, referred to as the priority order.

Specifically, there are 27 university seats available. Students will choose university seats directly when it is their turn to choose. The allocation procedure follows the steps below (all the steps take place without any interactions with the other students):

Step 1. The student with **priority 1** gets to choose in the 1st round:

- her most preferred university seat among all 27 seats.

Step 2. The student with **priority 2** gets to choose in the 2nd round:

- her most preferred university seat among the remaining 26 seats.

Step 3. The student with **priority 3** gets to choose in the 3rd round:

- her most preferred university seat among the remaining 25 seats.

…

Step 4-26. Students with **priority 4 through 26** gets to choose in the 4th to the 26th rounds:

- their most preferred university seat among the remaining available seats, which will reduce from 24 down to 2 as the process progresses.

Step 27. The student with **priority 27** is assigned in the 27th round:

- the last remaining university seat.

To summarize, students with higher priority are considered earlier in the procedure. In each step, the student being considered will be admitted to her chosen university seat.

For example, suppose you are chosen to make a choice in a specific admission process in the first round, then your priority order will be **priority 1**, which means you have the first priority to choose your most preferred university seat among all 27 seats. The university seat you choose is [University A, Economics, No Tuition]. Then your final allocation outcome in this admission process will be [University A, Economics, No Tuition]. If you are chosen to make a choice in a specific admission process in the 26th round, then your priority order will be **priority 26**. The last two remaining seats are [University C, Economics, Full Tuition] and [University B, Finance, Full Tuition]. If you choose [University C, Economics, Full Tuition], then your final allocation outcome for that admission process will be [University C, Economics, Full Tuition].

Summary

- The experiment consists of 12 independent admission processes, you will make one choice for each admission process, resulting in 12 choices throughout the experiment.
- For each admission process, you will face a unique choice situation where you will learn the formula that translates university prestige, field fit and tuition into preference scores. The higher the score of the university seat, the higher is the ranking in your preferences, and thus the higher the payoff holding the seat can bring to you.

- The allocation procedure allocates 27 students, including you, to 27 university seats in each admission process.
- The choices submitted by you and the other 26 students in each admission process will determine the outcome of the allocation procedure at the end of the experiment.

Explanation for the screenshot

Below is a screenshot of this experiment where you can learn your preferences for each university seat and submit your preferences to the allocation system. In each admission process, you will have at most 3 minutes to submit your choice, and the decision screen will display your remaining time. First you can learn the information about university prestige, your field fit and tuition from the box in the top left corner. Note that only the value of field fit may vary among the 27 students in each admission process, the parameters of university prestige and tuition is the same for every student. Second, the top right corner shows your personal preference score formula, and you can use the calculator in the middle to obtain your preference scores for each university seat. Third, you can see a list of university seats available for you to choose. Finally, you can submit your preference decision in the box at the bottom, where you choose your most preferred university seat from the available options. Afterward, you can click the submit button at the bottom right to complete your decision for this round.

[Treatment: ACCURACY]

Welcome! This is an experiment in the economics of decision making. If you follow the instructions carefully, you may earn a considerable amount of money. These instructions are identical for every participant. Please turn off your electronic devices. Do not communicate with each other or ask questions aloud during the experiment. If you have questions at any point, raise your hand and we will come to you and answer them.

Overview

In this experiment, you are required to report your preferences over 27 university seats as accurately as possible.

- All participants will be independently making decisions as students.
- The 27 university seats are all different, and each university seat is characterized by three features: university, field of study and tuition. There are three universities, three fields of study, and three levels of tuition. Each university offers one seat for each field of study at each tuition level. Thus, the three scenarios each for university, field and tuition make up the $3^3 = 27$ university seats available.
- The experiment consists of twelve independent rounds. In each round, you will have a 8-minute countdown to submit a **ranking list** including the 27 university seats.
- Your payoff will depend on how accurately your **submitted ranking list** matches your **true preference ranking list**. Your true preference ranking list will be determined based on your personal preference score formula (as explained in detail below).
- The accuracy will be assessed using the **Kendall Score**, which measures how closely your submitted ranking list aligns with your true preference ranking list. After each round, you will learn your Kendall Score, ranging from 0 to 1. A lower Kendall Score indicates a more accurate match between your submitted ranking list and your true preference ranking list.
- At the end of the experiment, one of the 12 rounds will be randomly selected to determine your payment. Your final payment will be calculated as the following formula: **Final Payoff = CNY 160 × (1 - Kendall Score)**, with a maximum possible payoff of 160 and a minimum of 0.

Your Preferences over University Seats

- You can obtain a higher payoff if you report your preferences over the 27 university seats more accurately.
- Your preferences over university seats are determined by three elements: university prestige, field fit, and tuition. Your **preference score** for each university seat can be obtained based on your personal preference score formula. For each round, you will learn a new personal

preference score formula. The seat with the highest score is your 1st preference, the seat with the second highest score is your 2nd preference, and so on. The ranking list generated according to your personal preference score formula will be considered your **true preference ranking list**.

- For instance, your preference score formula could look like that:

$$Score = 25 * University\ prestige + 40 * Field\ fit - 10 * Tuition$$

- For each round, each student has a **different** <u>formula</u> that translates university prestige, field fit and tuition into preference scores.
  - Each student has the **same** <u>university prestige</u> and <u>tuition</u>.
  - Students may have **different** <u>field fits</u>.
  - Thus, students' preference ranks over each university seats might be different.
- During each round, you will know university prestige, your field fit, and tuition for each university seat.
- During each round, you will have a calculator to help you obtain your own preference scores for each university seat.

<u>Your Submission Decisions</u>

In each round, you will submit a **ranking list** including the 27 university seats. Other participants cannot observe your decision.

University Listing

- You will be asked to reveal your preferences over the university seats by ranking all 27 combinations of the university, field and tuition, <u>listing them as your 1st, 2nd, 3rd, 4th, 5th, …. 27th choices</u>.

- These submitted ranking lists will be directly compared with your true preference ranking lists to calculate your Kendall Score for each round. After each round, you will know your Kendall Score, which reflects how closely your submitted ranking list matches your true preference ranking list based on your preference scores.

- Your Kendall Score will range from 0 to 1, rounded to three decimal places. A lower score indicates a closer match between your submitted and true preference ranking lists. For instance, a score of 0 indicates a perfect match, while a score of 1 indicates that the two ranking lists are exactly opposite.
- At the end of the experiment, one of the 12 rounds will be randomly selected to determine your payment. Your final payment will be calculated as the following formula: **Final Payoff**

= CNY 160 × (1 - Kendall Score), with a maximum possible payoff of 160 and a minimum of 0.

For example, if the 3rd round is randomly selected for payment and your Kendall Score in that round is 0.70, your payoff will be CNY 160 × (1 - 0.70) = CNY 48.

Summary

- The experiment consists of 12 independent rounds. In each round, you will submit a ranking list to show your preferences over 27 university seats. Your payoff will be determined by the accuracy of your submitted ranking list compared to your true preference ranking list, based on your personal preference score formula in each round.
- The accuracy will be measured using the Kendall Score, which shows how closely your submitted ranking matches with your true ranking. A lower score indicates a closer match between the two ranking lists.
- In each round, you will be provided with different preference score formulas that translated university prestige, field fit and tuition into the preference scores. The higher the score of a university seat, the higher it ranks in your true preference ranking list.
- After each round, you will learn your Kendall Score for that round, reflecting the accuracy of your submitted ranking list. Your final payoff will be determined by the Kendall Score from one randomly selected round at the end of the experiment.

Explanation for the screenshot

Below is a screenshot of this experiment where you can learn your preferences for each university seat and submit your ranking list to the system. In each round, you will have at most 8 minutes to submit, and the decision screen will display your remaining time. First, you can view information about university prestige, your field fit and tuition in the box located in the top left corner. Note that only the value of field fit may vary in each round, while the parameters of university prestige and tuition remain the same for all students. Second, the top right corner shows your personal preference score formula. You can use the calculator in the middle to obtain your preference scores for each university seat. Finally, you can submit your ranking list in the box at the bottom.

实验开始: 第 1 轮

剩余时间：6 分钟 05 秒

以下三个列表列出了 **实验的参数**：

大学声望	**专业匹配度**	**学费**
A大学 - 500	经济学 - 300	全额学费 - 500
B大学 - 200	金融学 - 500	半额学费 - 250
C大学 - 600	法学 - 700	免学费 - 0

您的个人偏好分数由以下公式决定:

偏好分数 = 33.2 * 大学声望 + 35.6 * 专业匹配度 - 39.6 * 学费 - 0.5 * 大学声望 * 学费

你可以在这里计算你可能的偏好分数。

计算器: 33.2 * 大学声望 + 35.6 * 专业匹配度 - 39.6 * 学费 - 0.5 * 大学声望 * 学费 = 偏好分数

33.2 * ____ + 35.6 * ____ - 39.6 * ____ - 0.5 * ____ * ____ = ____

计算 重置

偏好提交板块

Rank1: ----	Rank2: ----	Rank3: ----
Rank4: ----	Rank5: ----	Rank6: ----
Rank7: ----	Rank8: ----	Rank9: ----
Rank10: ----	Rank11: ----	Rank12: ----
Rank13: ----	Rank14: ----	Rank15: ----
Rank16: ----	Rank17: ----	Rank18: ----
Rank19: ----	Rank20: ----	Rank21: ----
Rank22: ----	Rank23: ----	Rank24: ----
Rank25: ----	Rank26: ----	Rank27: ----

提交

[Treatment: SD-WEIGHT]

The instructions for the SD-WEIGHT treatment mirror those of the SD-DIRECT treatment, except for the "Your Submission Decisions" section and the "Explanation for the Screenshot" section, which are as follows:

Your Submission Decisions

Before the allocation procedure, you will make a submission decision consisting of three ranking lists and three importance weights. Other participants cannot observe your decision.

Preference submission of universities, fields, and tuition

- First, you will be asked to create three ranking lists for university, field fit, and tuition separately, to reveal your preferences over the three universities, three fields and three modes of tuition.
- Additionally, you will be asked to assign a weight to each of the ranking lists. The weight value should be a number between 1 and 100.

The computer will construct your ranking over 27 university seats based on your submitted ranking lists and importance weights of university, field fit, and tuition, calculating a point for each university seat. The formula used is:

$$\begin{aligned} Points \ = \ & (Rank\ of\ University \ \times \ Weight\ of\ University) \\ & + \ (Rank\ of\ Field \ \times \ Weight\ of\ Field) \\ & + \ (Rank\ of\ Tuition \ \times \ Weight\ of\ Tuition). \end{aligned}$$

A lower point means a higher ranking for that university seat, i.e. the university seat with the lowest points will be the 1st choice in the ranking, while the university seat with the most points will be the last choice in the ranking. In cases where points are equal, the tie will be broken randomly. Therefore, for university, field and tuition, the higher the weight you assign to, the more the element influences the final submission ranking.

For example, if University A, Economics, and No Tuition are respectively ranked the first in your university ranking list, and Finance ranked the second in your field fit list, with importance weights assigned as 50 for university, 20 for field fit, and 70 for tuition, then the points for University A, Economics, and No tuition would be $1 \times 50 + 1 \times 20 + 1 \times 70 = 140$, while the points for University A, Finance, and No tuition would be $1 \times 50 + 2 \times 20 + 1 \times 70 = 160$. Remember, the higher the points, the lower the final submission ranking of the university seat. These submission rankings will be directly used in the allocation procedure.

Explanation for the screenshot

Below is a screenshot of this experiment where you can learn your preferences for each university seat and submit your preferences to the allocation system. First you can learn the information about university prestige, your field fit and tuition from the box in the top left corner. Note that only the value of field fit may vary among the 27 students, the parameters of university prestige and tuition is the same for every student. Second, the top right corner shows your personal preference score formula, and you can use the calculator in the middle to obtain your preference scores for each university seat. Finally, you can submit your preference-revealing decision in the box at the bottom, consisting of three ranking lists and importance weights of university prestige, field fit and tuition respectively.

[Treatment: SD-LEX]

The instructions for the SD-LEX treatment mirror those of the SD-DIRECT treatment, except for the "Your Submission Decisions" section and the "Explanation for the Screenshot" section, which are as follows:

Your Submission Decisions

Before the allocation procedure, you will make a submission decision consisting of three ranking lists. Other participants cannot observe your decision.

Preference submission of universities, fields, and tuition

- You will be asked to create three ranking lists for university, field fit, and tuition separately, to reveal your preferences over the three universities, three fields and three modes of tuition.

The computer will construct your ranking over 27 university seats using this method: First, it considers your university ranking list. All seats at each university, regardless of field or tuition, are ranked collectively. Seats at the highest-ranked university are placed above those at the second-ranked university, and so on. Second, within each university, the computer ranks seats based on your field ranking list. Seats in the most preferred fields are ranked higher than those in less preferred fields. Finally, within each field, seats are ranked based on your tuition ranking list, with preferences for tuition determining their orders. The formula used is:

$$Points = (Rank\ of\ University \times 100) + (Rank\ of\ Field \times 10) + (Rank\ of\ Tuition \times 1).$$

A lower point means a higher ranking for that university seat, i.e. the university seat with the lowest points will be the 1st choice in the ranking, while the university seat with the most points will be the last choice in the ranking. For example, if University A, Economics, No Tuition are respectively ranked the first in your university ranking list, and Finance ranked the second in your field fit list, then the points for [University A, Economics, No tuition] would be $1 \times 100 + 1 \times 10 + 1 \times 1 = 111$, while the points for [University A, Finance, No tuition] would be $1 \times 100 + 2 \times 10 + 1 \times 1 = 121$. Remember, the higher the points, the lower the final submission ranking of the university seat. The resulting submission rankings of 27 university seats will be directly used in the allocation procedure.

Explanation for the screenshot

Below is a screenshot of this experiment where you can learn your preferences for each university seat and submit your preferences to the allocation system. First you can learn the information about university prestige, your field fit and tuition from the box in the top left corner. Note that only the value of field fit may vary among the 27 students, the parameters of university prestige and tuition is the same for every student. Second, the top right corner shows your personal preference score formula, and you can use the calculator in the middle to obtain your preference scores for each university seat.

Finally, you can submit your preference-revealing decision in the box at the bottom, consisting of three ranking lists of university prestige, field fit and tuition respectively.