ISSN (Print) 0473-453X
Discussion Paper No. 1148
ISSN (Online) 2435-0982

# BEHAVIORAL CHANGES IN DIFFERENT DESIGNS OF SEARCH EXPERIMENTS

Yuta Kittaka
Ryo Mikami
Natsumi Shimada

Revised June 2022
November 2021

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

# Behavioral Changes in Different Designs of Search Experiments*

Yuta Kittaka † Ryo Mikami‡ Natsumi Shimada§

June 9, 2022

**Abstract**

While search experiments are available in several designs, growing experimental evidence suggests that individual search behavior depends on design details. We conduct an experiment providing the first categorization and comparison of several search experiment designs widely accepted in search studies. These designs can be categorized as passive, quasi-active, and active, according to the degree of flexibility in decision-making regarding the search. Despite the experiment being based on an identical model, we found significant differences at the aggregate- and individual-level in the results across designs. The average number of searches was the highest and closest to the theoretical value in the active design. Compared with the active design, subjects searched significantly less in the quasi-active and passive designs. The results indicate that the widely accepted design, wherein subjects make decisions based on a given offer rather than choosing among potential alternatives themselves, may have unexpected effects on subjects' behavior. Furthermore, subjects' risk aversion has a significant effect only in the passive design, suggesting that out-of-model factors specific to that design may influence behavior through risk preferences. Other methodological implications for search experiments are also provided.

**Keywords:** search experiment, consumer search, labor search, risk preference experimental design, online experiment

**JEL classification:** C91, D12, D83, J60

* We thank Ben Casner, Nobuyuki Hanaki, Morimitsu Kurino, Takehito Masuda, Noriaki Matsushima, Takahiro Miura, and Masaru Sasaki for their valuable comments and suggestions. We also thank Syumei Hirai, Fumio Ohtake, Yuta Shimodaira, and the seminar and workshop participants at the Ministry of Economy, Trade and Industry, Shinshu University, the Japan Association for Applied Economics 2021 Autumn Meeting, the Applied Regional Science Conference 35th Annual Meeting, and the 24th Experimental Social Science Conference. We gratefully acknowledge the financial support from the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant numbers JP20H05631, JP21J00008) and Joint Usage/Research Grants 2021 at the Institute of Social and Economic Research, Osaka University.

† Corresponding author: Graduate School of Economics, Kobe University, 2-1, Rokkodai-cho, Nada-ku, Kobe, Hyogo, 657-8501, Japan. E-mail: kittaka@econ-io.jp, ORCID: 0000-0002-9876-4492.

‡ Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka, 560-0043, Japan. E-mail: mikamiryomyself@gmail.com, ORCID: 0000-0002-4927-5174.

§ Institute of Social and Economic Research, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan. E-mail: nshimada@iser.osaka-u.ac.jp, ORCID: 0000-0003-1824-0590.

# 1 Introduction

Daily economic behavior often entails sequential information search behavior with multiple options, such as purchasing goods and/or services or job-seeking. For example, we often do not know product prices or job wages beforehand and thus search for their information. Based on the seminal search models of Lippman and McCall (1979) and Weitzman (1979), scholars have conducted many experimental studies investigating people's search behavior across research fields (e.g., Schunk, 2009; Schunk and Winter, 2009 in general studies; Caplin et al., 2011; Kittaka and Mikami, 2020; Casner, 2021; Jhunjhunwala, 2021 in consumer studies; and Brown et al., 2011; Miura et al., 2017 in labor studies).[1] These experimental studies are advantageous at connecting and tracking individual characteristics and search behavior in a specified environment and provide valuable implications in various fields.

Depending on the research field and objectives, search experiment designs often differ significantly, despite their identical theoretical models (e.g., differences in the process of action to obtain information and conditions for the termination of a search). Nevertheless, surprisingly, there has been no consensus on the preferable design of search experiments. Although some experimental evidence shows that different experimental designs can lead to considerably different results (e.g., Pedroni et al., 2017),[2] little attention has been paid to this issue in the field of search studies.[3] Because search behavior is closely related to one's daily activities, experimental results can be applied to real-world policies and management decisions. Hence, the existence of various experimental designs that can yield substantially different results may cause confusion or even produce undesirable consequences. We address this gap in our study.

In this study, we investigate whether there are differences in subjects' search behavior

[1] For other examples, see Schotter and Braunstein (1981); Hey (1987); Bearden et al. (2006) for general studies; Braunstein and Schotter (1982); Cox and Oaxaca (1989) for labor studies; and Kogut (1990); Zwick et al. (2003) for consumer studies. These studies differ in purpose and focus but are based on the same theory and employ (almost) the same model.

[2] They showed that different experimental methods elicit different risk preferences. See also Holzmeister and Stefan (2021).

[3] Because the search theory predicts that search behavior depends on individual risk preferences (cf. Nachman, 1972; Lippman and McCall, 1979), differences in search experiment designs may cause different search behavior.

among the different search experiment designs through an experiment. To this end, we categorize existing search experiment designs according to the degree of flexibility of the subjects' decision-making, which, to the best of our knowledge, is a novel attempt in search experiments. We focus on search experiments based on the theoretical work of Lippman and McCall (1979) and Weitzman (1979), such as those based on the "one-step-forward" strategy, and categorize the experiments into *passive*, *quasi-active*, and *active*. (i) Passive search design is a relatively new design in search studies (e.g., Brown et al., 2011; Jhunjhunwala, 2021; Casner, 2021), wherein the subjects choose a cut-off (reservation) value in each search, and then a computer compare the value with a random suggested value. This design differs from the other two designs in that the subject cannot terminate the search until the computer suggests a value higher than or equal to the subject's choice; hence, we call this design (relatively) "passive."[4] (ii) Quasi-active search design is a major design in search experiments (e.g., Schotter and Braunstein, 1981; Sonnemans, 1998; Zwick et al., 2003; Schunk, 2009; Miura et al., 2017), wherein the subjects are offered *one* random value at each step, and they decide whether to continue or terminate the search based on that value. (iii) Active search design is motivated by Caplin et al. (2011); Kittaka and Mikami (2020); Bhatia et al. (2021), wherein the subjects have *multiple* options (they do not know the values in advance) and are free to decide what options to observe or when to terminate the search.[5] Table 1 summarizes the above designs. Based on the findings of Pedroni et al. (2017) and Holzmeister and Stefan (2021), we suspect that such design differences might affect the results through risk preferences. Hence, we conduct an additional risk elicitation task using the multiple price list (MPL) based on the work of Holt and Laury (2002).[6]

[4] A similar experimental design is called the strategy method ("cold" in psychology), wherein the subject describes value(s) at the beginning of each search spell. See Section 3.5 for a detailed discussion.

[5] For other similar experimental design examples (regarding choices, but not exactly the same model), see Gabaix et al. (2006) and Reutskaja et al. (2011). Note that such multiple-choice designs are commonly used in psychology, such as in the Columbia card task (Figner et al., 2009).

[6] In the context of risk and search, Schunk (2009) found that loss aversion, rather than subjects' risk preferences, affected search results, while Miura et al. (2017) reported the opposite results. However, these studies differ in their setting, and hence the relationship between risk preference and search results is still unclear (see Section 5 for more details.)

Table 1: Summary of the designs

| Design | Search decision | Offer | Timing of termination |
|---|---|---|---|
| Passive | state a cutoff value before each search | receive one offer | when the received offer exceeds the stated value |
| Quasi-active | accept or reject the current maximum offer | receive one offer | at any time when clicking the "Finish" button |
| Active | accept or reject the current maximum offer | select one from the potential offers | at any time when clicking the "Finish" button |

We find that the results varied significantly at both the aggregate and individual levels across the designs (treatments). The average number of searches (search duration) for the active search treatment was highest and closest to the theoretically predicted value. Compared with the design, subjects searched significantly less in the quasi-active and passive search treatments. It implies that the widely accepted design, wherein subjects make decisions based on a given offer rather than choosing among potential alternatives themselves, may have unexpectedly suppressed subjects' search; this suggests that the search design is one of the causes of the well-known under-search problem (cf. Schotter and Braunstein, 1981; Hey, 1987; Cox and Oaxaca, 1989). As for the subjects' behavior, the prevalence of recall (i.e., choosing an option observed in the past over a new one) was 12% and almost constant for all treatments. The termination rate (the probability that the subjects would terminate the search when they observed a new offer) was also constant within each treatment for most of the subjects. This suggests that most subjects followed a "one-step-forward" strategy with a constant reservation value, at least within the same experimental design. However, we find significant differences in the termination rates between the passive and active treatments. Risk preference, as measured by the MPL method, had a significant effect on the number of searches only in the passive search design; this may partly explain the difference in behaviors across the designs.

To the best of our knowledge, this is the first study to comprehensively examine search experiments to analyze the effect of different designs in the literature. We contribute to the literature by illustrating that the differences in flexibility that are disregarded in the search model and experiments significantly affect subjects' search behavior even when using the same model. Therefore, the choice of the experimental design is crucial. We present several experimental design choices that can be used for different purposes and in different environments, and provide novel, rich comparisons of search experiments in various designs. Thus, we contribute to subsequent research related to search theory and experiments.

Our methodological implications and arguments are as follows.[7] The active design is superior when the search duration is essential; otherwise, the passive design is preferable, as it can directly track the subject's behavior. The out-of-model factors that are different and usually not considered in various search experiments might be crucial (e.g., whether the offer is given or the subject chooses from potential offers, the timing of decisions, or recall settings). Therefore, experimenters should be careful regarding such out-of-model factors, and our results provide a rationale about what aspects should be considered. Moreover, although many theoretical and experimental studies assume risk neutrality, individual risk preferences may potentially affect search behavior depending on the design. However, the most prevalent method, the MPL, may not be appropriate for the risk elicitation method, at least for the active and quasi-active designs. Thus, researchers need to select the method carefully when investigating the relationship between subjects' risk preferences and their search behavior.

These findings can potentially be applied to various real-world market environments. In the case of job search, some search for multiple candidate firms one by one (i.e., active search), while some use agents to find firms (i.e., passive search). Similarly, when looking for a spouse, people search for a partner from among multiple candidates or utilize the services of a matchmaker for an arranged marriage (e.g., the traditional Japanese custom of Miai). On shopping platforms, consumers may search for products themselves, or the platform may suggest products (e.g., BuyBox on Amazon). Although previous search studies discard this difference in search forms, our results show that such differences in form affect search behavior, and planners can partially control people's behavior by using different search forms.

[7] The detailed discussions are given in Section 5.

With regard to managerial and policy implications, we suggest that a passive search environment is not desirable from a search efficiency perspective because it could create (relatively) severe under-search problems if subjects follow the same search rule. Of course, there are trade-offs between search efficiency and search cost; however, the under-search problem could reduce the search efficiency. Planners should avoid, as much as possible, search environments that reduce individuals' decision-making flexibility (e.g., delegating the search to others or forcing them to think and commit in advance). For example, given cost-efficiency, the benefit of specialization, or the burden of forward-thinking, people tend to delegate search to agents in a job matching or matchmaking environment. However, such delegation may also reduce search efficiency by causing people to terminate the search more quickly; thus, total efficiency could be improved by designing an environment that provides greater search discretion, for example, by reducing the search burden. Conversely, this implies that such a passive search design might address important problems caused by lengthening searches, such as the unemployment problem, even albeit at the cost of efficiency.

The remainder of the study proceeds as follows. Section 2 presents the theoretical model. Section 3 describes the experimental designs. Section 4 reports the results of the experiment. Section 5 proposes methodological implications. Section 6 concludes.

## 2 The model

We first describe a generalized form of a well-known standard search model (cf. Lippman and McCall, 1979) that can be applicable to both labor and consumer studies and all the search experiments we conduct.

Consider a scenario where an individual is searching for information (hereafter, we refer to it as "offers") that will positively affect their utility, such as "wages" or "match utilities." The settings are as follows. In each search, each individual obtains exactly one offer, denoted by $w$, and decides whether to accept the offer (or returns to previous offers, if any) or reject it, and continues the search. We assume that $w$ is randomly drawn from the same twice differentiable cumulative distribution function $F(w)$ with an interval $[\underline{w}, \bar{w}]$ and that each individual is aware of the distribution. The individual incurs a time-independent positive cost $c$ at each search period, which can be interpreted as a *search cost* (including trip costs and opportunity costs) to obtain new information. We suppose that both $w$ and $c$ are independent of each other and of any previous and remaining values of offers, time, and number of searches, as is common in search theory. A search problem is assumed to be an infinite horizon (i.e., the number of offers is infinite). We also assume that each individual search is in a random order.

Consider an individual who follows an optimal search policy to maximize their expected net surplus, given the maximum sampled offer $x$ ($x_0 = 0$). For now, suppose that an individual has a liner utility function. Then, the value of the search (i.e., the value from rejecting offer $x$ and performing another search) is given by

$$V(x) = -c + E\max[w, V(x)]. \tag{1}$$

The above expression can be written as

$$V(x) = -c + \int_{\underline{w}}^{\bar{w}} \max[w, V(x)]dF(w) = -c + \int_{V(x)}^{\bar{w}} (w - V(x))dF(w) + V(x), \tag{2}$$

which yields,

$$c = \int_{V(x)}^{\bar{w}} (w - V(x))dF(w). \tag{3}$$

Expression (3) implicitly and uniquely defines a value of $V(x)$.

Now, consider a hypothetical maximum sampled offer $x^*$ such that the individual is indifferent to the choice between accepting $x^*$ and rejecting it. Then, as is well known, each individual's optimal search rule can be described as follows (cf. Lippman and McCall, 1979; Weitzman, 1979): (i) they should further search if $w < x^*$; (ii) otherwise, they should stop and accept $w$, where $x^*$ is uniquely and implicitly defined by the following expression:

$$c = \int_{x^*}^{\bar{w}} (w - x^*)dF(w). \tag{4}$$

Hereafter, we refer to $x^*$ as a *reservation value*. The expression (4) and an optimal search policy imply that, in each period of the search, an individual compares the search cost (the left-hand side of (4)) and benefits from an additional search (the right-hand side), and then decides whether to search further based on the time-independent variable $x^*$. Therefore, each individual's search behavior is *myopic*. Hence, we refer to this rule as "one-step-forward" search rule with a constant reservation value.

Here, we consider that each individual with an arbitrary, concave, and non-decreasing utility function $u(\cdot)$ follows an optimal search policy that maximizes $E[u(\cdot)]$. Then, from the arguments above, we have

$$u(x) = F(x)u(x - c) + \int_{x}^{\bar{w}} u(w - c)dF(w), \tag{5}$$

and this expression implicitly determines the individual's reservation value $x^*$. It is known that the solution of (5) has the same constant reservation property as seen above; therefore, the "one-step-forward" search rule will persist (cf. Schunk, 2009).

The reservation value $x^*$ described above is a good approximation for capturing individuals' search behavior. Here, we consider another more easily observable (and widely used) measure, such as the number of searches (search duration). As each individual stops with probability $1 - F(x^*)$ (wherein $w > x^*$), the expected search duration can be simply summarized by the following expression:

$$E[N] = \frac{1}{1 - F(x^*)}, \tag{6}$$

where $N$ is the number of searches. Notably, Nachman (1972) showed that more risk-averse individuals have lower $x^*$, requiring them to terminate the search sooner than less risk-averse individuals.

We emphasize that, in this setting, given that an individual follows such an optimal search policy with a constant reservation value, they never choose a *recall* option through which they can return and accept previously rejected offers (cf. Lippman and McCall, 1979). More importantly, as long as the experimental design follows this model, the theory predicts that an individual's search behavior will be characterized by the same reservation value $x^*$ and the same number of searches, regardless of the experimental design.

## 3 Experimental design

We conducted an experiment consisting of two parts. Part 1 is a search task consisting of three treatments, and Part 2 is a risk elicitation task.

The experiment was conducted entirely online at Osaka University, Japan, in July and August 2021 using oTree (Chen et al., 2016); the subjects participated in the experiments using their own computers.[8,9] The total number of subjects was 103, consisting of undergraduate and graduate students aged 18 to 32 years (average age: 21) from various departments of Osaka University. They were recruited through the ORSEE platform (Greiner, 2015). We employed a within-subject design, and each of the three search treatments in Part 1 was repeated for 20 spells (plus five practice spells per treatment). The order of the treatments was randomized by sessions to remove learning effects. The average time required for the experiment was about 60 minutes. The subject's reward was the sum of one randomly selected payoff (times 60) from each of the three search treatments and the payoff of Part 2, and the average earnings was JPY 2,000 (about $18.24).[10] The subjects were informed of their reward after completing the entire session.

Before each treatment, a pre-recorded instruction clip was played, and the subjects could

[8] Common concerns in online experiments include lack of attention and increased noise because of the participation of various subjects (Snowberg and Yariv 2021; Gupta et al. 2021). Snowberg and Yariv (2021) showed that observations involving university-student subjects are less noisy than those collected using other methods (such as MTurk). In addition, Li et al. (2021) showed that online experiments on individual decision-making tasks perform equally well as lab experiments, especially when subjects are observed via webcam. Given their findings, we recruited university students and asked subjects to turn on their webcams and share their computer's screen during the experiment to ensure their complete concentration.

[9] To ensure a unified environment for the subjects, we only permitted them to use personal computers (not tablets or smartphones) and the same software (Google Chrome and Zoom). We set up several personal computers and assigned each participant a room using the Zoom room feature. During the experiment, we stayed connected with the participants via a zoom chat and over the phone.

[10] We adopted this random incentive system wherein a single random outcome is chosen to avoid risk hedging behavior (see Heinemann et al. (2009); Charness et al. (2016) for more details). Note that, in many search experiments, the reward is the sum of all the spells.

start their tasks at their own time after they thoroughly understood the details.[11] In addition, the subjects could read the instructions at any time during each treatment. After the experiment, we administered a questionnaire to determine whether the subjects had understood the experiment correctly on an eight-point Likert scale. Ninety one percent of the subjects provided an answer of seven or eight points (the average point was 7.74). All experiments were conducted in Japanese (the instructions and screens provided in this study were translated into English). As is often the case with search experiments, through the instructions, each subject was informed that they would participate in an experiment designed to investigate how they search for wages or prices.

In each of the three search treatments, subjects make sequential decisions based on offers, simulating typical information search behavior. For consistency, the offer the subjects received was displayed as a number on the card for all designs. This card design was motivated by Sugden et al. (2019) and Kittaka and Mikami (2020).[12] As is common in search treatments, each offer is an independent random draw from a discretized exponential distribution with $\lambda = 0.15$, truncated at 20. At the beginning of the session, using a graph and a table, we explained to the subjects that the offers were drawn from the same distribution across the three search experiments (the subjects could access the details of the distribution at any point). Table 2 shows the distribution we used in Part 1.

Table 2: Probability of drawing each value (always available to subjects)

| Value | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Prob. | 14.66% | 12.62% | 10.86% | 9.35% | 8.05% | 6.92% | 5.96% | 5.13% | 4.42% | 3.80% |
| 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | Total |
| 3.27% | 2.82% | 2.42% | 2.09% | 1.80% | 1.55% | 1.33% | 1.14% | 0.99% | 0.85% | 100% |

Subjects had to pay one search cost to obtain information on one offer. Our design potentially allowed subjects to search infinitely (there was no time limit), and the presence

[11] The instructions used in the experiment are provided in the Appendix.

[12] In addition to search experiments, there are other examples where boxes or images are used to indicate multiple options, such as in Reutskaja et al. (2011) and Deck et al. (2013).

of search costs could lead to negative rewards. However, the expected number of searches was 4.537; with a few exceptions, no subjects conducted an extensive search, and they all received positive rewards. The payoff for each search spell equals the maximum offer that the subject found during each spell minus the total search cost paid (i.e., the total number of searches) during the spell. After each search spell was terminated (conditions for termination are described below), the payoff for that spell was displayed at the bottom, which signaled the subject to proceed to the next spell by clicking on the "Next" button. When the subject clicked on the "Next" button after the spell was repeated 20 times, they were led to the next treatment instruction screen (if any).

A feature of our experiment is that the process of information acquisition and the conditions for terminating the search differ across the three search treatments. In the following subsections, we describe the design of each of the three search treatments.

## 3.1 Passive search

The "passive" search design is popular in more recent search experiments (e.g., Brown et al., 2011; Casner, 2021; Jhunjhunwala, 2021). Unlike in the other two designs, in this design, the subjects select the reservation value directly at each search period.[13] Compared with the active or quasi-active search, this design can be considered as mimicking a real-life situation where the principal delegates the search to an agent (e.g., a recruitment consultant).

Fig. 1 is a screenshot of the passive search treatment. The number of spells is displayed at the top, and the current offer is shown as a card number at the center (the initial offer is zero). Subjects selected a cutoff value (i.e., a reservation value) from a list of numbers from 1 to 20. When a subject selects a reservation value and chooses to continue ("Ask to search" button on the screen), a new offer is presented, incurring a search cost. If the offer is higher than or equal to the reservation value selected, the search is terminated. Conversely, if the offer is strictly lower than the reservation value, the search continues, and the subject is required to select the reservation value again (which can be different from the previous values). Unlike

[13] This type of procedure is called the Becker-DeGroot-Marschak method by Becker et al. (1964).

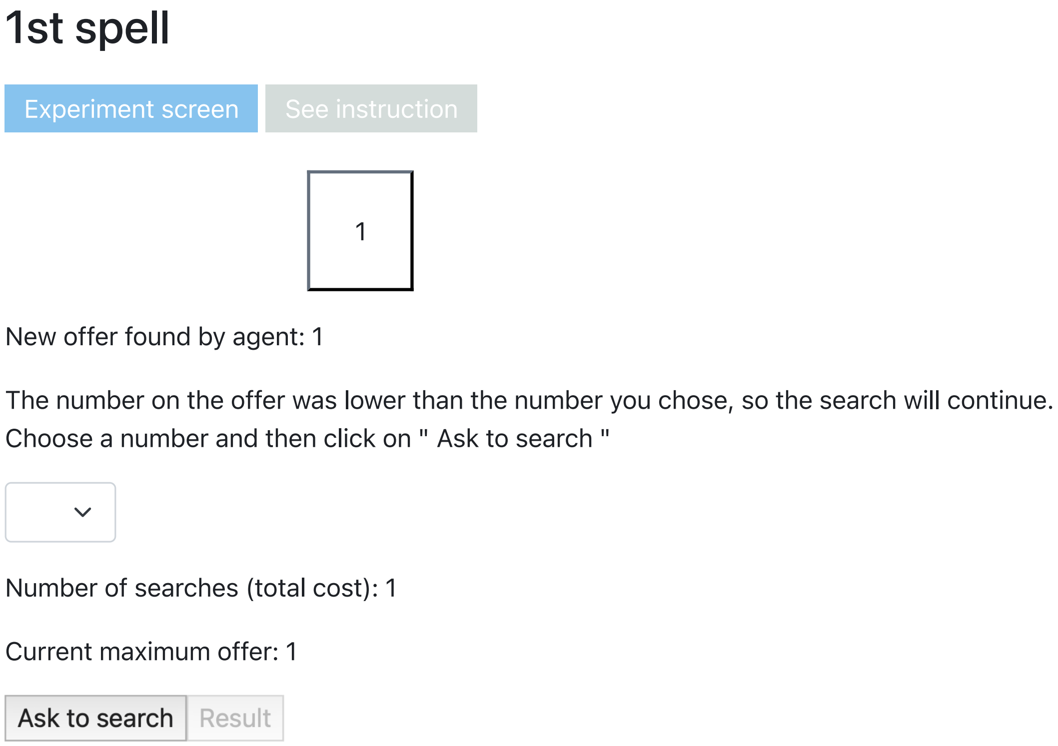



**Fig.** 1: The main screen of the passive search design

some previous studies, we allow recall; that is, the subject's payoff is not the offer that exceeds the reservation value at the time, but the maximum value up to that point.[14,15]

### 3.2 Quasi-active search

To our knowledge, the "quasi-active" search design is the most popular among the experiments in search theory. In this treatment, one offer is always available on the screen. The offer

[14] Casner (2021) allowed subjects to recall, which is similar to our study. By contrast, Brown et al. (2011) explicitly prohibited recall options, and Jhunjhunwala (2021) also partially excluded recall. The latter two designs are appropriate for their purposes; however, the reservation value may differ under the no recall setting. Hence, we allow recall to compare the results of the three designs.

[15] Unlike in the other two treatments, recall is costly in the passive search design. The costly recall setting could lead to a different reservation value. We emphasize that, because we (potentially) allow subjects to search infinitely numerous times, the reservation value may be constant (cf. Janssen and Parakhonyak, 2014). Moreover, as subjects never choose the recall option theoretically (described in the model section), and the intuition is identical, we follow prior studies and employ this design for comparison with other designs even if the costly recall complicates the problem.

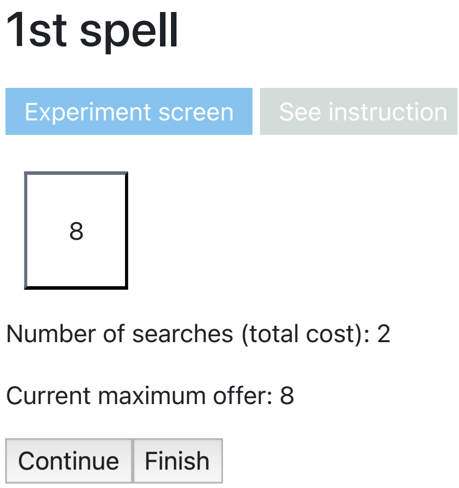



**Fig.** 2: The main screen of the quasi-active search design

is randomly determined according to a predefined distribution, and the subjects choose to continue the search or accept the offer. When a subject clicks on the "Continue" button, a new offer is presented. Once they choose to terminate the search (clicked on the "Finish" button), the maximum offer minus the total search cost equals the payoff for that spell. In terms of a retained maximum value, the subject can return to the previous offers (i.e., recall is allowed). This design (and the active search design) differs from the passive search design in that subjects do not have to choose a reservation value; further, they can terminate the search at any time even if the current offer displayed is lower than the current maximum value. This design corresponds to real-world situations, such as job matching, and the design is often used in the context of labor search (e.g., Miura et al., 2017).

### 3.3 Active search

The "active" search design is inspired by the works of Gabaix et al. (2006); Caplin et al. (2011); Kittaka and Mikami (2020); Bhatia et al. (2021). The difference between a quasi-active search and an active search is that, if a subject decides to search, multiple options (with values written at the back) are displayed, and they are free to choose one. Specifically,

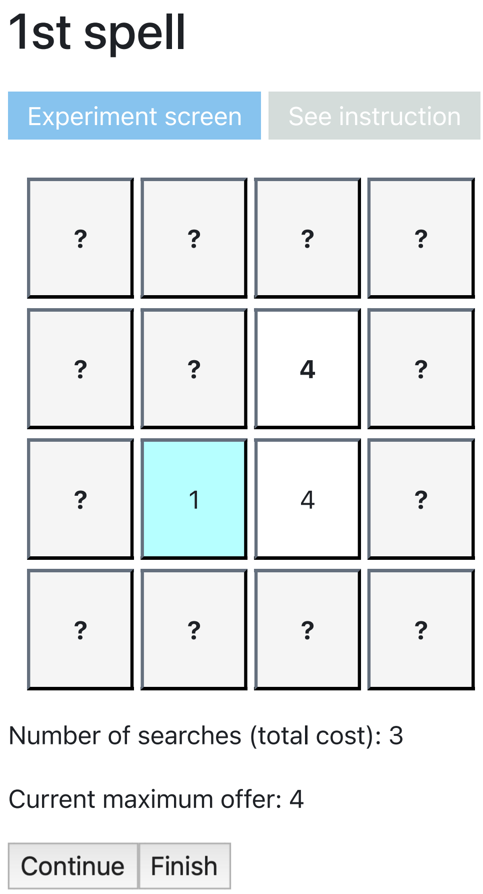



**Fig.** 3: The main screen of the active search design

in the active search design, each subject is presented with 16 available offers and an option to terminate on the display. If the subject chooses all 16 offers, they obtain a new set of 16 offers, while the previous maximum offer is recorded. The value of each offer is not initially known, and subjects can discover the value of their chosen offers by incurring a search cost per offer (they are informed that all offer values will be realized independently from the same distribution). Subjects are also free to choose to terminate their search with the current offer(s) at any point, without any cost. We call this design "active," as subjects are free to decide which offers they want to know the value of and under what conditions they want to terminate their search.

Fig. 3 is the screen of the active search treatment. When a subject clicks on the "Continue" button and chooses one of the cards (offers), the value of the chosen offer is displayed, incurring one search cost. When the subject clicks on the "Finish" button, the search process is terminated, and the maximum current offer minus the total search costs is displayed and recorded as the payoff for that spell (i.e., recall is allowed). Although consumer search experiments often employ quasi-active or passive designs as described above, we believe that this active search design is the closest to the context of real-world consumers' choice behavior.

### 3.4 Risk elicitation task

We also elicited subjects' risk preferences in Part 2 by MPL based on Holt and Laury (2002). MPL is one of the risk elicitation methods that is widely used not only in search experiments but also in various research for measuring individual risk preference (e.g., Tversky and Kahneman, 1992; Andersen et al., 2008; Beauchamp et al., 2020).

Fig. 4 displays a part of the content of the questionnaire. There are 20 pairs of two lotteries, and subjects choose between "Lottery A" and "Lottery B." The payoff depends on the probability ($p$) that the higher value of each lottery is realized. If subjects choose lottery A (safe lottery), they receive JPY 600 with probability $p$ or JPY 300 with probability $1-p$; if they choose lottery B (risky lottery), they receive JPY 900 with probability $p$ or JPY 0 with probability $1-p$. The probability $p$ starts at 5% and increases by 5% from the first row to the twentieth row. For example, in the first row, the expected payoff for the risky lottery becomes greater than that for the safe lottery (JPY 315 and JPY 45, respectively). For the MPL, in the top nine rows, the expected payoff of the safe lottery is greater than that of the risky one. The expected payoff of the safe lottery is equal to that of the risky one in the tenth row; therefore, the risk-neutral subjects switch from Lottery A to B in the tenth or eleventh rows. Because risk-averse subjects tend to choose a safer lottery, they would switch from A to B before the tenth row. Hence, we use the number of safe lotteries chosen as a proxy for each subject's risk aversion.

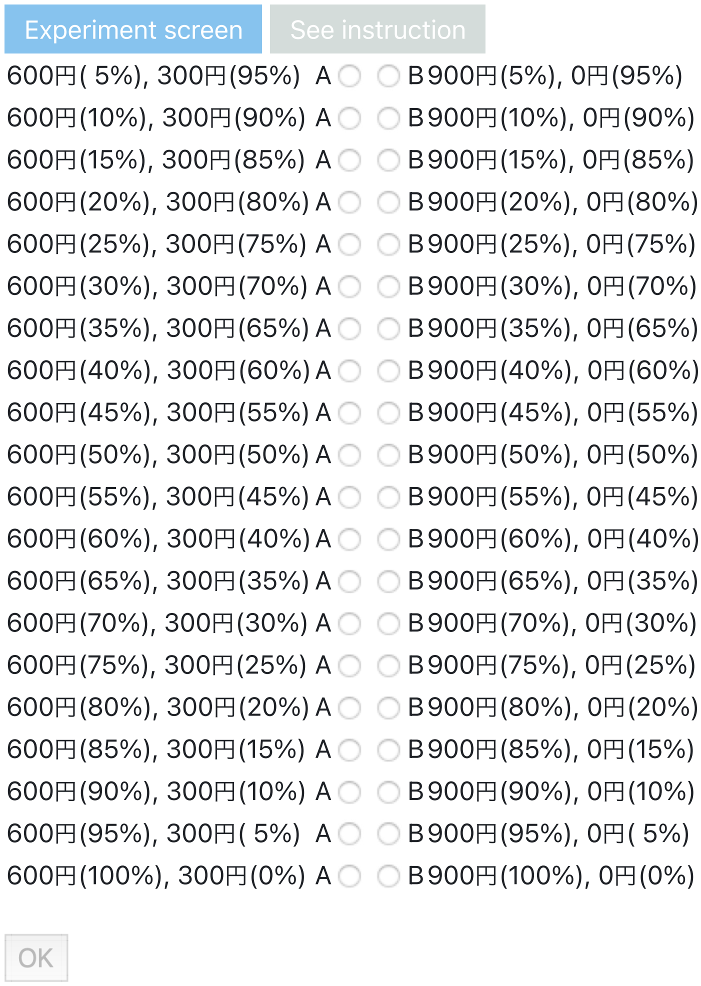



**Fig.** 4: The screen of the multiple price list task (the symbol “円” represents JPY)

### 3.5 Discussion: Search design using a strategy method

The search designs mentioned above, wherein players respond sequentially to the outcome of each search, are called *direct response methods* in the context of game theory experiments. Such methods are predominantly used in search experiments. Another well-known method in the context of game theory experiments is the *strategy method*, wherein the subject determines the responses to all possible outcomes in advance. Theoretically, in search environment, subjects have one reservation value, and therefore these differences in methods do not affect the experimental results. There are several search experiments based on the strategy method. Sonnemans (1998) and Asano et al. (2015) conducted an experiment based on the strategy method; however, the difference between this and the direct response method is ambiguous. We conducted a pre-experiment that included a design based on the strategy method and found potential differences from other designs. However, including the strategy method requires larger samples, and this method is not commonly used in recent search experiments. Furthermore, the difference between the strategy method and the passive design is that the subject only has to enter a value initially or can change the value depending on the results of each search. Therefore, we exclude the design from the main experiment.

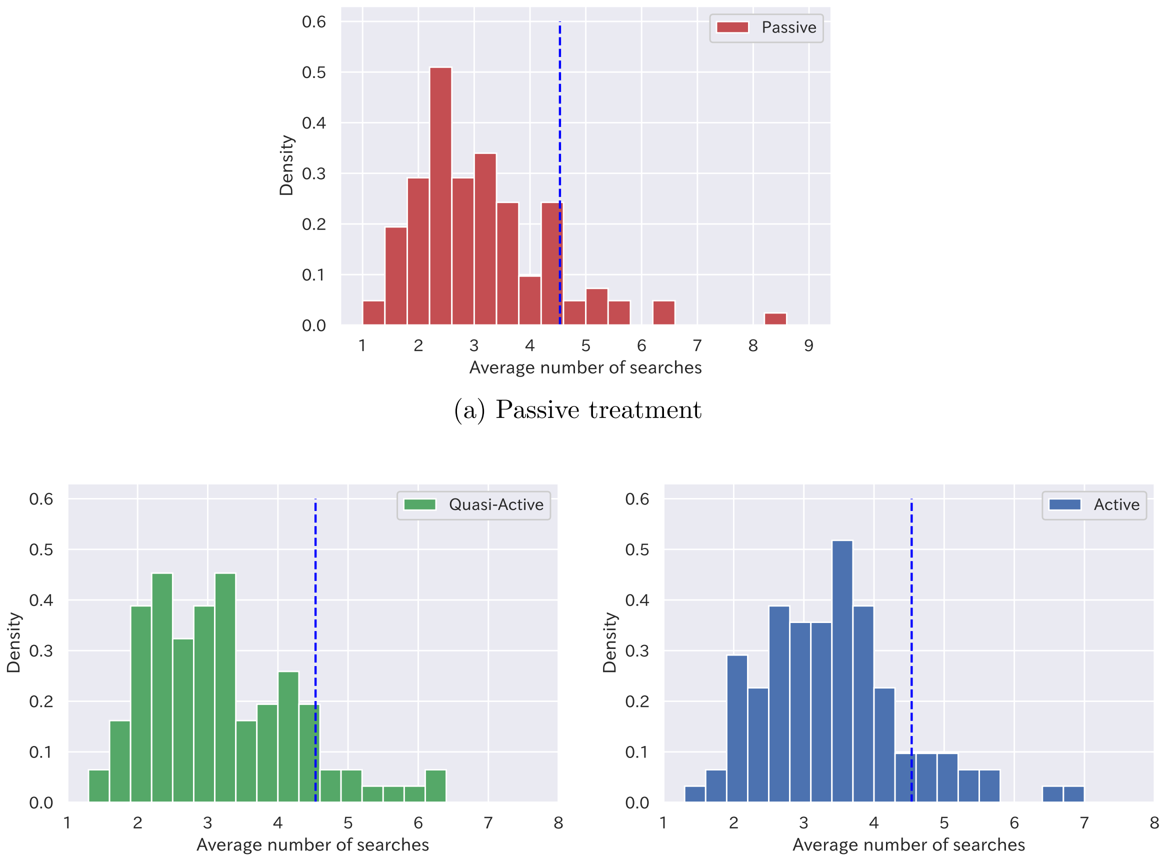
## 4 Results and discussion

As we are interested in determining whether there is a difference among the designs, we first compare the results across the treatments.

Table 3: Summary of data

| | | | | Theoretical | | |
|---|---|---|---|---|---|---|
| | Passive | Quasi-active | Active | Lower | Upper | Risk-neutral |
| Average payoff | 8.298 | 8.437 | 8.413 | 8.863 | 9.012 | 9.028 |
| | (0.028) | (0.026) | (0.028) | | | |
| Average number of searches | 3.095 | 3.133 | 3.380 | 3.665 | 3.993 | 4.537 |
| | (0.027) | (0.022) | (0.022) | | | |
| Number of observations | 2060 | 2060 | 2060 | 100 | 100 | – |

*Note*: Standard errors are presented in parentheses. The "Theoretical" columns present model-predicted values. The "Lower" and "Upper" columns show the mean of the model-predicted lower and upper values using the observed risk aversion parameters elicited by the risk elicitation task (we consider two values because subjects' risk preferences are expressed as intervals rather than point values in the task). The "Risk-neutral" column shows the model-predicted values for a risk-neutral subject. We exclude three subjects who chose safe options on all rows in the task because their risk aversion parameter cannot be identified.




(a) Passive treatment

(b) Quasi-active treatment

(c) Active treatment

**Fig.** 5: Histograms of the average number of searches for each treatment

## 4.1 Main results

Table 3 reports the summary statistics for each treatment. The average payoff (the maximum offer found minus the total search cost in each spell) was slightly lower in the passive treatment, but there were no statistical differences across treatments. However, we find that the average number of searches differed significantly across treatments. Specifically, the average number of searches in the active search treatment was significantly higher than in the other treatments at the 1% significance level.[16,17] Moreover, at the individual level analysis, we found that the average number of searches was also different between the active treatment and both the quasi-active and passive treatments at the 5% level.[18]

Recall that the only difference between the quasi-active and active treatments lies in whether subjects make decisions based on a single offer received or whether they can choose one from several potential offers (this difference also exists between active and passive treatments). Therefore, we conclude that how the experimenter presents potential offers to subjects in a search experiment has a noticeable influence on subjects' search behavior (whether subjects can see more than one potential offer or whether they are free to choose among them).

We also consider the tendencies in the average number of searches across designs in more detail. Fig. 5 illustrates the distribution of the number of searches averaged by subject. The dashed line represents the theoretically predicted number of searches for risk-neutral subjects. The figure shows that the distribution trend of the average number of searches also seemed to differ among the treatments. The peak of the average number of searches was around 3.6 in the active treatment (Fig. 5c) and 2.3 and 3.3 in the quasi-active treatments (Fig. 5b), while it was around 2.3 in the passive treatment (Fig. 5a). These results are consistent with the above arguments and suggest that, among all experimental designs, the passive treatment potentially suppresses the number of searches the most overall.

[16] We also conduct a non-parametric test for the average number of searches. The Mann–Whitney U (Wilcoxon rank-sum) test was significant at the 5% level for the differences between the active and quasi-active treatments and at the 1% level for the difference between the active and passive treatments. Conversely, there was no significant difference between the quasi-active and passive treatments ($p = 0.1001$).

[17] In the passive treatment, the top 10 spells with the highest number of searches were significantly high (the average number of top 10 spells was 26.00 in the passive treatment, while it was 18.30 in the active treatment and 19.90 in the quasi-active treatment). This tendency can also be seen in Fig. 5a (outliers in passive treatment). The average number of searches in the passive treatment was 3.006 after removing the data that exceed the maximum number of searches in other spells (21 in both the active and quasi-active treatments). Thus, we cannot conclude that there were no potential difference between quasi-active and passive treatments when we remove the outliers.

[18] The Wilcoxon signed-rank sum test for the differences between the active and quasi-active (passive) treatments was significant at the 5% level, while it was not significant between the quasi-active and passive treatments.

### 4.1.1 Discussion: Potential differences in the passive treatment

As there was no significant difference in the results in Table 3 between the quasi-active and passive designs, it seems that the differences between the two designs (whether to select the reservation value before the search or to respond to the result after the search; the timing of termination) may not affect subject behavior. However, Fig. 5 shows that there was a potential difference between the two designs. Before analyzing the behavioral differences among designs, we need to discuss some factors that are beyond theory when regarding the results of the passive treatment.

First, regarding the aspect of recall, subjects have to carry out an additional search in order to return to the previous offer in the passive treatment because of the difference in the timing of termination. Although subjects never choose recall in theory, recall behavior is often observed in search experiments, and such a recall design may cause the subject to search more (or possibly less) than in other treatments in the experiment. In fact, we found that some subjects selected an extremely low reservation value after the second search in the treatment, which implies that they may have carried out an additional search in order to terminate the search (see Section 4.3) and may have masked the differences in the number of searches. Some observations specific to the passive design may also relate to the point: There were a few outliers in the passive design that had a distinctly high number of searches (see Fig. 5a). Also, as we note in Footnote 17, there were some spells wherein the number of searches was fairly higher only in the passive treatment.

Second, each subject must write their reservation value before each search and must terminate the search as long as the offer exceeds the reservation value in the passive treatment; namely, the subject has to commit their search decision *ex-ante*. Such a pre-commitment may cause subjects to be more reluctant to search. While we did not control for these points, they can increase or decrease the average number of searches. Third, because the passive treatment is the only one that requires forward-thinking, one might think that the design requires more effort (and thus more time) in each search decision. We will discuss this issue in Section 5.

#### 4.1.2 Discussion: An under-search problem

We observe an overall "under-search" problem as in previous studies, where subjects searched less than theoretically predicted ("Theoretical" columns in Table 3). From the observations in the risk elicitation task (see Table 4), we found that the subjects were risk-averse overall (79 subjects chose the safe lottery more often than the risky lottery), which implies that the number of searches observed would be less than the risk-neutral theoretical value. Even when we considered the risk-averse subjects, the number of searches observed in the tasks were still lower than the theoretical value ("Lower" and "Upper" columns in Table 3). However, the subjects' behavior was closest to the theoretical predictions in the active search treatment. We thus argue that adjusting the degree of flexibility of subjects' decision-making (e.g., regarding potential offers) may be an effective remedy to partially solve the under-search problem.

Table 4: Observations in the risk elicitation task

| Number of safe lotteries | 1 | 2 − 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Number of observations | 1 | 0 | 4 | 1 | 3 | 15 | 10 | 7 | 16 | 11 | 9 | 8 | 5 | 4 | 6 | 3 |

*Note*: "Number of safe lotteries" is the number of safe options subjects choose in the risk elicitation task (taking values between 1 to 20). We use this variable as a proxy for measuring the subject's risk aversion.

### 4.2 Behavioral differences

In the last subsection, we have compared the results across treatments and confirmed that differences in design regarding the offer influence subjects' behavior. In this subsection, we consider whether there are differences in search behavior across designs within subjects.

#### 4.2.1 Spell-to-spell transitions and learning

Before exploring the search behavior in detail, we briefly examine whether there was a learning effect. As mentioned in Section 3, each treatment was conducted in a random order to eliminate learning effects across the treatments. In addition, Fig. 6 shows that there was no significant trend in the average number of searches across spells. These results suggest that the learning effect is not the underlying factor in the different results across the treatments.

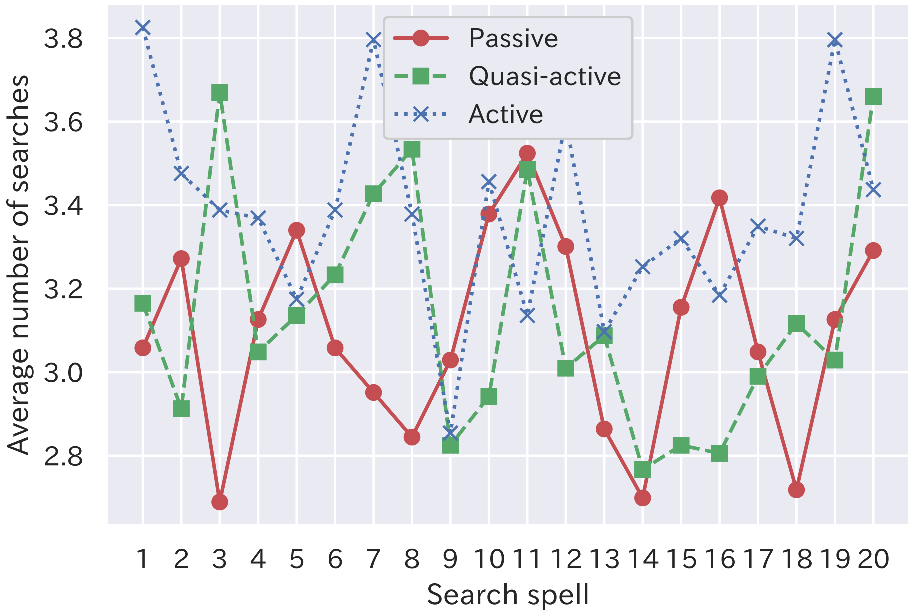



**Fig.** 6: The relationship between experience and the average number of searches

### 4.2.2 Search strategy

Some previous studies based on the passive design (Brown et al., 2011; Casner, 2021) suggested that a downward trend is observed in subjects' reservation value, and subjects do not follow a search strategy based on the constant reservation value, as described in Section 2.

Here, we investigate whether subjects' search strategies are based on the constant reservation value in all designs. We further explore whether the tendency regarding search strategies differs across designs.

First, we consider the subjects' search strategies from the viewpoint of recall behavior. The reason is as follows. Consider a subject who follows the optimal search strategy with a constant reservation value. If they decided to continue the search instead of terminating the search at the previous drawn offer value, then they would never choose the past offer value again. We define a subject as exhibits "recall" if they terminate the search even though the last offer is equal to or less than the previous maximum offer.

Table 5: Summary of data regarding recall

| Treat. | Number of recalls | Number of obs. | Prevalence of recall (%) |
|---|---|---|---|
| Passive | 245 | 2060 | 11.9% |
| Quasi-active | 264 | 2060 | 12.8% |
| Active | 231 | 2060 | 11.2% |
| Total | 740 | 6180 | 12.0% |

Table 5 reports a summary of our data regarding recall. The prevalence of recalls was not large throughout the three search treatments, and the trends were similar across the treatments. This suggests that a search strategy based on a constant reservation value is likely to be consistent with subject behavior. We analyze the subjects' search strategy in more detail below.

#### 4.2.3 Termination rate

Here, we consider the trend of the termination rate of the search because the rate should be almost constant throughout the game if subjects employ the constant reservation value strategy.

Fig. 7 shows the percentage of subjects who terminated the search among those who continued to search until a particular duration at the aggregate level up to the 10th search (note that less than 15% of the spells lasted for six or more searches and less than 4% lasted for 10 or more).[19] The termination rate of the subjects was approximately 30% for the active and quasi-active treatments overall. This result suggests that, irrespective of how long the search lasted, the variance in the reservation value is not large enough in these treatments under the condition that all subjects obey their own constant reservation value strategy. In the passive search process, the termination rate was constant for spells with up to six searches and decreased slightly for spells with six or more successive searches; that is, the variance in the reservation value was relatively large compared with that in the other treatments.

Table 6 shows the fraction of subjects whose termination rates changed with the search

[19] The number of observations after the 10th search was too small ($N < 60$) and was thus excluded.


**Fig.** 7: Percentage of subjects terminating the search per period across treatments

Table 6: Termination rate change among subjects ($N = 103$)

| Treat. | Termination rate (first search) | Number of subjects who differed in termination rate by search duration | | | |
|---|---|---|---|---|---|
| | | Second search | Third search | Fourth search | Fifth search |
| Passive | 31.6% | 12 | 7 | 6 | 8 |
| Quasi-active | 29.9% | 8 | 8 | 5 | 6 |
| Active | 28.4% | 9 | 8 | 4 | 7 |

*Note*: The "Termination rate" column indicates the average probability that the subjects terminated the search in the first search. Columns (3) to (6) show the number of subjects who would change the termination rate from the value of the first search in the number of searches, by using Fisher's exact test at the 5% significance level. Searches beyond the sixth search were not included, because of the small sample size and difficulty in testing.

duration compared with the initial (baseline) termination rate. Regarding whether there was a difference in the termination rate, we found no significant difference in the search duration up to the fifth search within treatments at the individual level, except in the case of a few subjects. This suggests that, at least for the first couple of searches, most subjects may employ a search strategy based on a constant reservation value.

However, for 13.5% of the subjects, we found significant differences in the baseline termination rate across treatments within individuals (calculated by Fisher's exact test at the 5% significance level). To investigate the causes of the high termination rate in the passive treatment, we analyze the relationship between risk preferences and termination rate.

#### 4.2.4 The effect of risk preferences

We conduct a regression analysis to investigate what might have caused the differences mentioned above. We employ a logit model with a random effect at the individual level using a dummy variable as a dependent variable; the variable takes the value of 0 if the subject continues the search in each search decision and 1 if the subject terminates the search. We analyzed this dependent variable via five regression models using three types of dependent variables, which are described below. We used the active search treatment as a baseline and created "Passive" and "Quasi-active" as treatment dummy variables. "Number of safe lotteries" is the number of safe options (lotteries A) that subjects choose in the risk elicitation task (taking values between 1 and 20), which is a time-invariant variable within individuals. "Number of searches" is the number of searches by the subject in each spell.

Table 7 reports the results of the logistic regression of the termination rate. Using two treatment dummy variables, Column (1) in Table 7 shows that the passive and quasi-active treatments positively affected the probability of termination at the 1% and 5% levels of significance. We then carried out an analysis to see the effect of the number of searches in Columns (2) and (3). From Column (2), we could not confirm an overall trend that the subject changed their search strategy throughout the search, which is consistent with prior arguments. By treatments, Column (3) shows that the coefficient was negative and significant only in the passive treatment, but the magnitude was relatively small. Finally, using the number of safe lotteries, we investigated the effect of risk preferences on the termination rate in Columns (4) and (5). Column (4) is consistent with the theoretical prediction that more risk-averse subjects tend to terminate the search quickly, although Column (5) shows that this result was not common across all designs. More specifically, in the passive search

Table 7: Estimates of the decision to terminate using logistic regression with a random effect model

| Covariates | Coefficients (SE) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Passive | $0.102^{***}$ | | $0.188^{***}$ | | $-0.229$ |
| | (0.036) | | (0.052) | | (0.145) |
| Quasi-active | $0.083^{**}$ | | $0.094^{*}$ | | $-0.062$ |
| | (0.036) | | (0.054) | | (0.142) |
| Number of safe lotteries | | | | $0.012^{*}$ | 0.001 |
| | | | | (0.007) | (0.010) |
| Passive × Number of safe lotteries | | | | | $0.025^{**}$ |
| | | | | | (0.011) |
| Quasi-active × Number of safe lotteries | | | | | 0.011 |
| | | | | | (0.010) |
| Number of searches | | $-0.005$ | 0.007 | | |
| | | (0.005) | (0.009) | | |
| Passive × Number of searches | | | $-0.027^{**}$ | | |
| | | | (0.012) | | |
| Quasi-active × Number of searches | | | $-0.003$ | | |
| | | | (0.013) | | |
| Constant | $-1.202^{***}$ | $-1.124^{***}$ | $-1.225^{***}$ | $-1.306^{***}$ | $-1.215^{***}$ |
| | (0.033) | (0.101) | (0.044) | (0.030) | (0.130) |
| Observations | 25,981 | 25,981 | 25,981 | 25,981 | 25,981 |

*Note*: $^{***}$, $^{**}$, and $^{*}$ indicate significance at the 1%, 5%, and 10% levels. Standard errors (SE) are in parentheses. We used the active search treatment as a baseline treatment. Passive and Quasi-active are dummy variables. Number of safe lotteries is the number of safe options subjects choose in the risk elicitation task (taking values between 1 and 20). We use this variable as a proxy for measuring the subject's risk aversion. Number of searches is the number of searches by the subject in each spell.

design, more risk-averse subjects tended to be more reluctant to search, unlike in the other designs. Thus, we argue that the different effects of the risk preferences can partially explain the differences in search behavior among designs.

## 4.3 Consistency with previous studies

Thus far, we have found no evidence that subjects had a decreasing reservation value trend, unlike previous studies using the passive design (e.g., Brown et al., 2011; Casner, 2021). There were no significant differences in the termination rates for the first couple of searches. Furthermore, if subjects had a decreasing reservation value trend, then the long-term termination rate trend must become higher. However, Fig. 7 and Column (3) in Table 7 show that the termination rate in the passive design was lower later in the search, which is inconsistent with the case of a long-term decreasing reservation value trend. We then pose a pertinent question: Does our study yield different results from previous studies?

Here, we used the passive design to investigate the changes in the reservation value in detail. We believe this to be a superior design for this purpose in that it allows us to directly observe the subject's reservation value. According to Brown et al. (2011), we compared the reservation value selected at the beginning of each spell with the last value selected to determine if they were the same or showed an increasing or decreasing trend, at an aggregate level. Although we found some patterns indicating that the trend changed mid-spell or was random, we only considered the first and last values because such patterns were found for all the three trends.

Table 8: Trend in the reservation value in the passive treatment

| Trend | Num. | Percentage | Avg.Rsv.Value |
|---|---|---|---|
| Only one choice | 652 | 31.6% | 7.12 |
| Same value | 651 | 31.6% | 8.91 |
| Decreasing | 593 | 28.7% | 9.87 |
| Increasing | 164 | 7.9% | 9.97 |
| Total | 2060 | | |

Table 8 shows the trend of the reservation values. In approximately 30% of the spells, the subjects followed a rigorous, constant reservation value strategy; that is, they continued choosing only one value within the same spell, which is consistent with the observation that the termination rate was constant.

Conversely, we also found a decreasing trend in the reservation value in some spells, similar to Brown et al. (2011) (but the percentage was much smaller than that of theirs). Further, the average reservation value selected in those spells was significantly higher than in the spells where the reservation value was constant (significant at the 1% level). To see how the downward trend of the subjects' reservation values takes shape, we consider the values that the subjects choose at the beginning of each search spell.

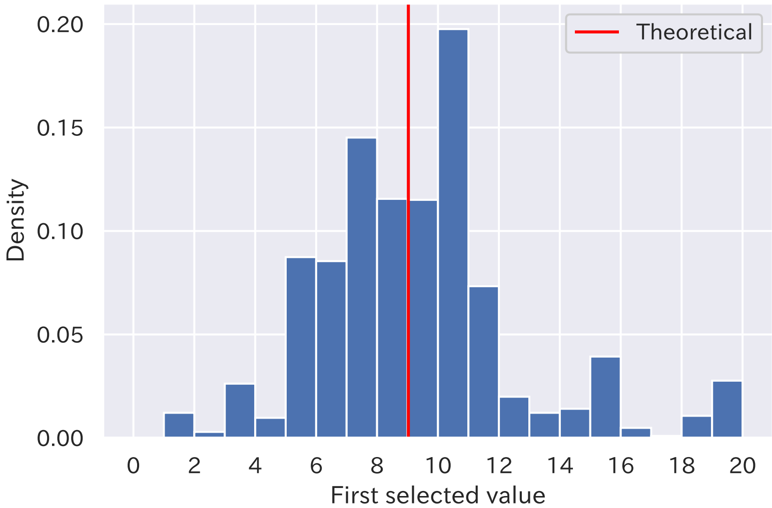



**Fig.** 8: Histograms of the first selected value

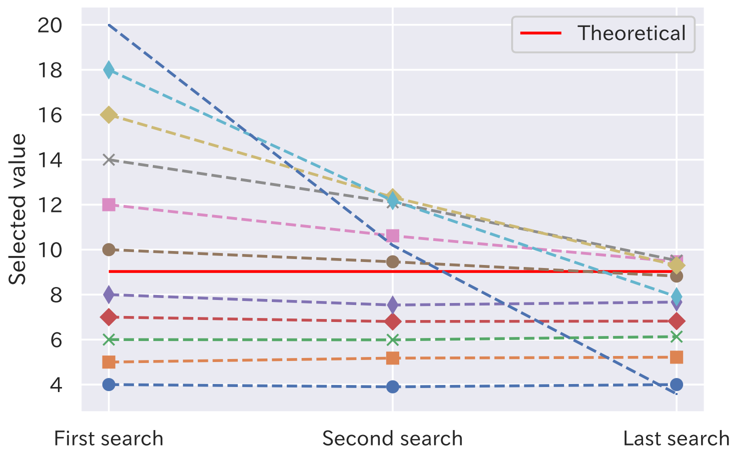



**Fig.** 9: Trends of selected values classified according to the first selected value

Fig. 8 shows that most of the subjects selected a value between 6 and 12 for most of the observations. However, we also found that there exist some spells with higher chosen values such as 16 or 20. Fig. 9 presents the trends of the first value, the next value, and the last value selected, classified according to the first value selected. There is no noticeable difference among the first, second, and last selected values in spells where the first selected value is less than 10. However, we found that the higher the first value selected, the lower the second value selected tends to be, and the final value tends to be drastically lower. For example, the average selected value was 3.6 at the end of the spells wherein the first selected value was 20. Such observations are consistent with those in Fig. 7; that is, the termination rates are higher in the earlier searches (except for the first search) and lower thereafter.[20] Those observations suggest that behavior such as selecting a fairly high reservation value in the first search, and then selecting a lower value is one of the causes of the downward trend in the reservation value in previous studies. However, in the present study, such behavior was not pronounced and did not have a significant influence overall. Although it cannot be stated with certainty whether this trend is unique to the passive treatment, Fig. 7 shows a similar trend of low termination rates for the first search for all treatments (but the fluctuation seem to be the largest in the passive treatment).

# 5 Discussion: The methodological implications

Here we discuss the appropriate experimental design and implications based on findings. First, we discuss when it is appropriate to choose each of the three search designs.

## 5.1 Appropriate experimental design

### 5.1.1 Active search design

An active search design may be an advantageous experimental method in situations where we want to measure the effect of some environmental change, such as that in marketing

[20] Such a heuristic wherein the subject chooses a reservation value higher than the theoretical value for the first few times, and later chooses a lower value, is similar to the heuristic detected by Casner (2021). Interestingly, in the researcher's experimental design, subjects had to acquire information about the distribution of offers. In our experiment, we found similar heuristics even without such a requirement.

strategies or policies. As the magnitude of the outcome values is more important than the detailed process of the search behavior in the situations, the under-search problem may have serious consequences and should be controlled as much as possible.

#### 5.1.2 Quasi-active search design

A quasi-active search design that has been used in many previous studies may be suitable for comparison with those studies. However, even when used for comparison, we argue that the detailed out-of-model setting may be sensitive to the design, as we discuss below, and thus comparisons require close attention. Therefore, for purposes other than comparison with previous studies, using a passive design with similar results on the degree of under-search, but with the advantage of observing reservation values directly, is more recommended.

#### 5.1.3 Passive search design

A passive design is suitable for studying search trends or heuristics, as it allows the direct elicitation of the subject's reservation values, which can be helpful for economics and management studies. However, the design causes the most severe under-search problems. Moreover, as described below, out-of-model factors and MPL effects need to be considered, especially in this design. Hence, improving the passive design is a good challenge for future studies.

### 5.2 Methodological implications

The methodological implications and suggestions of our findings for future (search) experimental designs are discussed below.

#### 5.2.1 Out-of-model factors

Since we found significant differences in the average number of searches between the active and the other designs, we emphasize the importance of the design regarding how the offer is presented. Our results suggest that the design employed by many previous studies, in which subjects make decisions based on a given offer rather than making choices on their own, has an unexpected effect on subjects' behavior. Thus, experimenters should be cautious in their designs regarding offers (especially when considering an under-search problem).

Furthermore, although we did not find a significant difference in the average number of searches between the quasi-active and passive treatments, the result of then regression analysis (Table 7) shows that the the effect of subjects' risk preferences differed between these designs. This implies that the outcomes of those two designs are potentially different, whereas the difference may be masked. As discussed in Section 4.1.1, there are several differences between the passive and other search designs, such as costly recall, forward-thinking, and pre-commitment, which could result in different outcomes (the no-recall setting and elapsed time should also be considered). One might think that these out-of-model factors canceled out and masked the difference (recall that Footnotes 16 and 17 imply that the difference between the two treatments was quite sensitive ). While the settings concerning these factors have often been disregarded in previous studies, we propose that these factors be controlled. In addition, comparisons with previous studies with different experimental designs are probably problematic; even with the same design but different detailed settings, the results could be different (which may not be noticeable with a small sample, and thus a much larger sample size is required). We emphasize that the experimenter must provide details of the experimental design as those out-of-model factors may influence subjects' search behavior.

#### 5.2.2 Time spent searching

The passive treatment is the only one that requires forward-thinking; one might think that the design requires more effort (and thus more time) in each search decision. Brown et al. (2011) and Casner (2021) investigated the effect of elapsed time and found that longer search may result in fatigue, which could lower reservation values (and thus shorter search). Since we only have flawed data on time owing to factors specific to online experiments, the following discussion is not a rigorous argument.[21] However, our data did suggest that the time spent

[21] It is difficult to control for problems specific to online experiments, such as differences in computer performance, Internet connection speed, and lag in program processing, which make an accurate analysis impossible. Face-to-face experiments are necessary to measure cognitive time.

for search (measured in seconds) in the passive search treatment was not longer than that in the active search treatment. Thus, in line with Brown et al. (2011), fatigue in the case of a longer search (and thus shorter search duration) may be rather less noticeable in the passive treatment than in other treatments. This argument is also supported by the observations that there was no difference in the average number of searches between quasi-active and passive treatments. However, we emphasize that the fatigue effect may be easily masked, and subsequent search experiments need to consider the effect of such decision-making time (this effect may be mitigated by less lengthy search spells, as in our experiment).[22]

#### 5.2.3 Termination rate

Previous studies have often used the average number of searches as a measure of subjects' search behavior. Here, we propose termination rates as a new measure that can be used to approximate subject behavior, such as fluctuations in reservation values within a search spell in the active and quasi-active search designs. The use of the termination rate can partially compensate for the disadvantage of these two designs in that the reservation value cannot be measured directly compared with the passive search design. Note that the use of the termination rate is not suitable for long search spells because it requires a large sample size, but it is suitable for checking short search spells (or at least the first few searches).

#### 5.2.4 Risk elicitation method

Regarding experiments in the (quasi-)active search design, similar to our study, Schunk and Winter (2009) also found that risk preference did not affect the search results using different risk elicitation tasks, while Miura et al. (2017) found that risk preferences, as measured by the MPL method, affected search outcomes (they used the quasi-active search design). Unlike Schunk and Winter (2009) (and ours), Miura et al. (2017) introduced a no-recall setting and search termination possibility. The results suggest that the MPL method is sensitive to those factors and may result in uncontrollable outcomes (our passive treatment results also indicate

[22] We plan to conduct an additional experiment to address this issue.

this). Therefore, experimenters should avoid using the usual MPL method as a risk elicitation method for comparisons, or use it very carefully. However, we emphasize that search theory and Bhatia et al. (2021)'s result predict that the risk preferences affect subjects' behavior. Exploring the appropriate risk elicitation tasks for search experiments will be a fruitful future research agenda.[23]

# 6 Conclusion

Growing experimental evidence suggests that individual search behavior depends on experimental design details, in contrast to theory. In the present study, we disentangle the effect of design details by employing three treatments to encompass previous studies—passive, quasi-active, and active—according to the degree of flexibility of the subjects' decision-making.

We found a significant difference in the results at both the aggregate and individual levels across the designs. The average number of searches for the active search design was the highest and closest to the model-predicted value and that for the passive search design was the lowest; this implies that the widely accepted design, wherein subjects make decisions based on a given offer rather than choosing among potential alternatives themselves, may have unexpectedly suppressed subjects' behavior. The prevalence of recall and termination rate analysis results support the fact that most subjects search according to a constant reservation value strategy within designs. However, there was a difference in the termination rate between the passive and active designs. Unlike the quasi-active and active designs, more risk-averse subjects tend to be more reluctant to search in the passive design, which may explain the different behaviors between the designs. Other various out-of-model factors might have also influenced the results. We, therefore, argue that researchers should be cautious when designing search experiments and provide the design details. These findings have many methodological implications that will be helpful for future experiments (not only for search experiments). Furthermore, our findings can be applied to various market environments that potentially take multiple search

[23] See Csermely and Rabas (2016) for an investigation and a comparison of the MPL method and its derivatives.

forms (e.g., job market, matchmaking market, and shopping). We believe that our findings benefit both researchers and practitioners.

## References


Andersen, Steffen, Glenn W Harrison, Morten I Lau, and E Elisabet Rutström (2008) "Eliciting risk and time preferences," *Econometrica*, 76 (3), 583–618.

Asano, Takao, Hiroko Okudaira, and Masaru Sasaki (2015) "An experimental test of a search model under ambiguity," *Theory and Decision*, 79 (4), 627–637.

Bearden, J Neil, Amnon Rapoport, and Ryan O Murphy (2006) "Sequential Observation and Selection with Rank-dependent Payoffs: An Experimental Study," *Management Science*, 52 (9), 1437–1449.

Beauchamp, Jonathan P, Daniel J Benjamin, David I Laibson, and Christopher F Chabris (2020) "Measuring and controlling for the compromise effect when estimating risk preference parameters," *Experimental Economics*, 23 (4), 1069–1099.

Becker, Gordon M, Morris H DeGroot, and Jacob Marschak (1964) "Measuring Utility by A Single-response Sequential Method," *Behavioral Science*, 9 (3), 226–232.

Bhatia, Sudeep, Lisheng He, Wenjia Joyce Zhao, and Pantelis P Analytis (2021) "Cognitive Models of Optimal Sequential Search with Recall," *Cognition*, 210, 104595.

Braunstein, Yale M and Andrew Schotter (1982) "Labor Market Search: An Experimental Study," *Economic Inquiry*, 20 (1), 133–144.

Brown, Meta, Christopher J Flinn, and Andrew Schotter (2011) "Real-time Search in the Laboratory and the Market," *The American Economic Review*, 101 (2), 948–974.

Caplin, Andrew, Mark Dean, and Daniel Martin (2011) "Search and Satisficing," *The American Economic Review*, 101 (7), 2899–2922.

Casner, Ben (2021) "Learning While Shopping: An Experimental Investigation into The Effect of Learning on Consumer Search," *Experimental Economics*, 24, 238–273.

Charness, Gary, Uri Gneezy, and Brianna Halladay (2016) "Experimental Methods: Pay One or Pay All," *Journal of Economic Behavior & Organization*, 131, 141–150.

Chen, Daniel L, Martin Schonger, and Chris Wickens (2016) "oTree—An Open-source Platform for Laboratory, Online, and Field Experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Cox, James C and Ronald L Oaxaca (1989) "Laboratory Experiments with a Finite-horizon Job-search Model," *Journal of Risk and Uncertainty*, 2 (3), 301–329.

Csermely, Tamás and Alexander Rabas (2016) "How to reveal people's preferences: Comparing time consistency and predictive power of multiple price list risk elicitation methods," *Journal of Risk and Uncertainty*, 53 (2), 107–136.

Deck, Cary, Jungmin Lee, Javier A Reyes, and Christopher C Rosen (2013) "A Failed Attempt To Explain Within Subject Variation in Risk Taking Behavior Using Domain Specific Risk Attitudes," *Journal of Economic Behavior & Organization*, 87, 1–24.

Figner, Bernd, Rachael J Mackinlay, Friedrich Wilkening, and Elke U Weber (2009) "Affective and Deliberative Processes in Risky Choice: Age Differences in Risk Taking in The Columbia Card Task.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35 (3), 709.

Gabaix, Xavier, David Laibson, Guillermo Moloche, and Stephen Weinberg (2006) "Costly Information Acquisition: Experimental Analysis of A Boundedly Rational Model," *The American Economic Review*, 96 (4), 1043–1068.

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

Gupta, Neeraja, Luca Rigott, and Alistair Wilson (2021) "The Experimenters' Dilemma: Inferential Preferences over Populations," *arXiv preprint arXiv:2107.05064*.

Heinemann, Frank, Rosemarie Nagel, and Peter Ockenfels (2009) "Measuring Strategic Uncertainty in Coordination Games," *The Review of Economic Studies*, 76 (1), 181–221.

Hey, John D (1987) "Still Searching," *Journal of Economic Behavior & Organization*, 8 (1), 137–144.

Holt, Charles A and Susan K Laury (2002) "Risk Aversion and Incentive Effects," *The American Economic Review*, 92 (5), 1644–1655.

Holzmeister, Felix and Matthias Stefan (2021) "The Risk Elicitation Puzzle Revisited: Across-methods (in) Consistency?" *Experimental Economics*, 24 (2), 593–616.

Janssen, Maarten CW and Alexei Parakhonyak (2014) "Consumer Search Markets with Costly Revisits," *Economic Theory*, 55 (2), 481–514.

Jhunjhunwala, Tanushree (2021) "Searching to Avoid Regret: An Experimental Evidence," *Journal of Economic Behavior & Organization*, 189, 298–319.

Kittaka, Yuta and Ryo Mikami (2020) "Consumer Search and Stock-out: A Laboratory Experiment," *Discussion Papers at ISER Osaka University* (1104).

Kogut, Carl A (1990) "Consumer Search Behavior and Sunk Costs," *Journal of Economic Behavior & Organization*, 14 (3), 381–392.

Li, Jiawei, Stephen Leider, Damian Beil, and Izak Duenyas (2021) "Running online experiments using web-conferencing software," *Journal of the Economic Science Association*, 7 (2), 167–183.

Lippman, Steven A and John Joseph McCall (1979) *Studies in the Economics of Search*, 123: North-Holland.

Miura, Takahiro, Keigo Inukai, Masaru Sasaki et al. (2017) "The Effect of Risk Attitudes on Search Behavior: A Laboratory Search Experiment," *Graduate School of Economics and Osaka School of International Public Policy (OSIPP) Osaka University Discussion Papers In Economics And Business*, 17, 1–16.

Nachman, David C (1972) "On Risk Aversion and Optimal Stopping,"Technical report, Discussion paper.

Pedroni, Andreas, Renato Frey, Adrian Bruhin, Gilles Dutilh, Ralph Hertwig, and Jörg Rieskamp (2017) "The Risk Elicitation Puzzle," *Nature Human Behaviour*, 1 (11), 803–809.

Reutskaja, Elena, Rosemarie Nagel, Colin F Camerer, and Antonio Rangel (2011) "Search Dynamics in Consumer Choice Under Time Pressure: An Eye-tracking Study," *The American Economic Review*, 101 (2), 900–926.

Schotter, Andrew and Yale M Braunstein (1981) "Economic Search: An Experimental Study," *Economic Inquiry*, 19 (1), 1–25.

Schunk, Daniel (2009) "Behavioral Heterogeneity in Dynamic Search Situations: Theory and Experimental Evidence," *Journal of Economic Dynamics and Control*, 33 (9), 1719–1738.

Schunk, Daniel and Joachim Winter (2009) "The Relationship Between Risk Attitudes and Heuristics in Search Tasks: A Laboratory Experiment," *Journal of Economic Behavior & Organization*, 71 (2), 347–360.

Snowberg, Erik and Leeat Yariv (2021) "Testing the Waters: Behavior Across Participant Pools," *The American Economic Review*, 111 (2), 687–719.

Sonnemans, Joep (1998) "Strategies of Search," *Journal of Economic Behavior & Organization*, 35 (3), 309–332.

Sugden, Robert, Mengjie Wang, and Daniel John Zizzo (2019) "Take It or Leave It: Experimental Evidence on The Effect of Time-limited Offers on Consumer Behaviour," *Journal of Economic Behavior & Organization*, 168, 1–23.

Tversky, Amos and Daniel Kahneman (1992) "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5 (4), 297–323.

Weitzman, Martin L (1979) "Optimal Search for the Best Alternative," *Econometrica: Journal of the Econometric Society*, 47 (3), 641–654.

Zwick, Rami, Amnon Rapoport, Alison King Chung Lo, and AV Muthukrishnan (2003) "Consumer Sequential Search: Not Enough or Too Much?" *Marketing Science*, 22 (4), 503–519.

# Appendix

We conducted our experiment using oTree (Chen et al., 2016) and a web-server (Heroku and oTree hub). After the subjects gained access to Zoom, we shared the URL (using Zoom's room feature) via chat. The instructions were recorded on video, and the subjects could view it at their own pace and convenience. At the beginning of each experiment, subjects were required to enter passwords that could only be found by reading each instruction carefully. First, we explained the common instructions, and then conducted each experiment. The order of treatments was changed in each slot. The instructions, including the common ones, could be referred to at any point during each experiment. The details of the instructions are described below.[24]

## A.1 Common instructions

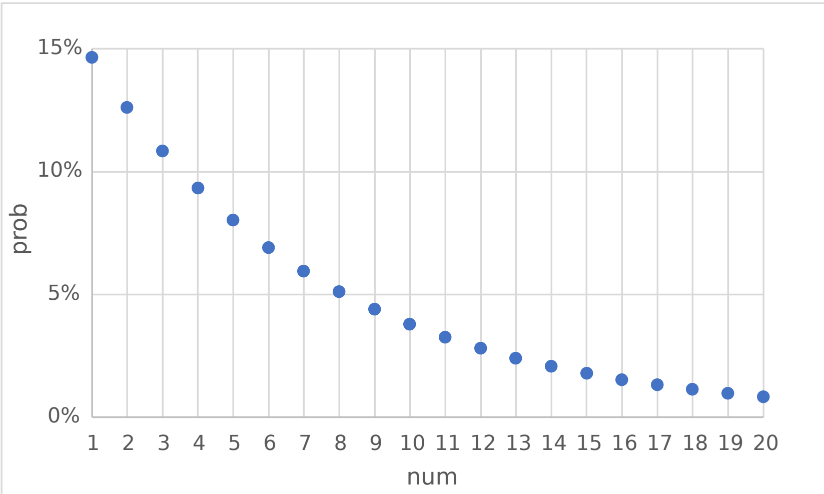
In parts 1 to 3 of the experiment, you will make 20 decisions each. During the experiment, the following cards will appear.

[24] All experiments were conducted in Japanese (instructions and screens provided in Appendix were translated into English).


The cards have numbers on them, and the results of each session depend on the numbers you find during the experiment.

### A.1.1 Regarding the numbers on the cards

The numbers on the cards follow the following probabilities in the interval [1–20]. The number on each card is independent of the number of previous cards found.




| num | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | total |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| prob | 14.66% | 12.62% | 10.86% | 9.35% | 8.05% | 6.92% | 5.96% | 5.13% | 4.42% | 3.80% | 3.27% | 2.82% | 2.42% | 2.09% | 1.80% | 1.55% | 1.33% | 1.14% | 0.99% | 0.85% | 100% |

### A.1.2 Rewards

After all experiments are completed, the computer randomly selects one out of the 20 times. The result of the chosen spell will be multiplied by 60 yen to give you a reward for that part of the experiment.

### A.2 Part 1 (Passive search)

**Experiment overview**

Assume a situation wherein you are paying someone (an agent) to find what you want. For example, you might ask someone to buy a product for you in a shop, or you might ask an agent to help you find a job. The experiment is carried out 20 times in total. Assume that you are looking for something you want 20 times.

**Experiment description**

You tell the agent your "cut off value." The agent looks for a card each time a cutoff value is given. The cost of drawing a card is 1 each time, and you pay the agent at the end. The cards have numbers on them, and the highest number the agent finds on a card is recorded as the maximum point for that spell. If your maximum points are above the cut-off value set, the spell will end automatically, and you will proceed to the next spell. If the maximum points are less than the cut-off value set, you will have to pay again to have an agent find a card for you. (You can set a different cut-off value). The result for each spell is the maximum point for that spell minus the cost of paying the agent (total cost).

**Example**

At the beginning, you set the cut off value to 10. Then, the agent finds a value of 5. The maximum point is updated to 5. At this time, 5 is less than the cut-off value you had set (10); therefore, you cannot complete this spell. The total cost you will incur is 1. Next, you set the reservation value as 15. Then, the agent finds a value of 6. The maximum point is updated from 5 to 6. At this time, 6 is less than the cut-off value that you had set (15); therefore, you cannot complete this spell. The total cost you incur is 2. Next, you set the cut off value as 6. Then, the agent finds a value of 6. The maximum points remain at 6. At this time, 6 is greater than or equal to the standard value set by you (6); therefore, you can finish this spell. The total cost to you is 3. The result of the spell is 3 (=6[maximum point]–3[total cost]).

**Display on the screen**

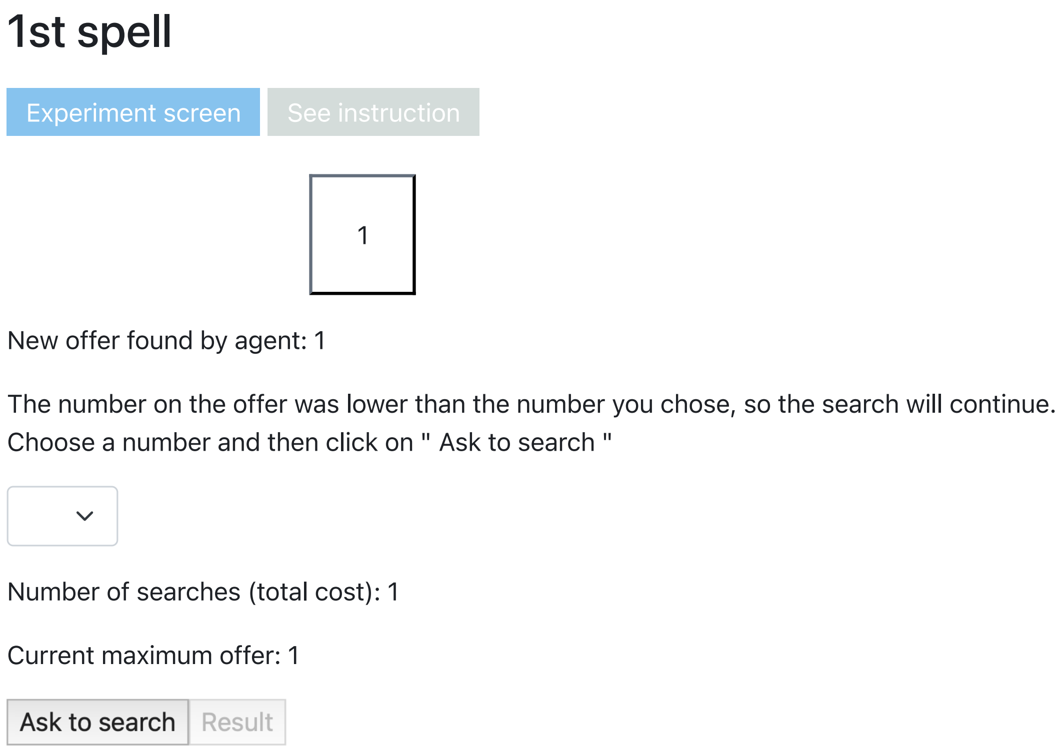



**Fig.** 10: Display on the screen for the passive search treatment

The number of times the agent has searched for a card (total cost) is shown below the card, and each time the agent searches for a card, a cost of 1 is added. The maximum points at this time are the highest number of cards that the agent has found.

## A.3 Part 2 (Quasi-active search)

### Experiment overview

Assume a situation where someone proposes something to you. For example, consider a shop assistant offering a recommendation or a job offer. You will be asked to make a decision. The experiment is carried out 20 times in total. Assume that you are looking for something you want 20 times.

**Experiment description**

There is one card on the screen. You can ask for this card to be changed to another card. The higher the number, the better the goods or the company you have found. The cost of drawing a card is 1 each time. You can choose to draw any card from the ones you have not yet drawn. You can draw as many offers as you want. The payoff for each search spell is recorded as the maximum value that you find during each spell minus the total search cost paid (i.e., the total number of searches) during the spell. After each spell, the results are displayed.

**Display on the screen**

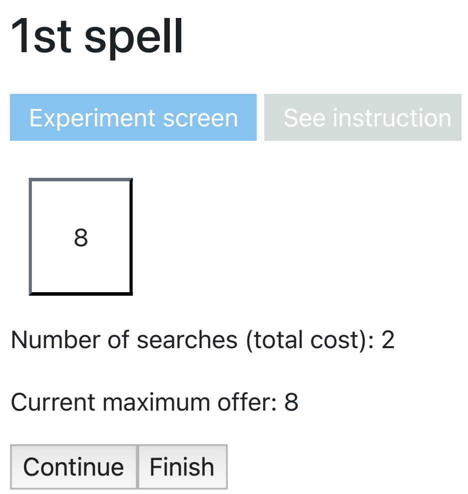



**Fig.** 11: Display on the screen for the quasi-active search treatment

The number of times you changed the card (total cost) is shown below the card, and the maximum point (current maximum offer) for this spell is the highest number of cards.

## A.4 Part 3 (Active search)

The cards are presented on the screen. You can turn over the cards one at a time. Each card has a number on it, and the highest number on the cards you turn over is recorded as the maximum point for that session. The higher the number, the better the goods or the company you have found. The cost of drawing a card is 1 each time. You can choose to draw any card from the ones you have not yet drawn. You can draw as many offers as you want. The payoff for each search spell is recorded as the maximum value that you found during each spell minus the total search cost paid (i.e., the total number of searches) during the spell. After each spell, the results are displayed.

**Display on the screen**

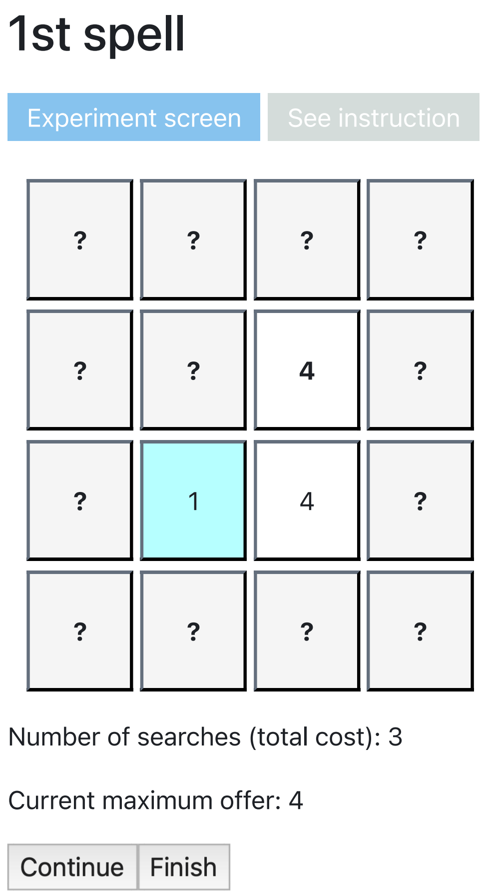



**Fig.** 12: Display on the screen for the active search treatment

There are 16 cards on the screen. When you draw all the cards, a new card will be set automatically. Card [?]: it has not been drawn yet. Card [number]: it has already been drawn. The card you have just drawn will be displayed in light blue.

## A.5 Part 4 (Risk elicitation task by the multiple price list (MPL))

In this part of the experiment, you will be asked to decide at which point you will switch to A or B for a total of 20 questions. Each question is a choice between A and B. A and B are lotteries. Example. A: 10% chance of getting JPY 600 and 90% chance of getting JPY 300. B: 10% chance of getting JPY 900 and 90% chance of getting JPY 0. Please choose which lottery you prefer, A or B.

**Display on the screen**

There are 20 questions on the screen.

**Operation description**

Click on one button anywhere, and A or B will be set automatically. The other choices, apart from the one you selected, will be the same. You can then choose again for each question. Once you have made your decision, click on the "Next" button.

**Rewards**

At the end of the experiment, you will be asked to choose one of the questions from 1 to 20. The reward for this part of the experiment is determined according to the probability of the lottery of the question you had chosen. You will be rewarded for this part of the experiment. At the end of Experiment 4, the rewards for all the experiments are determined.

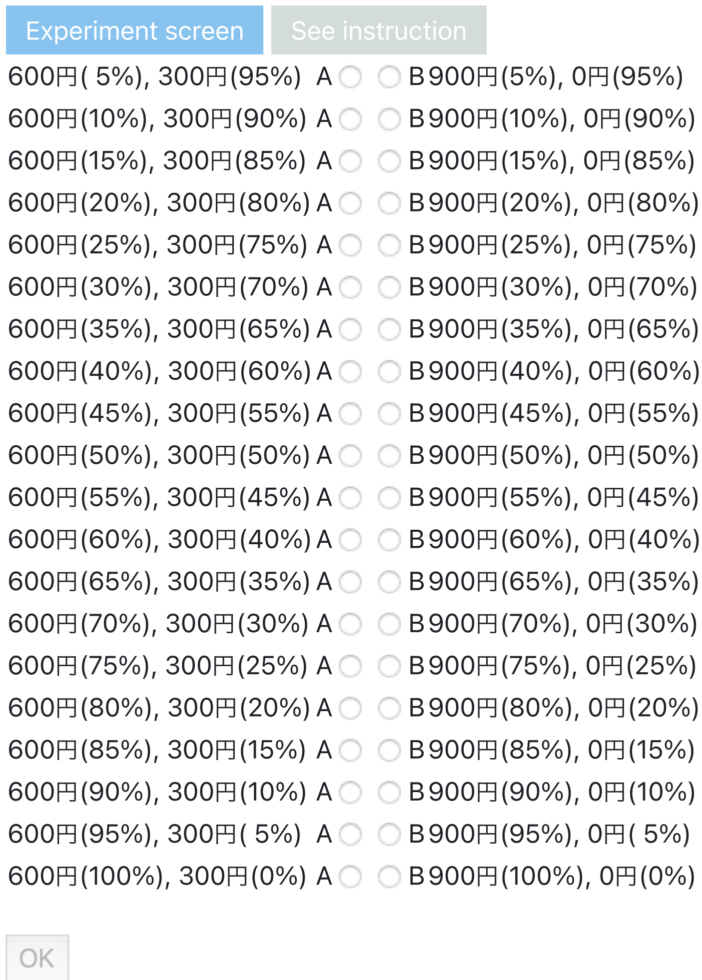



**Fig.** 13: Display on the screen for MPL

(*Note*: The symbol “円” represents JPY).