

Passive or Active? Behavioral changes in different designs of search experiments*

Yuta Kittaka[†] Ryo Mikami[‡] Natsumi Shimada[§]

March 14, 2022

Abstract

While search experiments are available in several designs, accumulating 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. Despite identical models, we found individual- and aggregate-level significant differences in the results across designs. In the passive design, subjects tended to be more reluctant to search than the active design, especially for risk-averse subjects. Our results highlight the importance and potential of controlling the search designs in search environments.

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

JEL classification: C91, D12, D83, J60

* We would like to 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 seminar and workshop participants at 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 (Behavioral Economics) at ISER Osaka University.

[†] Graduate School of Economics, Kobe University, 2-1, Rokkodai-cho, Nada-ku, Kobe, Hyogo Prefecture, 657-8501, Japan. E-mail: kittaka@econ-io.jp.

[‡] Graduate School of Economics, Osaka University, 1-7 Machikaneyama-cho, Toyonaka-shi, Osaka, 560-0043, Japan. E-mail: ryomysself@gmail.com.

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

1 Introduction

Daily economic behavior often entails sequential information search behavior with multiple options, such as purchasing goods, services or job-seeking. Based on the seminal models of Lippman and McCall (1979) and Weitzman (1979), many experimental studies have been investigated people’s search behavior in many research fields (e.g., Schunk, 2009; Schunk and Winter, 2009 in general studies; Brown et al., 2011; Miura et al., 2017 in labor studies; and Caplin et al., 2011; Kittaka and Mikami, 2020; Casner, 2021; Jhunjhunwala, 2021 in consumer studies).¹ These experimental studies are advantageous at connecting and tracking individual characteristics and search behavior in a specified environment and have provided many 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 the information and conditions for termination of the search). Nevertheless, surprisingly, there has been no consensus on the optimal design for search experiments of the proposed categories. Although some experimental evidence shows that different experimental designs can lead to considerably different results (e.g., Pedroni et al., 2017; Holzmeister and Stefan, 2021),² little attention has been paid to this issue in the field of search studies.³ Because search behavior is closely connected to our daily activities, the experimental results are likely to be applied to real-world policy and management decisions. Hence, experimental results based on a variety of experimental designs that may yield substantially different results may cause confusion or even produce undesirable consequences. We address this gap in our study.

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.

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

In this study, we investigate whether and what causes differences in subjects’ search behavior among these 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 is the first 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 three categories: *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), in which the subjects choose a cut-off (reservation) value each search, and then the computer compares 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 one 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 (not knowing the values in advance) and are free to decide what options to choose or when to terminate the search.⁵ Table 1 summarizes the designs mentioned in above arguments. 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

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 the Columbia card task (Figner et al., 2009).

using the multiple price list (MPL) based on Holt and Laury (2002).⁶

Table 1: Summary of designs

Design	Search decision	Offer	Timing of termination
Passive	state a cutoff (reservation) value	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 individual and aggregate levels across the designs (treatments). The average number of searches (search duration) was lower than the theoretically predicted value for all treatments, even when considering subjects’ risk preferences. The result for the active search treatment was the highest and closest to the predicted values, though it was the lowest for the passive search treatment. This result suggests that the search design is one of the causes of a 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 choosing a past option over a new one was twelve percent 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

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.

explain the difference in behaviors and results 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 covering the literature. We contribute to the literature by illustrating that the differences in flexibility 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 a crucial issue. 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. Furthermore, as with the degree of flexibility in decision making, the out-of-model factors that are different and usually not considered in various search experiments might be crucial (e.g., the timing of decisions or recall settings). Therefore, experimenters should be careful with such out-of-model factors, and our results provide a rationale about what 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 method, may not be appropriate for the risk elicitation method, at least for the active and quasi-active designs. Thus, researchers need to thoroughly consider when investigating the relationship between subjects' risk preferences and their search behavior.

These findings can potentially be applied to various real-world market environments. Consider job search in the form that people search for multiple candidate firms one by one (i.e., active search), and another form of search includes the cases where people

⁷ The detailed discussions are given in Section 5.

use agents to find firms (i.e., passive search). Similarly, in the marriage market, people search for a partner from multiple candidates or use a matchmaker for an arranged marriage (such as the traditional Japanese custom of *Miai*). In shopping platforms, consumers may search for products yourself, or the platform may offer products (e.g., *BuyBox* on Amazon). Those difference in search form is discarded in the search model, and our results suggest that in such markets with multiple search forms, planners can control people’s behavior by using different forms of search. When considering such a market, for example, variety of choice, price, and efficiency have been the main topics, we argue that the form of search is also important.

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 the environment to make it easier for people to 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 at a slight 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 addressed in 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 rejects 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 the moment, 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

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

The above expression can be written as

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

which yields,

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

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 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:

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

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 $Eu(\cdot)$. Then, from the arguments above, we have

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

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 (duration of the search). Because each individual stops with probability $1 - F(x^*)$ (in which $w > x^*$), the expected search duration can be simply summarized by the following expression:

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

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.⁸⁹ The total number of subjects was 103, consisting

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 comput-

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 within-subject design, and each of the three search treatments in Part 1 was repeated 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 2,000 JPY (about \$18.24).¹⁰ The subjects were informed of their reward for completing the entire session. The instructions are provided in the Appendix.

Before each treatment, a pre-recorded instruction clip was played, and the subjects could start their tasks at their own time after they thoroughly understood the details. 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 answered seven or eight point (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

ers (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 in which 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.

design was motivated by Caplin et al. (2011), Sugden et al. (2019), and Kittaka and Mikami (2020).¹¹ 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 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, and the presence 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

¹¹ 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).

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.¹² 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).

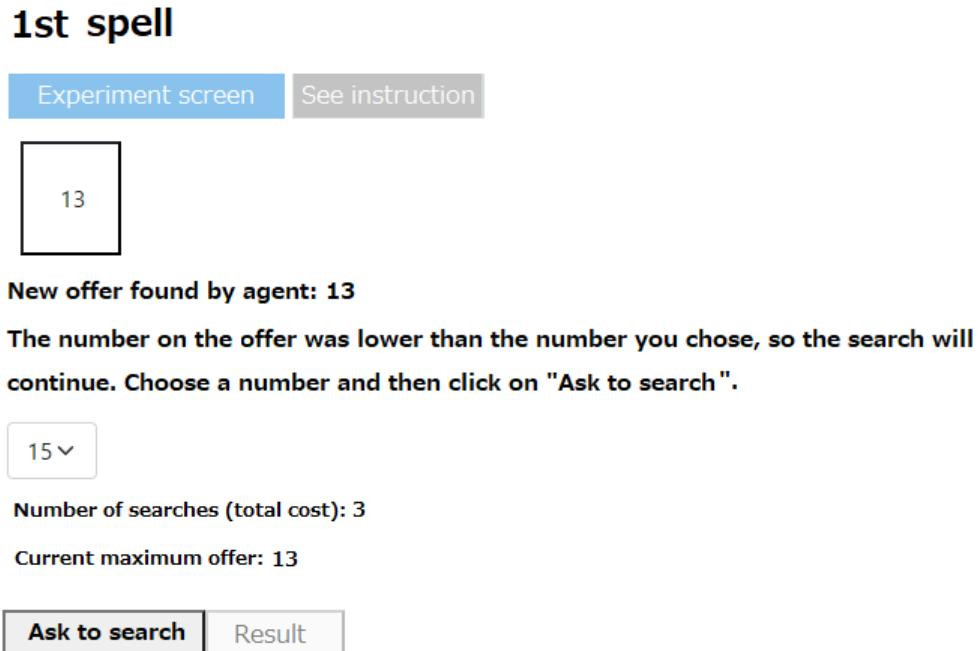


Figure 1: The main screen of the passive search design

Figure 1 is the screen shot 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

¹² This type of procedure is called the BDM procedure by Becker et al. (1964).

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 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.¹³¹⁴

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 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, and they can terminate the search at any time even if the current

¹³ Casner (2021) allowed subjects to recall, which is similar to our study. In 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 no recall setting. Hence, we allow recall to compare the results among three designs.

¹⁴ Unlike in the other two treatments, recall is costly in the passive search design. The costly recall setting could lead to a different (and possibly not constant) 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 at the model section, and the intuition is identical, although the costly recall complicates the problem, similar to all previous studies mentioned above, we employ this design to compare the design to others.

1st spell

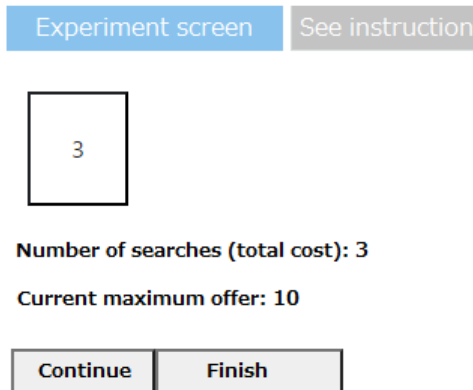


Figure 2: The main screen of the quasi-active search design

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 (values are written on the back of it) are displayed, and they are free to choose one. Specifically, in the active search design, each subject is presented with 16 available offers and a terminating option 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. Therefore, subjects can draw as many offers as they want. 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

1st spell

Experiment screen See instruction

?	?	?	?
?	?	4	?
?	1	?	10
?	?	?	?

Number of searches (total cost): 3
Current maximum offer: 10

Continue Finish

Figure 3: The main screen of the active search design

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.

Figure 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 (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

Experiment screen
See instruction

600円(5%), 300円(95%) A B 900円(5%), 0円(95%)

600円(10%), 300円(90%) A B 900円(10%), 0円(90%)

600円(15%), 300円(85%) A B 900円(15%), 0円(85%)

600円(20%), 300円(80%) A B 900円(20%), 0円(80%)

600円(25%), 300円(75%) A B 900円(25%), 0円(75%)

600円(30%), 300円(70%) A B 900円(30%), 0円(70%)

600円(35%), 300円(65%) A B 900円(35%), 0円(65%)

600円(40%), 300円(60%) A B 900円(40%), 0円(60%)

600円(45%), 300円(55%) A B 900円(45%), 0円(55%)

600円(50%), 300円(50%) A B 900円(50%), 0円(50%)

•
•
•

600円(100%), 300円(0%) A B 900円(100%), 0円(0%)

OK

The symbol "円" means the Japanese Yen (1 Yen
≐ 0.009 USD).

Figure 4: The screen of the multiple price list task (We add a note and omit some parts)

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 which 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).

Figure 4 displays a part of contents 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 600 JPY with probability p or 300 JPY with probability $1 - p$; if they choose lottery B (risky lottery), they receive 900 JPY with probability p or 0 JPY with probability $1 - p$. The probability p , beginning with five percent, increases by five percent 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 (315 JPY and 45 JPY, respectively). In the MPL we utilized, the top nine rows, the expected payoff of the safe lottery is greater than the risky one. The expected payoff of the safe lottery is equal to the risky one in the tenth row, so the risk-neutral subjects switch from Lottery A to B on tenth or eleventh rows. Because risk-averse subjects tend to choose a more safe lottery, they would switch A to B earlier than the tenth row. Hence, we use the number of safe lotteries chosen as a proxy for each subject's risk aversion.

3.5 Discussion: Search design using strategic method

We have a remark on the search design. Search designs mentioned above, in which players respond sequentially to the outcome of each search, are called *direct response methods* in the context of game theory experiments. Such methods are the majority in search experiments. Another well-known method in the context of game theory experiments is the *strategy method*, in which the subject determine the responses to all possible outcomes in advance. Theoretically, subjects have one reservation value, and therefore these differences in method do not affect the experimental results. There are several search experiments based on the strategy method. Sonnemans (1998) and Asano et al. (2015) conduct an experiment based on the strategy method in their experiments (however, the difference with the direct response method is ambiguous). We conducted a pre-experiment that included the design based on the strategy method and found potential differences from other designs. However, including the strategy method requires more large 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

Because we are interested in determining whether there is a difference among the designs, we first compare the results at the aggregate level analysis to verify the presence of a difference.

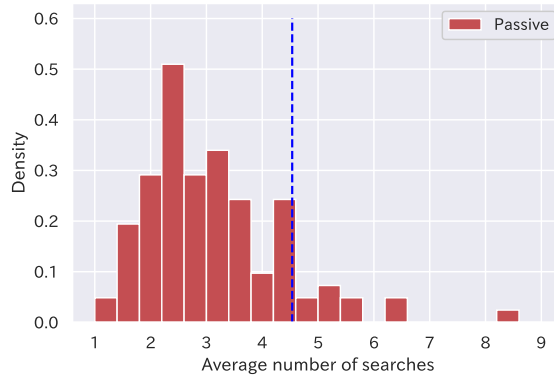
Table 3: Summary of data

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

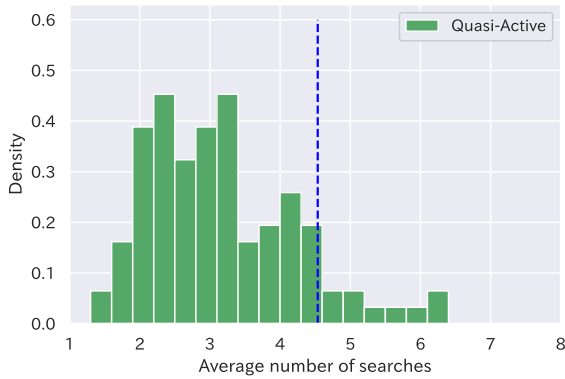
Note: Standard errors are presented in parentheses. The “Theoretical” columns present model-predicted values. “Lower” and “Upper” columns show the mean of model-predicted lower and upper values using observed risk aversion parameters elicited by the risk elicitation task (we consider two values because 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 can not be identified.

4.1 Main results

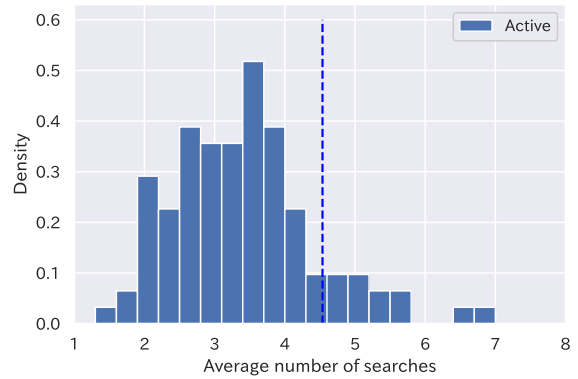
Table 3 reports the summary statistics for each treatment. The average payoff (the maximum offer found minus the total search cost in each spell) was slightly lower in the passive treatment, but there were no statistical differences across treatments. However, we find that the average number of searches differed significantly across treatments. Specifically, the average number of searches in the active search treatment was signifi-



(a) Passive treatment



(b) Quasi-active treatment



(c) Active treatment

Figure 5: Histograms of the average number of searches for each treatment (the dashed line represents the theoretically predicted number of searches for risk neutral subjects).

cantly higher than in the other treatments at the 1% significance level.^{15 16}

At the individual level analysis, the average number of searches was also different

¹⁵ We also conduct a non-parametric test for the average number of searches. The Mann–Whitney U (Wilcoxon rank-sum) test for differences between the active and quasi-active treatments was significant at the 5% level and was significant at the 1% level in the active and passive treatments. Conversely, there was no significant difference between the quasi-active and passive treatments ($p = 0.1001$).

¹⁶ In the passive treatment, the top 10 spells with the highest number of searches were significantly high (the average number of top 10 spells is 26.00 in the passive treatment, while it was 18.30 in the active treatment and 19.90 in the quasi-active treatment); the average number of searches in the passive treatment was 3.006 when removing the data that exceed the maximum number of searches in other spells (21 in both active and quasi-active treatments).

in the active and both the quasi-active and passive treatments at the 5% level.¹⁷ This implies that the design of the information acquisition process is a crucial aspect in search experiments, and that making one’s own decision regarding potential offers affects the search decision.

Figure 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 and 2.3 and 3.3 in the quasi-active treatments, while it was around 2.3 in the passive treatment. These results suggest that, among all experimental designs, the passive treatment suppresses the number of searches the most overall.

We observe an “under-search” problem as in previous studies, where subjects searched less than the theoretically predicted number of searches (“Theoretical” columns in Table 3). From the risk elicitation task (see Table 4), we found that 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 in the risk-neutral theoretical value. Nevertheless, even when we considered the risk-averse subjects, the numbers of searches observed in the tasks were still lower than the theoretical value (“Lower” and “Upper” columns in Table 3). However, subjects’ behavior was the closest to the theoretical predictions in the active search treatment. We thus argue that adjusting the degree of flexibility of subjects’ decision-making may be an effective remedy to partially solve the under-search problem.

¹⁷ The Wilcoxon signed-rank sum test for differences in the active and quasi-active (passive) treatments was significant at the 5% level, while it was not significant in the quasi-active and passive treatments.

Table 4: Observations in risk the 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.

4.1.1 Discussion: Potential differences in the passive treatment

So far, we have discussed the differences in the results among designs in terms of the degree of decision-making flexibility. Regarding the results of passive treatment, we need to consider other factors beyond theory as the reason for the difference.

First, each subject must write her reservation value before each search and has to terminate the search as long as the offer exceeds the reservation value in the passive treatment. Namely, the subject has to commit their search *ex-ante*. Such a pre-commitment may cause subjects to be more reluctant to search. Second, subjects have to carry out an additional search in order to return to the previous offer in the treatment. Although subjects never choose recall in theory, such a recall design may make the subject search more or less than in other treatments in the experiment. As we note in Footnote 16, there were some spells in which the number of searches was longer only in the passive treatment. And 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 to terminate the search.¹⁸ We did not control for the those points, but it should be noted that these points can increase or decrease the average number of searches.

Third, because the passive treatment is the only one requiring forward-thinking, one might think that the design requires more effort (and thus more time) for each search decision. We will discuss this issue in Section 5.

¹⁸ See Section 4.3. In the section, we discuss the possibility of heuristics, which tend to appear in the passive treatment, possibly increasing the variance of the number of searches.

4.2 Behavioral differences

In this section, we consider what might have caused the differences in the results between the designs, as seen in the previous subsection, from the perspective of subjects' behavior.

4.2.1 Spell-to-spell transitions and learning

Before considering 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, Figure 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.

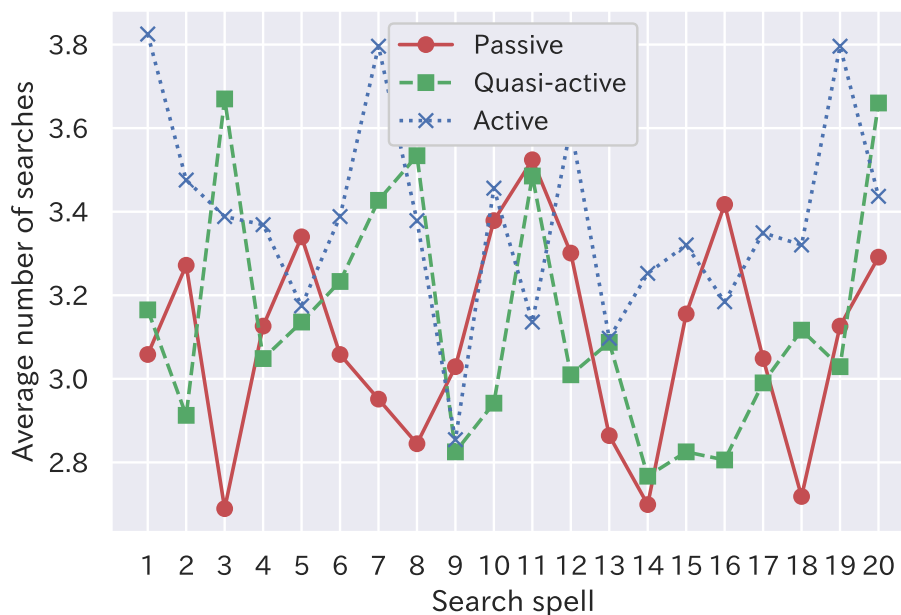


Figure 6: The relationship between experience and the average number of searches

4.2.2 Search strategy

Some previous studies (e.g., Brown et al., 2011; Casner, 2021) based on the passive design suggest that subjects exhibit a downward trend in reservation value, rather than following a search strategy based on constant reservation value, as described in Section 2.

Here, we investigate whether or not subjects’ search strategies are based on the constant reservation value in all designs. We further explore whether the tendency of search strategies differs across designs, that is, whether it drives differences in the average number of searches 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 she decided to continue the search instead of terminating the search at the previous drawn offer value, then she would never choose the past offer value again. We define a subject as exhibits “recall” if the subject terminates the search even though the last offer is equal or less than the previous maximum offer.

Table 5: Summary of data regarding recall

Treat.	Number of Recall	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; therefore, 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.

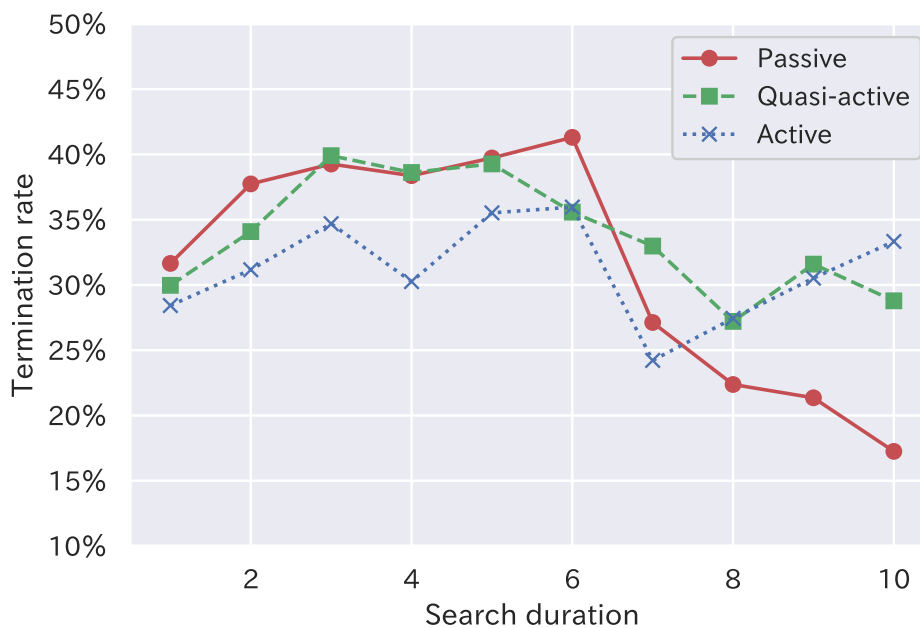


Figure 7: Percentage of subjects terminating the search per period across treatments

Figure 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. 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

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

rate was constant for a spell with up to six searches and decreased slightly for spells with six or more duration in a row. That is, the variance in the reservation value was relatively large compared with the other 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 on the first search. Columns 3 to 6 shows 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 of testing.

Table 6 shows the fraction of subjects whose termination rate changed with the search duration compared with the initial (baseline) termination rate. Regarding whether there was a difference in the termination rate, we found no significant difference for the search duration up to the fifth search within treatments at the individual level, except for 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 low termination rate of search in the passive treatment, we analyze the relationship between risk preferences and termination rate.

4.2.4 The effect of risk preferences

We create “termination rate” as a dependent variable, which takes the value of 0 if the subject continues the search in each search decision and takes the value of 1 if the

subject terminates. We analyzed this dependent variable by the five regression models, using three types of dependent variables, which are described below. We used an 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 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 a logistic regression regarding 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 impact of the number of searches is in Columns (2) and (3). From Column (2), we could not confirm an overall trend that subject changed their search strategy throughout the search, which is consistent with previous 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 design, more risk-averse subjects tended to be more reluctant to search, unlike in the other designs. Given the statistical difference in the termination rates between the active and passive treatments, this result implies that the reason for the lower average number of searches in the passive treatment is that more risk-averse subjects are more likely to terminate their search in each search decision.

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

Covariates	Coefficients (SE)				
	(1)	(2)	(3)	(4)	(5)
Passive	0.102*** (0.036)		0.188*** (0.052)		-0.229 (0.145)
Quasi-active	0.083** (0.036)		0.094* (0.054)		-0.062 (0.142)
Number of safe lotteries				0.012* (0.007)	0.001 (0.010)
Passive \times Number of safe lotteries					0.025** (0.011)
Quasi-active \times Number of safe lotteries					0.011 (0.010)
Number of searches		-0.005 (0.005)	0.007 (0.009)		
Passive \times Number of searches			-0.027** (0.012)		
Quasi-active \times Number of searches			-0.003 (0.013)		
Constant	-1.202*** (0.033)	-1.124*** (0.101)	-1.225*** (0.044)	-1.306*** (0.030)	-1.215*** (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 variables are dummy variables. Number of safe lotteries is the number of safe options subjects chosen 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.

4.3 Consistency with previous studies

Thus far, we found no evidence that subjects had a decreasing reservation value trend at an individual level, unlike previous studies (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 a long-term termination rate trend must increase. However, Figure 7 and Column (3) in Table 7 show that the termination rate is lower later in the search, which is inconsistent in 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, 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 trend of decreasing the reservation value in some spells, similar to Brown et al. (2011) (but the percentage was much smaller than theirs). Further, the average reservation value selected in those spells was significantly higher than in spells where the reservation value was constant (significant at the 1% level).

To see how the downward trend of the subjects' reservation values shapes, we see the values that the subjects choose at the beginning of each search spell.

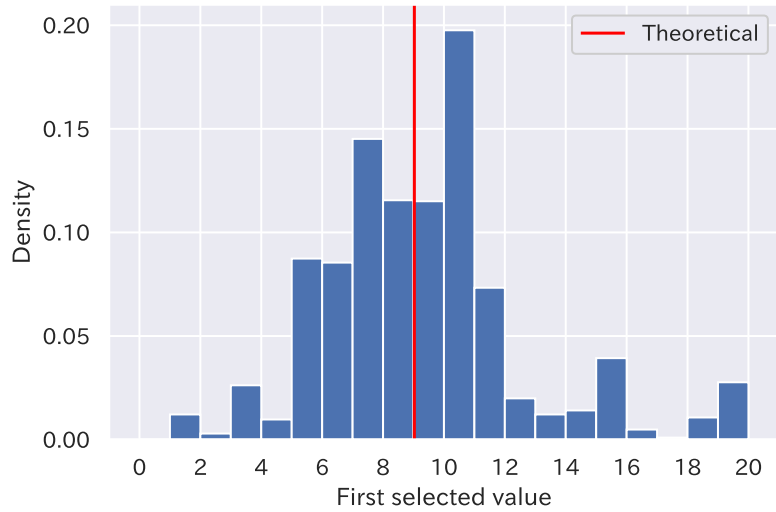


Figure 8: Histograms of the first selected value

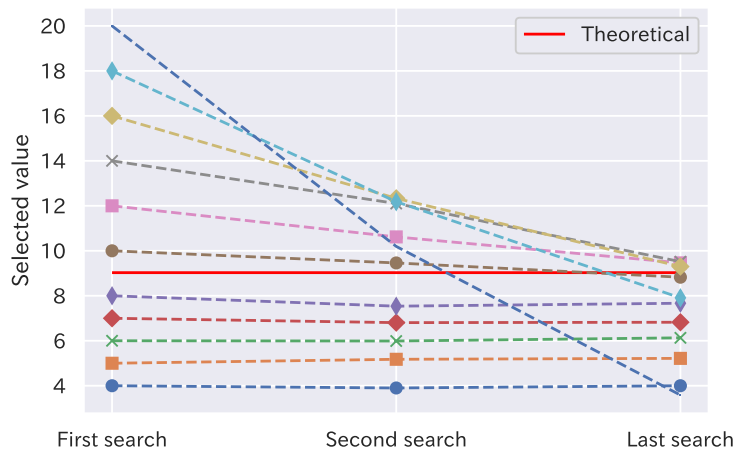


Figure 9: Trends of selected values classified according to the first selected value

Figure 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 are some spells choosing higher

values such that 16 or 20. Also, Figure 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 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 much 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, in which the first selected value was 20. Such observations are consistent with those in Figure 7 that termination rates are higher in 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 reservation value in previous studies. However, in the present study, such behavior was not pronounced at the individual level and did not have a significant influence overall. Although it cannot be stated with certainty whether this trend is unique to the passive treatment, Figure 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

Active search design: An active search design may be an advantageous experimental method in situations where we want to measure the impact of some environmental

²⁰ Such a heuristic in which the subject chooses a reservation value higher than the theoretical value for the first few times, and later lowers it, 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 learning.

change, such as marketing strategy or policy. Because the magnitude of the outcome values is more important than the detailed process of search behavior, the under-search problem may have serious consequences in the situations.

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 great attention. Therefore, for purposes other than comparison with previous studies, using a passive design with similar results on the degree of under-search but more advantages is more recommended.

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 will contribute to the study in economics and management. However, the design causes the most severe under search problems. In addition, as described below, the out-of-model factors and the MPL impacts mentioned below would need to be considered, especially in this design. Hence, improving the passive design is a nice challenge for the future.

5.2 Methodological implications

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

Out-of-model factors: Although we did not find a significant difference in the average number of searches between the quasi-active and passive treatments, the result of regression analysis (Table 7) suggests that the outcomes of those two designs are potentially different. As noted in Section 4.1.1, there are several differences in the passive search design such as elapsed time, costly-recall, forward-thinking, and pre-commitment

may have caused subjects, which could result in different outcomes (no-recall setting should also be considered). One might think that those out-of-model factors canceled out and masked the difference (recall that Footnotes 15 and 16 imply that the difference between the two treatments was pretty sensitive). While the settings concerning these factors often seem disregarded in previous studies, we propose that these factors be controlled. Furthermore, comparisons with previous studies with different experimental designs are probably problematic, and moreover, 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 would require a much larger sample size).

Time spent searching: The passive treatment is the only one requiring forward-thinking; one might think that the design requires more effort (and thus more time) for each search decision. In Brown et al. (2011) and Casner (2021), they investigated the effect of elapsed time and found that longer search may result in fatigue, which could lower reservation values. Since we only have flawed data on time due to factors specific to online experiments, the following discussion is not a rigorous argument.²¹ But our data did suggest that the search duration in the passive search treatment was not longer than in the active search treatment. Thus, in the sense of Brown et al. (2011), fatigue in longer search (and thus shorter search duration) may be rather less noticeable in the passive treatment than other treatments. Furthermore, we found no significant difference in the average search duration between the quasi-active and passive treatments, which implies that the degree of flexibility in decision making is more important rather than that the time spent unique to the passive treatment suppressed the search duration. However, we emphasize that the fatigue effect may be easily masked, and subsequent search experiments need to consider the impact of such decision-making time

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

(this impact may be mitigated by less lengthy search spells, as in our experiment).²²

Termination rate: 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 to the passive search design. Note that the use of 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).

Risk elicitation method: As experiments in the active search design, both Schunk and Winter (2009) and we 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. Unlike the former two studies, Miura et al. (2017) introduced no-recall setting and search termination possibility (common in labor search experiments). Those results suggest that the MPL method is sensitive to those factors and may result in uncontrollable outcomes. Therefore, experimenters should avoid using the usual MPL method as a risk elicitation method for comparisons or at least 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 (our passive treatment results also suggest it). Exploring the appropriate risk elicitation tasks for search experiments will be fruitful future research.²³

²² We are planning to conduct an additional experiment to address this issue.

²³ See Csermely and Rabas (2016) for investigation and comparison the MPL method and its derivatives.

6 Conclusion

Accumulating 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 cover the 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 individual and aggregate levels across the designs. The average number of searches for the active search design was highest and closest to the model-predicted value and lowest for the passive search design. Prevalence of recall and termination rate analysis results support that most subjects search according to a constant reservation value strategy within treatments. However, there was a difference in the termination rate between the passive and active treatments. Unlike the quasi-active and active treatments, more risk-averse subjects tend to be more reluctant to search in the passive treatment, which may explain the difference between the designs. Other various out-of-model factors might also be influencing the results. We, therefore, argue that researchers should be cautious in designing search experiments. 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 forms (e.g., job market, matchmaking market, 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*, Vol. 76, pp. 583–618.
- Asano, Takao, Hiroko Okudaira, and Masaru Sasaki (2015) “An experimental test of a search model under ambiguity,” *Theory and Decision*, Vol. 79, pp. 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*, Vol. 52, pp. 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*, Vol. 23, pp. 1069–1099.
- Becker, Gordon M, Morris H DeGroot, and Jacob Marschak (1964) “Measuring Utility by A Single-response Sequential Method,” *Behavioral Science*, Vol. 9, pp. 226–232.
- Bhatia, Sudeep, Lisheng He, Wenjia Joyce Zhao, and Pantelis P Analytis (2021) “Cognitive Models of Optimal Sequential Search with Recall,” *Cognition*, Vol. 210, p. 104595.
- Braunstein, Yale M and Andrew Schotter (1982) “Labor Market Search: An Experimental Study,” *Economic Inquiry*, Vol. 20, pp. 133–144.
- Brown, Meta, Christopher J Flinn, and Andrew Schotter (2011) “Real-time Search in the Laboratory and the Market,” *The American Economic Review*, Vol. 101, pp. 948–974.
- Caplin, Andrew, Mark Dean, and Daniel Martin (2011) “Search and Satisficing,” *The American Economic Review*, Vol. 101, pp. 2899–2922.

- Casner, Ben (2021) “Learning While Shopping: An Experimental Investigation into The Effect of Learning on Consumer Search,” *Experimental Economics*, Vol. 24, pp. 238–273.
- Charness, Gary, Uri Gneezy, and Brianna Halladay (2016) “Experimental Methods: Pay One or Pay All,” *Journal of Economic Behavior & Organization*, Vol. 131, pp. 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*, Vol. 9, pp. 88–97.
- Cox, James C and Ronald L Oaxaca (1989) “Laboratory Experiments with a Finite-horizon Job-search Model,” *Journal of Risk and Uncertainty*, Vol. 2, pp. 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*, Vol. 53, pp. 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*, Vol. 87, pp. 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*, Vol. 35, p. 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*, Vol. 96, pp. 1043–1068.

- Greiner, Ben (2015) “Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE,” *Journal of the Economic Science Association*, Vol. 1, pp. 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*, Vol. 76, pp. 181–221.
- Hey, John D (1987) “Still Searching,” *Journal of Economic Behavior & Organization*, Vol. 8, pp. 137–144.
- Holt, Charles A and Susan K Laury (2002) “Risk Aversion and Incentive Effects,” *The American Economic Review*, Vol. 92, pp. 1644–1655.
- Holzmeister, Felix and Matthias Stefan (2021) “The Risk Elicitation Puzzle Revisited: Across-methods (in) Consistency?” *Experimental Economics*, Vol. 24, pp. 593–616.
- Janssen, Maarten CW and Alexei Parakhonyak (2014) “Consumer Search Markets with Costly Revisits,” *Economic Theory*, Vol. 55, pp. 481–514.
- Jhunjhunwala, Tanushree (2021) “Searching to Avoid Regret: An Experimental Evidence,” *Journal of Economic Behavior & Organization*, Vol. 189, pp. 298–319.
- Kittaka, Yuta and Ryo Mikami (2020) “Consumer Search and Stock-out: A Laboratory Experiment,” *Discussion Papers at ISER Osaka University*.
- Kogut, Carl A (1990) “Consumer Search Behavior and Sunk Costs,” *Journal of Economic Behavior & Organization*, Vol. 14, pp. 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*, Vol. 7, pp. 167–183.
- Lippman, Steven A and John Joseph McCall (1979) *Studies in the Economics of Search*, Vol. 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*, Vol. 17, pp. 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*, Vol. 1, pp. 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*, Vol. 101, pp. 900–926.
- Schotter, Andrew and Yale M Braunstein (1981) “Economic Search: An Experimental Study,” *Economic Inquiry*, Vol. 19, pp. 1–25.
- Schunk, Daniel (2009) “Behavioral Heterogeneity in Dynamic Search Situations: Theory and Experimental Evidence,” *Journal of Economic Dynamics and Control*, Vol. 33, pp. 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*, Vol. 71, pp. 347–360.
- Snowberg, Erik and Leeat Yariv (2021) “Testing the Waters: Behavior Across Participant Pools,” *The American Economic Review*, Vol. 111, pp. 687–719.
- Sonnemans, Joep (1998) “Strategies of Search,” *Journal of Economic Behavior & Organization*, Vol. 35, pp. 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*, Vol. 168, pp. 1–23.
- Tversky, Amos and Daniel Kahneman (1992) “Advances in prospect theory: Cumulative representation of uncertainty,” *Journal of Risk and uncertainty*, Vol. 5, pp. 297–323.
- Weitzman, Martin L (1979) “Optimal Search for the Best Alternative,” *Econometrica: Journal of the Econometric Society*, Vol. 47, pp. 641–654.
- Zwick, Rami, Amnon Rapoport, Alison King Chung Lo, and AV Muthukrishnan (2003) “Consumer Sequential Search: Not Enough or Too Much?” *Marketing Science*, Vol. 22, pp. 503–519.

Appendix A (For Online Publication)

We conducted our experiment using oTree (Chen et al., 2016) and 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 in a 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.



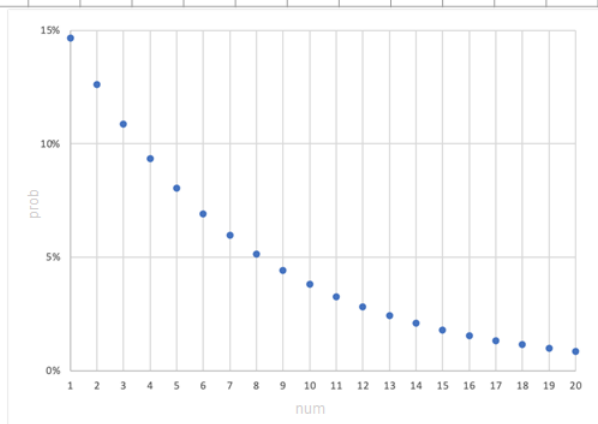
The cards have numbers on them, and the results of each session depend on the numbers you find during the experiment.

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 to the number of previous cards found.

²⁴ All experiments were conducted in Japanese (those instructions and screens provided in the appendix are translated into English).

num	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	計
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 in which 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 ($=6[\text{maximum point}] - 3[\text{total cost}]$).

Display on the screen

The screenshot shows an experiment interface. At the top, it says "1st spell" in bold. Below this are two buttons: "Experiment screen" (highlighted in blue) and "See instruction" (grey). A box contains the number "2". Below the box, it says "New offer found by agent: 2" and "The number on the offer was lower than the number you chose, so the search will continue. Choose a number and then click on 'Ask to search'". There is a dropdown menu with a downward arrow. Below that, it says "Number of searches (total cost): 1" and "Current maximum offer: 2". At the bottom, there are two buttons: "Ask to search" (highlighted in black) and "Result" (grey).

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

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

1st spell

Experiment screen

See instruction

0

Number of searches (total cost): 0

Current maximum offer: 0

Continue

Finish

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

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 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 600 yen and 90% chance of getting 300 yen. B: 10% chance of getting 900 yen and 90% chance of getting 0 yen. Please choose which lottery you prefer, A or B.

1st spell

Experiment screen See instruction

?	?	?	?
?	?	82	?
?	4	50	?
?	?	?	?

Number of searches (total cost): 3

Current maximum offer: 82

Continue Finish

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

Experiment screen

See instruction

- 600円(5%), 300円(95%) A B 900円(5%), 0円(95%)
600円(10%), 300円(90%) A B 900円(10%), 0円(90%)
600円(15%), 300円(85%) A B 900円(15%), 0円(85%)
600円(20%), 300円(80%) A B 900円(20%), 0円(80%)
600円(25%), 300円(75%) A B 900円(25%), 0円(75%)
600円(30%), 300円(70%) A B 900円(30%), 0円(70%)
600円(35%), 300円(65%) A B 900円(35%), 0円(65%)
600円(40%), 300円(60%) A B 900円(40%), 0円(60%)
600円(45%), 300円(55%) A B 900円(45%), 0円(55%)
600円(50%), 300円(50%) A B 900円(50%), 0円(50%)
600円(55%), 300円(45%) A B 900円(55%), 0円(45%)
600円(60%), 300円(40%) A B 900円(60%), 0円(40%)
600円(65%), 300円(35%) A B 900円(65%), 0円(35%)
600円(70%), 300円(30%) A B 900円(70%), 0円(30%)
600円(75%), 300円(25%) A B 900円(75%), 0円(25%)
600円(80%), 300円(20%) A B 900円(80%), 0円(20%)
600円(85%), 300円(15%) A B 900円(85%), 0円(15%)
600円(90%), 300円(10%) A B 900円(90%), 0円(10%)
600円(95%), 300円(5%) A B 900円(95%), 0円(5%)
600円(100%), 300円(0%) A B 900円(100%), 0円(0%)

OK

are determined.