



 Universität Trier



Quantitative Finance
and Risk Analysis

Working Paper Series N° 20 - 02:

Pre-Decisional Information Acquisition: Do We Pay Too
Much for Information?

Marc Oliver Rieger

Mei Wang

Daniel Hausmann

Pre-decisional information acquisition: do we pay too much for information?

Marc Oliver Rieger¹ Mei Wang² Daniel Hausmann³

Abstract

It is a common phenomenon that people tend to acquire more information in a decision task than a rational benchmark would predict. What is the reason behind this? To answer this question we conducted an information acquisition experiment that has been carefully designed to disentangle several plausible reasons for information overpurchasing before decision-making. A within-subject experiment with a simple basic information acquisition task on an investment project, equivalent formulated lotteries, estimations of probability, and an additional option to satisfy one's curiosity was used to test five different potential reasons. The results show that overpurchasing of information can be explained nearly entirely by systematic information-processing errors (*misestimation* or incorrect Bayesian updating). Other factors, such as *overoptimism* on the validity of new information, *risk aversion*, *ambiguity aversion*, and *curiosity* for (irrelevant) information, play at most a minor role. Our results imply that overinvestment in information acquisition can be mostly avoided if more detailed information is given to decision makers on how much (or little) further information can improve the decision quality.

Keywords: sequential information acquisition, ambiguity, Bayesian updating, financial decision-making.

¹ Corresponding Author: Marc Oliver Rieger. Fachbereich IV, Department of Business Administration, University of Trier. Email: mrieger@uni-trier.de

² WHU-Otto Beisheim School of Management, Chair of Behavioral Finance, Vallendar, 54296 Germany.

³ Department of Psychology, University of Zurich, Switzerland.

1 Introduction

We often need to collect and evaluate information before making final decisions, e.g., to choose between different investment options. Although further information can be valuable, it is often costly and noisy. Acquiring further information does not necessarily lead to optimal decisions. On the contrary, one should not collect further information if the cost of collecting additional information is higher than the potential benefit (e.g., Edwards, 1965). Therefore, it is crucial to know the optimal time to stop information acquisition.

Searching for information is usually a prerequisite for a successful decision (Edwards, 1965). It is a common phenomenon that people tend to search for much more information than a rational benchmark would predict (e.g. Connolly & Serre, 1984; Hershman, 1970; Newell & Shanks, 2003; Newell, Weston & Shanks, 2003; Zwick, Rapoport, Lo & Muthukrishnan, 2003; Diehl, 2005; Hausmann, Christen & Läge, 2006).

Various behavioral heuristics and biases have been suggested to explain over- or under-acquisition of information. For example, people deviate systematically from normative probabilistic thinking based on Bayesian updating rules by either ignoring the base rate or being too conservative to update belief (Tversky & Kahneman, 1974; Grether, 1980). In the former case, they overestimate the validity of new information, leading to overacquisition of information. In the latter case, people undervalue new information, and would make decisions based on insufficient information acquisition. Behavioral biases such as overconfidence and overoptimism about private prior information may also lead to a disregard of further information and therefore to underacquisition of information (Moore and Cain, 2007; Weinberg, 2009), whereas overoptimism about the value of new information can also lead to overacquisition of additional information.

While in the past many experiments have been done to investigate conditions under which oversearching seems to be stronger or weaker, these studies typically had rather complicated

decision situations where it is often impossible to compute optimal decisions that could serve as a benchmark (e.g., Hausmann & Läge, 2008). It is therefore difficult to measure the precise extent of oversearching and sometimes even difficult to prove the occurrence of oversearching. In this paper we follow a slightly different road and try to measure oversearching in a parsimonious experimental design where it is possible to compute the optimal searching behavior. The use of this optimal searching behavior as a rational benchmark is pivotal in determining the reasons for deviations. We also carefully design our experiment so that we can disentangle the most important potential factors that may explain oversearching. In particular, we test five plausible reasons for oversearching:

- *overoptimism* about the quality of new information,
- *misestimation of probabilities*,
- individual differences like *risk aversion* or
- *ambiguity aversion*, and finally
- simple *curiosity* about new information.

We will discuss the theories behind these reasons in more details in Section 2.

The main idea of a parsimonious and controlled experiment is to reduce the information search process and the choice to its minimum. This means in our case: to make a decision in a financial scenario *with* or *without* further information (evaluations). This binary choice about the information request will be compared with choice tasks about equivalent lotteries where the (real or estimated) validity of acquired information is translated into the probability to win. Participant-specific behavior – searching too much (overpurchasing) or too less (underpurchasing) – will then be compared with different, but specific benchmarks. We will explain this in details in Section 3.

The experiments that we conduct (described in Section 4) show that overpurchasing of information is mostly a consequence of overestimating the probability that the additional information improves the chances for an optimal decision (Section 5).

We will finally summarize the results and their limitations in the concluding Section 6.

2 Potential explanations of information overpurchasing

Information acquisition decisions involve several steps, each of those potentially affected by biases or preferences. Fig.1 summarizes these steps. We will now take a closer look at each of them separately.

Figure 1: Illustration of the decision process. An optimal decision about information purchasing needs to take into account the validity of the information, estimate the probability that it turns out to be useful, assess its expected utility, and finally come to a decision. In reality, a number of biases and preferences will influence this process, as illustrated in this figure.



2.1 Overoptimism about the quality of new information

Overestimation or overconfidence with implications to judgment and decision making has been demonstrated, e.g. for unrealistic optimism about future life events (Weinstein, 1980), for estimation involving complex or hard tasks (Grieco & Hogarth, 2009), for rare events in judgment tasks (Barron & Yechiam, 2009), and for investors in financial markets (Allen &

Evans, 2005), and it can be a source of incorrect judgments, e.g. diagnostic errors in Medicine (McIsaac & Hunchak, 2011; Berner & Graber, 2008).⁴

If a person has the choice to acquire new information that might (or might not) be helpful for reaching an optimal decision, the probability with which this new information turns out to be helpful is obviously an important factor. In experiments, this probability is often (directly or indirectly) given. But still, a person might be overly optimistic about his or her luck to obtain a useful information. Forecasts are also often estimated too narrow (Klayman, Soll, Gonzalez-Vallejo & Barlas, 1999). This overoptimism about the quality of potential information might therefore lead to oversearching, as the quality of the information is systematically overestimated. This effect should be particularly strong when the validity of the information is not well known (i.e., when ambiguity exists).

2.2 Information processing mistakes and misestimation

Even if a person is not overoptimistic, he or she might simply misestimate the probability for information to be useful, since in general (including in our own experiment) the probability that the information is useful (even after considering searching costs) is not provided. When aggregating the prior probability and new information, people tend to deviate from the Bayesian updating model by using mental short cuts, such as representative heuristics, availability heuristics, anchor and adjustment (Tversky & Kahneman, 1974). The new information may be either over- or underweighted, depending on the saliency and other factors. In most previous experiments, the true probability, however, is not even precisely known to the experimenter, as most experimental setups are too complicated to allow for its computation (e.g. Gilliland, Schmitt & Wood, 1993; Saad & Russo, 1996; Zwick, Rapoport, Lo & Muthukrishnan, 2003; Moore & Cain, 2007; Blavatsky, 2009).

⁴ Moore and Cain (2007), however, raise objection against the generality of overconfidence (see also Weinberg, 2009; Grieco & Hogarth, 2009).

In our experiment, however, this probability *can* indeed be computed. Although the setup is as simple as possible, the computation of this probability is still quite difficult and can therefore not be performed by the subjects during the experiment. We will see that our results suggest that information processing mistakes lead to systematic overestimation of the usefulness of information and seem to be the main causes of oversearching.

2.3 Individual Preferences and risk aversion

In decision experiments, and in particular in information searching tasks, it is often implicitly assumed that subjects are risk-neutral, i.e. that their goal is to maximize the expected value. While this viewpoint, given the relatively small payments that subjects usually obtain, might be justified from a normative point of view (Rabin 2000), empirical evidence is strongly suggesting that actual decisions are different (Kahneman and Tversky 1979, Tversky and Kahneman 1992). Moreover, risk neutrality becomes even from a normative point of view questionable if we assume that subjects do not aim to maximize their actual payoff from the laboratory experiment, but instead behave according to the actual framing of the experimental questions where decisions are usually framed in much larger monetary amounts.

But how can risk aversion (seemingly) lead to oversearching? In experiments, obtaining more information usually reduces the probability of a wrong decision. Even if this reduction is so costly that the expected value of the outcome is decreased, the *risk* is reduced at the same time. Thus for a risk averse person it might be perfectly rational to keep searching for information, while a risk neutral person would already stop. When using a risk neutral model as a benchmark, the risk averse subjects would *seemingly* oversearch.

2.4 Ambiguity aversion

Ambiguity aversion describes the fact that many people tend to avoid situations that involve risks of an unknown probability (Ellsberg, 1961). As Ellsberg demonstrated with a classical paradox, this preference cannot be compatible with the axioms of rational decision making in

the spirit of expected utility theory. In information search tasks, ambiguity aversion could theoretically contribute to oversearching, since any additional information reduces uncertainty.

2.5 Curiosity about new information

A further explanation of oversearching is that people are just curious about incomplete and hidden information, or an unchosen alternative (van Dijk & Zeelenberg, 2007; for an overview see Loewenstein, 1994). To fill the “information-gap” and satisfy their curiosity people are willing to pay for additional information, even if they do not assume this information to be useful enough to justify its acquisition. While we all experience this effect from time to time (e.g., when searching for information on the web and finally ending up with acquiring interesting information of which we, however, know that it will not be useful for our decision making), it will turn out that within our experiment we can exclude this explanation for oversearching.

3 Experimental design

The information search tasks in the previous literature tend to be too complicated to calculate a rational benchmark. We therefore simplified the task of information search to its minimum. Our participants were asked to choose between two different locations for opening a new branch office based on certain independent evaluations (cues) where one of the two locations is recommended. Each evaluation has a specific probability of being correct (validity or trustworthiness) and has a certain cost. To reduce the complexity of the task, the participants were facing only two options: (a) stop searching, i.e. to buy only one evaluation; or (b) continue searching, i.e. to buy two further evaluations in addition to the first evaluation. Appendix D shows typical tasks in the experiment, Appendix E the parameters used. After deciding whether to continue searching or not, the participants then chose the location based on either one cue (if they bought only one evaluation) or three cues (if they bought all three evaluations).

When comparing the frequency with which subjects decided to acquire more cues (denoted as S_1) to the rational risk-neutral benchmark S_0 (computed in Appendix A), we find as expected a strong amount of oversearching.

We let the participants also estimate the overall probability that based on three evaluations the prediction would be correct. We denote this estimated probability as P_e and the correct probability as P_0 .

Next we presented the subjects with a choice between two lotteries: one had the outcomes and their respective probabilities of the investment decision (in the case where no more cues have been acquired), the other one had the outcomes (including the costs) with their respective probabilities for the case where more cues had been acquired (P_0). We denote the frequency that subjects chose the second lottery as S_2 .

We also presented the subjects with two lotteries where we replaced P_0 with P_e , the estimated probability. We denote the frequency of subjects choosing the second lottery as S_3 .

In this way, we can eliminate several factors that may explain oversearching one by one:

- If $S_1 > S_2$, the sole reason for oversearching cannot be *risk preferences* alone, since the real payoff distribution for the lotteries and the information search task are identical, however subjects prefer to search more, but do not choose the associated lottery.
- If $S_1 > S_3$, taking into account *probability misestimation* (i.e. *information processing mistakes*) can still not explain oversearching, since now the *estimated* payoff distribution for the lotteries and the information search task are identical.

To distinguish between *information processing mistakes* and *over-optimism* we can look at the estimated probabilities P_e for the usefulness of information: if $P_e \approx P_0$, this would point to over-optimism: the subject knows that the odds aren't good, but thinks he will be lucky. If, however, we find over-estimation ($P_e > P_0$) for this probability, this points to information processing mistakes.

With introducing unknown probabilities (in form of a probability range) for some of the tasks we can calculate the difference to the basic evaluation task with fixed or known probabilities. If *ambiguity aversion* causes oversearching, we would expect it to be higher in the ambiguity case. We can test this difference empirically as well.

By offering an additional option, namely seeing all three evaluations while deciding on the basis of only one, we can test the relevance of *curiosity*: if this option is never or rarely chosen, even though its costs are set to be much lower than actual collection of more information, curiosity cannot play a significant role in explaining oversearching.

To summarize: We created a within-subject design with three different tasks (a basic evaluation task, equivalent lotteries for testing *risk aversion*, and a probability estimation task for testing *misestimation*), two different conditions (known versus unknown probabilities; see also Appendix A) for testing the influence of *overoptimism*, and three decisional actions (choosing on the basis of one versus three evaluations with different probabilities and search costs for testing the rational benchmark; and an additional option for testing the relevance of *curiosity*). The Tables B.1 and B.2 in the Appendix B summarize this conceptual setup. Also see Appendix C for a concrete example of the experimental questions encountered by the participants.

4 Methodology

4.1 Participants

$N=93$ undergraduate students from the University of Zurich took part in the experiment (46% of them were male). They had a mean age of 25.6 years ($SD = 4.2$). Most of the participants majored in psychology or economics.

4.2 Stimuli and design

The experiment was programmed in Microsoft Visual Basic 6 and run on PCs in a university lab. Participants were asked to make a series of decisions about opening a branch in one of two Asian cities with fictional names. For each decision problem, participants were provided with recommendations of one of the two cities with a specific validity. In the following text, we will refer the recommendations as cues, and the corresponding cue validity was simply defined as probability of the recommendation being correct. The participants can obtain the first cue (i.e., recommendation) with no cost. Then they decide whether they want to choose the city based on the first recommendation, or whether they want to obtain the second *and* third recommendations with additional predefined costs before they make the final choice.

To simplify the experimental setting, the validities of the second and third cues were always identical, but could differ from the validity of the first cue. We had two conditions for the validities of the second and third cues:

- (1) In the unambiguous condition, the validity could be .65, .70, .80, or .90,
- (2) In the ambiguous-condition, the possible ranges of validity were .50–.70, .50–.80, or .50–.90.

In our setting, the second and third cues are useful if they both recommend a different city than the first cue. In contrast, if the second or third cue agrees with the first, the additional information is useless, because in such cases, the subject should follow the first recommendation anyway. Appendix A provides the mathematical formulations of the setting and calculation of corresponding probabilities of winning for each condition. After each choice (either on the basis of one or three recommendations), the probability that the chosen city is more profitable was computed according to a random number generator and Bayes' Rule, assuming all three cues are independent regarding the probability of being correct. The costs for obtaining two further recommendations varied from 2000 to 80,000 EC (Experimental

Currencies). To control for curiosity effect, participants were allowed to pay a small amount (amounts from 200 EC to 4000 EC) to inquire the second and third recommendations after the decision has been made based on only the first recommendation.

4.3 Procedure

The first type of tasks was the *information acquisition* task (compare also Appendix D): Our participants were asked to imagine themselves as a manager, and they needed to choose between two locations (cities) to open a branch of their company “Newproducts.” They were told that one of the locations will yield a profit of 100,000 EC, whereas the other of zero. To explore which location is more profitable, up to three evaluations could be obtained. For each decision problem, the first evaluation was automatically offered for free. The relevant information for the first evaluation, such as the name of the recommended city and the corresponding validity (probability of being correct) were displayed on the information board. Moreover, the validities (in the unambiguous condition) or the validity range (in the ambiguous condition) of the second and third evaluation were also provided, but without revealing the recommendation. Participants then could choose one of three different options:

- a. Decide without purchasing further information: to decide on the city based on the recommendation from the first evaluation.
- b. Decide after purchasing the additional two recommendations: to obtain the recommendation information from the second and third evaluations by paying a certain amount of cost, then decide on the city based on all three evaluations.
- c. Decide without purchasing further information, but pay for “curiosity”: to decide on the city based on the first recommendation (as in option a.), but to inquire about the two additional recommendations afterwards. A small fee had to be paid for this *ex post* “curiosity,” which was much lower than the information cost in option b.

After having chosen one of these options in each round, the participants were provided with a brief feedback whether their choices were right or wrong, as well as the gross profit, the search cost, and the net profit or loss within this round.

The participants were also asked to complete the *binary lottery* tasks, where the two alternatives (purchasing information vs. not purchasing) were reformulated into two equivalent lotteries. One lottery corresponded to the not-purchasing alternative in the previous information acquisition task (i.e., deciding on the basis of the first recommendation), namely with a probability p one could win 100,000 ECs, or win nothing with a probability of $1-p$, where p is the validity of the cue. The other lottery corresponded to the purchasing information option in the information acquisition task (i.e., purchasing second and third recommendations and making decisions accordingly). In this case, the subject had to pay the amount c to participate the following lottery: to win 100,000 EC with a probability of r , or to win nothing with a probability of $1-r$, where the probability r was the probability to choose the correct city based on all three evaluations, and c were the searching costs. (For the computation of this probability see Appendix A).

All probabilities were written down as text and additionally visually displayed as colored pies within a circle. Participants were asked to choose one of these two displayed lotteries by clicking on the corresponding button of the lotteries. The chosen lottery was highlighted in yellow. After pressing a second button the amount of the earned money (information cost and the resulting gain or loss in ECs) was displayed.

The third type of tasks was the *probability estimation task*. The same information about the first recommendation (city name and validity, e.g. $p = .90$) and the next two recommendations (validities or a range of validities without revealing the city name, e.g. $q = .80$ for the second and third recommendations respectively) was displayed. Instead of a decision task, the participants were asked to estimate a probability such as: “If you make your decision only

based on the first evaluation, your choice will be correct with a probability of 90%. If you get two more recommendations (each of them will be correct with a probability of 80%), what do you think is the probability of making the correct decision?" To reduce errors, the computer only accepted a value between p (the validity of the first evaluation) and 100 %.

A total of 45 information acquisition problems and 64 lottery questions alternated in blocks of 8 to 18 trials. 10 probability estimation tasks were distributed as single trials within information acquisition task blocks, followed by the corresponding information acquisition problem. The within-block order of information acquisition and lottery trials was randomized. The arrangement of equivalent lotteries was counterbalanced for each trial.

After having finished all 119 trials, a feedback was given about the numbers of correct and wrong choices as well as the final payoffs. The amount of experimental currencies (ECs) was divided by a factor (540,000) to convert to Swiss Francs, and the participants were paid in cash accordingly. The average payment was 7.54 Swiss Francs (ranging from 0.30 to 35 Swiss Francs) plus a flat show-up fee of 10 Swiss Francs.

5 Results

5.1 Descriptive results

In total, 93 participants completed $N=4024$ information searching tasks. In 70.7% of the cases, the subjects decided to stop after the first recommendation without choosing the "curiosity option," whereas 3.2% chose to stop, but also chose the "curiosity option" and looked at the additional information after their decision. In the remaining 26.2% cases, the subjects continued to purchase the two additional recommendations before they made the final decisions.

When the information acquisition decisions were reformulated as equivalent lotteries, in 59.8% of the cases participants decided in favor of the lottery that was equivalent to stop acquiring

further information, and in 16.6% of the cases they chose the lottery that was equivalent to purchasing the second and third recommendations. In the remaining 23.6% of the cases the “purchasing” option was dominated by the “stop” option, because the cost of purchasing further information could not increase the probability of success. Since the dominance structure was transparent in the lottery questions, we assumed the “stop” equivalent lottery as the optimal choice regardless of risk aversion attitudes.

5.2 Information overpurchasing by comparing with different benchmarks

The relatively low number of curiosity decisions seems to suggest that curiosity does not contribute much to overpurchasing information. The further results show that other factors are of much higher importance:

5.2.1 Risk-neutral benchmark

We define overpurchasing as purchasing further information when the optimal strategy is to stop, and underpurchasing as stopping when it is optimal to purchase further information. When comparing the actual decisions with the risk-neutral benchmark, we find 19.5% cases of overpurchasing and 12.3% cases of underpurchasing, resulting in net oversearching (i.e. oversearching minus undersearching) of 7.2% (see Table 1), which is significantly larger than zero (paired t-test, $t(4023) = 8.09, p < .001$).

5.2.2 The role of risk aversion

What would happen if we take into account the risk aversion attitudes of the subjects? We can control for this effect by looking at the lotteries that were equivalent to the searching decisions. When taking into account the risk attitudes as reflected in the lottery tasks, we find an even slightly larger net oversearching of 10.9% (paired t-test, $t(4023) = 10.89, p < .001$). It suggests that risk preferences do not seem to explain the more frequent occurrence of overpurchasing of additional information.

5.2.3 The role of probability misestimation

In the next step, we controlled for information processing mistakes (which in our setup means probability misestimation). When we asked the participants to decide about the lotteries with the probability as given by their own previous estimate in the information acquisition task, then the lottery decisions became very similar to the decisions made in the information acquisition task. The amount of overpurchasing as compared to this lottery benchmark (which controls for risk preferences as well as for probability estimation errors) was even negative: we found a small net *underpurchasing* of 2.4% (paired t-test, $t(4023) = 2.40$, $p < .01$). This implies that when we consider the subjective probability misestimation, the systematic overpurchasing behavior disappears.

Participants also tend to slightly overestimate the probability of success with all cues, i.e., the mean estimation error was positive (1.8%, paired t-test, $t(4023) = 14.1$, $p < .001$). We also found that in cases where subjects had overpurchased, the probability misestimation was stronger (mean difference was 1.9%, paired t-test, $t(784) = 5.99$, $p < .001$).

All these results together also imply that overoptimism on the value of additional information is unlikely to be the cause of oversearching: If a subject expects to be “lucky”, i.e. is overoptimistic about the (random) results of additional information, there should be a distinct difference to the lottery choice task (where none of the two lotteries gives reason to feel “luckier” than the other).

Table 1. Information Overpurchasing with different benchmarks

Benchmark	Net overpurchase	t-value	df	p-value
Risk neutral	7.2%	8.09	4023	<.001
Risk aversion	10.9%	10.89	4023	<.001
Subjective probability	-2.4%	2.40	4023	<.01

5.3 The role of ambiguity aversion

To check for a potential influence of ambiguity aversion in our experiment, we tested whether there are significant differences in overpurchasing between the non-ambiguous and the ambiguous cases. In the non-ambiguous cases, the validity of further cues were known, whereas in the ambiguous cases, only a range of probabilities was given as the validity of further cues. We see that the average probability estimation error was bigger for ambiguous cases, as it could have been expected given their increased complexity (Table 2). When compared to the risk neutral or risk aversion benchmark, there are no significant differences of overpurchasing between non-ambiguous and ambiguous conditions. However, there is a strong increase in overpurchasing when comparing to the lottery with estimated probabilities: in the ambiguous cases, the net overpurchasing is 3%, whereas there is a net underpurchasing of 4% in the unambiguous cases. The difference between these two conditions is significant at the 1% level ($t(977)=3.25$).

This result is difficult to interpret. A possible explanation would be that ambiguity plays a twofold role: On the one hand, it induces a larger overestimation of the validity of further information and thus makes information acquisition more attractive. On the other hand, if the validity of further information is ambiguous, ambiguity aversion might make subjects refrain from choosing to continue. Both effects seem to cancel out each other, so that there is no significant effect of ambiguity aversion on overpurchasing. These two opposing effects can indeed be seen from a linear regression with overpurchasing as dependent and probability misestimation and ambiguity as independent variables (Table 3).⁵

Further experiments might address these questions in more detail.

⁵ It is, however, still difficult to explain why ambiguity leads to overpurchasing as compared to the lottery with estimated probabilities.

Table 2. Difference between ambiguity conditions

	Ambiguity	N	Mean	Std. Dev.	t-value	p-value
Probability misestimation	Yes	592	.0279	.048	5.02	<.01
	No	2452	.0155	.074		
Net overpurchasing with risk neutral benchmark	Yes	978	.09	.563	1.13	.257
	No	3046	.07	.558		
Net overpurchasing with risk aversion benchmark	Yes	978	.12	.529	1.70	.089
	No	3046	.09	.570		
Net overpurchasing with subjective probability benchmark	Yes	978	.03	.604	3.25	<.01
	No	3046	-.04	.604		

Table 3. Impacts of ambiguity and probability misestimation on information overpurchasing

Dependent variable: information overpurchasing	
Ambiguity (1=ambiguous validity)	-0.059** (-3.201)
Probability overestimation	0.092** (5.041)
R ²	1%
N	3044

6 General discussion

Our experiment allows to discriminate between a number of possible explanations for the frequently observed phenomenon of **information overpurchasing** in decision making. In our setup we found strong evidence that this is mostly driven by **information processing mistakes**. More specifically, subjects tend to **overestimate the probability** that further information

collection will help them to find a better decision. Other possible explanations contribute very little or not at all to oversearching:

In our setup, we can exclude that curiosity can explain a substantial amount of overpurchasing, since the participants had a cheap possibility to acquire complete information without using it. They therefore could satisfy their curiosity without having to continue their information acquisition. If they did continue, curiosity could not have been the main reason for that decision.⁶ This also implies that participants do believe the additional information offers added value for their decision making.

Several other potential explanations for overpurchasing also turned out to be irrelevant in our experiment. Most notably, risk preferences, curiosity, overoptimism or a simple craving for more information. In fact, we found that information overpurchasing entirely vanished when controlling for misestimation: more precisely, the decisions corresponded actually to a slightly larger amount of searching than we observed in the experiment. One could therefore say that considering risk preferences and misestimation the participants were net underpurchasing, albeit not very much. – This also implies that our experiment did not leave out other important factors that might induce overpurchasing.

Although our results are very clear-cut on the causes of overpurchasing, there is a natural limitation to our study: It might be possible that some of the alternative explanations would play a role in a different information searching task. A problem that can of course never really be ruled out in experimental research.

There might be other possible explanations of overpurchasing: for example, when decision problems are more complex or presented in different ways, other factors could become more

⁶ At least not in our experiment: we cannot exclude that in other situations, curiosity might be a larger contributor to information overpurchasing.

important. Future studies could try to address this open question by following our general experimental setup, but changing the decision problem in its degree of complexity or flexibility. Given, however, the clear result of our experiment, we would be surprised if in other information searching tasks the influence of information processing mistakes were not significant as well, in particular since more complex tasks usually lead to more, but not fewer, information processing mistakes.

REFERENCES

- Allen, W. D., & Evans, D. A. (2005). Bidding and overconfidence in experimental financial markets. *The Journal of Behavioral Finance*, 6, 108-120.
- Barron, G., & Yechiam, E. (2009). The coexistence of overestimation and underweighting of rare events and the contingent recency effect. *Judgment and Decision Making*, 4, 447-460.
- Blavatsky, P. R. (2009). Betting on own knowledge: Experimental test of overconfidence. *Journal of Risk and Uncertainty*, 38, 39-49.
- Berner, E. S., & Graber, M. L. (2008). Overconfidence as a cause of diagnostic error in Medicine. *The American Journal of Medicine*, 121, S2-S23.
- Caldwell, D. F., & Burger, J. M. (2009). Learning about unchosen alternatives: When does curiosity overcome regret avoidance? *Cognition and Emotion*, 23, 1630-1639.
- Connolly, T., & Serre, P. (1984). Information search in judgment tasks: the effects of unequal cue validity and cost. *Organizational Behavior and Human Performance*, 34, 387-401.
- Diehl, K. (2005). When two rights make a wrong: Searching too much in ordered environments. *Journal of Marketing Research*, 42, 313-322.
- Edwards, W. (1965). Optimal strategies for seeking information. *Journal of Mathematical Psychology*, 2, 312-329.
- Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, 75, 643-669.
- Enke, B., & Zimmermann, F. (2019). Correlation neglect in belief formation. *The Review of Economic Studies*, 86(1), 313-332.
- Eyster, E., & Weizsacker, G. (2010). Correlation neglect in financial decision-making, DIW Discussion Papers, No. 1104, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin.

- Gilliland, S. W., Schmitt, N., & Wood, L. (1993). Cost-benefit determinants of decision process and accuracy. *Organizational Behavior and Human Decision Processes*, 56, 308-330.
- Grether, D. M. (1980). Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95, 537-557.
- Grieco, D., & Hogarth, R. M. (2009). Overconfidence in absolute and relative performance: The regression hypothesis and Bayesian updating. *Journal of Economic Psychology*, 30, 756-771.
- Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making*, 3, 229-243.
- Hausmann, D., Christen, S., & Läge, D. (2006). Urteils-Sicherheit um jeden Preis? Verletzung der Stopregel in Abhängigkeit von der Höhe der Informationssuchkosten. In B. Gula, R. Alexandrowicz, S. Strauss, E. Brunner, B. Jenull-Schiefer, & O. Vitouch (Hrsg.), *Perspektiven psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie* (S. 45-51). Lengerich, Pabst.
- Hedesström, T. M., Svedsäter, H., & Gärling, T. (2006). Covariation neglect among novice investors. *Journal of Experimental Psychology: Applied*, 12(3), 155.
- Hershman, R. L. (1970) Deviations from optimum information-purchase strategies in human decision-making, *Organizational Behavior and Human Performance*, 5, 313–329.
- Kahneman, D., & Tversky, A. (1979). Prospect theory – Analysis of decision under risk. *Econometrica*, 47, 263-291.
- Klayman, J., Soll, J. B., Gonzalez-Vallejo, C., & Barlas, S. (1999). Overconfidence: It depends on how, what, and whom you ask. *Organizational Behavior and Human Decision Processes*, 79, 216-247.

- Loewenstein, G. (1994). The psychology of curiosity. A review and reinterpretation. *Psychological Bulletin*, 116, 75-98.
- Maner, J. K., & Gerend, M. A. (2007). Motivationally selective risk judgments: Do fear and curiosity boost the boons or the banes? *Organizational Behavior and Human Decision Processes*, 103, 256-267.
- McIsaac, W. J., & Hunchak, C. L. (2011). Overestimation error and unnecessary antibiotic prescriptions for acute cystitis in adult woman. *Medical Decision Making*, 31, 405-411.
- Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes*, 103, 197-213.
- Newell, B. R., & Shanks, D. R. (2003). Take the Best or Look at the Rest? Factors Influencing “One-Reason” Decision Making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 53-65.
- Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and frugal heuristic: Not everyone “takes-the-best”. *Organizational Behavior and Human Decision Processes*, 91, 82-96.
- Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica*, 68, 1281-1292.
- Raiffa, H., & Schleifer, R. (1972) Applied statistical decision theory. Cambridge, MA: Harvard University Press.
- Saad, G., & Russo, J. E. (1996). Stopping criteria in sequential choice. *Organizational Behavior and Human Decision Processes*, 67, 258-270.
- Tversky, A., Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases, *Science*, 185, 1124–1131.

- Tversky, A., & Kahneman, D. (1992). Advances in Prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.
- Van Dijk, E., & Zeelenberg, M. (2007). When curiosity killed regret: Avoiding or seeking the unknown in decision-making under uncertainty. *Journal of Experimental Social Psychology*, 43, 656-662.
- Weinberg, B. A. (2009). A model of overconfidence. *Pacific Economic Review*, 14, 502-515.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, 39, 806-820.
- Zwick, R., Rapoport, A., Lo, A. K. C., & Muthukrishnan, A. V. (2003). Consumer sequential search: Not enough or too much? *Marketing Science*, 22, 503-519.

**APPENDIX A. Computation of the real probability to obtain a correct decision
following three cues**

We consider the following problem: A decision between two options has to be made. We have already one cue which is correct with a probability $p \in (0.5, 1)$ (p is called the *validity* of the cue). We have the option of collecting two more cues. Both of them have the validity $q \in (0.5, p)$, where we assume that all probabilities are independent. We distinguish two cases: first, q is known *a priori*; second, q is unknown.

A.1 Validity of cues is known

If we decide to stop, it is obviously best to follow the first cue (since $p > 0.5$) and our situation can therefore be described equivalently by a lottery where we win with a probability p and lose with a probability $1 - p$. If we decide to proceed, there is only *one* possibility that we will change our mind, given the new information, namely when *both* new cues contradict the original cue. (If only one of them contradicts the first cue, we will ultimately still follow the first cue, since the validity of the second and the validity of the third cue are equal.) The two last cues could still be irrelevant if their validity is too low, such that we still better follow the first cue. In other words, the real winning probability is always at least as big as the validity of the first cue.

In the case that the second and third cue are relevant, we will choose the correct option if one of the following two cases holds:

- The first cue was correct *and* at least one of the following two cues was correct.

• The first cue was incorrect *and* the second *and* the third cue were correct. Given the independence of the probabilities⁷, the probability for the former case is $p(1-(1-q)^2) = 2pq-pq^2$, the probability for the latter case is $(1-p)q^2$. Thus the equivalent lottery has the winning probability

$$(1) p_{\text{win}}(p,q)=\max [p,2pq-pq^2 +(1-p)q^2]$$

and the losing probability

$$(2) 1 - p_{\text{win}}(p, q).$$

The outcomes of the lottery are of course lowered by twice the cost of getting one cue. We see that $p_{\text{win}}(p, q) = p$ if and only if q is sufficiently low. More precisely we compute the critical value q^* of q below which $p_{\text{win}}(p,q) = p$ as follows: set $2pq^* - p(q^*)^2 + (1-p)(q^*)^2 = p$, then we rewrite this as

$$(q^*)^2 + \frac{2pq^*}{1-2p} - \frac{p}{1-2p} = 0$$

and solving this quadratic equation in q^* we arrive at

$$q^* = \frac{-p + \sqrt{p - p^2}}{1 - 2p},$$

where we have neglected the negative solution of the quadratic equation.

A.2 Validity of cues is unknown

We suppose now that q is chosen randomly with equal probability distribution such that $q \in (0.5, p)$. First, we show that $q^* \geq 0.5$, i.e. that

⁷ To be more precise, we did not state explicitly that the probabilities of the cues are statistically independent, previous research, however, had shown that this is the generic assumption, and non-zero correlations are usually even neglected when explicitly given (Hedesstrom, 2006; Eyster, 2010; Rieger, 2010; Enke, 2017).

$$q^* - 0.5 = \frac{\sqrt{p - p^2} - 0.5}{1 - 2p}$$

is non-negative: since $p \geq 0.5$, we have $p - p^2 \leq 1/4$, thus $\sqrt{p - p^2} - 0.5 \leq 0$. The numerator of (1) is therefore non-positive. The denominator is also non-positive, thus $q^* - 0.5 \geq 0$.

When we compute the winning probability, we have therefore always to distinguish the cases where q is larger or smaller than q^* . The winning probability is hence slightly more complicated than in the first case and becomes

$$\begin{aligned} p_{\text{win}}(p) &= \frac{1}{p - 0.5} \int_{0.5}^p p_{\text{win}}(p, q) dq \\ &= \frac{1}{p - 0.5} \left(\int_{0.5}^{q^*} p dq + \int_{q^*}^p (2pq - 2pq^2 + q^2) dq \right) \\ &= \frac{1}{p - 0.5} \left(q^* p - \frac{p}{2} + \left[pq^2 - \frac{2}{3} pq^3 + \frac{1}{3} q^3 \right]_{q^*}^p \right) \\ &= \frac{1}{p - 0.5} \left(q^* p - \frac{p}{2} + \frac{4}{3} p^3 - \frac{2}{3} p^4 - p(q^*)^2 + \frac{2}{3} p(q^*)^3 - \frac{1}{3} (q^*)^3 \right). \end{aligned}$$

APPENDIX B. Experimental tasks and summary of results

TABLE B1. Tasks and variable names

Task	Choice	Variable Names
I. Stop or continue searching	binary	S_1 (frequency of information purchase)
II. Equivalent lottery with correct probability	binary	S_2 (frequency of choice of lotteries corresponding to information purchase)
III. Estimated probability	probability	P_e (estimated probability for correct choice after information purchase)
IV. Equivalent lottery with estimated probability	binary	S_3 (frequency of choice of lotteries)

		corresponding to information purchase)
	Not measured, but computed	S_0 (EV benchmark: optimal frequency of information purchase without risk aversion)
	Not measured, but computed	P_0 (probability for correct choice after information purchase)

TABLE B2. Results, compared to a rational EV benchmark S_0 (computed, not measured), $P_e > P_0$ = overoptimism

Task	Possible explanation	Results	Statistical Significance
I. Stop or continue searching	“classical” oversearching (all reasons possible)	$S_1 > S_0$	yes
II. Equivalent lottery with correct probability	oversearching considering preferences (risk aversion, probability weighting, ...)	$S_1 > S_2$	yes
III. Estimated probability	oversearching considering preferences and misestimation (e.g. through overoptimism)	$S_1 > S_3$	no
IV. Equivalent lottery with estimated probability	Probability misestimation	$P_e > P_0$	yes
Inferred from differences between ambiguous and non-ambiguous tasks	ambiguity aversion	$(S_1 - S_0)_{amb} < (S_1 - S_0)_{non-amb}$	yes
	overoptimism for ambiguity lotteries	$(P_e - P_0)_{amb} > (P_e - P_0)_{non-amb}$	yes

APPENDIX C. An example of the questions in the experiment

Let us consider the following concrete example where a hypothetical subject, Bob, is participating in our experiment. Let us focus on his answers to one set of questions. (In the experiment these questions would have been interrupted by other questions with other parameters.)

First question:

Bob has to decide between the two cities of Qifang and Hungchih. If correct, the profit will be 100'000 EC. The first cue points to Qifang. The validity of the cues is 70%. Would he want to collect two more cues (price: 10'000 EC)? Let us say, Bob decides to buy the additional cues.

Second question:

Now, Bob has to estimate the probability with which he will choose the right city (after having all three cues).

He estimates the probability to be 90%.

Third question:

Bob has to decide between two lotteries:

(A) Win 100'000 EC with a probability of 70% and win nothing with a probability of 30%.

(B) Win 90'000 EC with a probability of 78.4% and lose 10'000 EC with a probability of 21.6%.

He decides for lottery (A).

Fourth question:

Bob has again to decide between two lotteries:

(A) Win 100'000 EC with a probability of 70% and win nothing with a probability of 30%.

(B) Win 90'000 EC with a probability of 90% and lose 10'000 EC with a probability of 10%.

He decides for lottery (B).

What could we derive from his answers without knowing anything about his preferences and how to model them?

First, the probabilities and outcomes in the third question are not random: lottery (A) corresponds to his decision problem without additional cues, lottery (B) corresponds to his decision problem with the two additional cues. The probabilities of (B) have been computed using the formulas (1) and (2) in Appendix A.

He prefers lottery (A), but at the same time decided to look for two more cues in the original formulation, leading to exactly the same result as lottery (B). Irregardless of his underlying preferences, there must be something wrong with Bob's choices. Risk preferences, loss aversion or similar concepts related to risk preferences can therefore not explain his choices.

The second question, however, reveals that he overestimates the success rate of collecting more information, since 90% is much higher than the correct value of 78.4%. Could this overestimation explain his choices?

In question four, he has again to decide on two lotteries. This time the probabilities of lottery (B) are the ones from his own probability estimate of question 2. Lottery (B) therefore corresponds to the decision situation that Bob thinks he is facing. In this case he prefers lottery (B) which corresponds to collecting more cues.

Probability misestimation can therefore explain his initial decision.

Finally, curiosity might play a role: Bob might just collect more cues because he wants to know what they would say! In order to exclude this, we offered the “cheap” option of just seeing this information after the decision. Bob did not pick that, so curiosity is not the driving force behind his choice.

In summary: Bob’s probability misestimation is the best explanation for his behavior.

Our example so far only used questions with known validity. In the questions with uncertain validity, there are two more potential factors to influence Bob’s behavior: ambiguity aversion and overoptimism. If we find a significant difference in his behavior between questions with known and questions with uncertain validity, in one or the other direction, these factors could explain it.

APPENDIX D: Experimental instructions

At the beginning of the experiment, the participants were randomly placed on computers in the experimental laboratory of the psychology department. They then read the following starting message on their screens (original in German, we show here the English translation):

Dear participant,

welcome and thank you very much for making yourself available for this little experiment. All data will be treated with absolute confidentiality.

We can only tell you what we are investigating when the following series of experiments has been completely solved.

What you have to do now is written on the following page. Please press “Next” to start. After pressing “Next” they were instructed with the following explanations about the procedure of the following experiment:

Investment Decisions

The company “Newproducts” wants to expand and subsequently opens numerous new branches in the booming Asian market.

Your task as project manager in this company is to decide in which of two cities a branch should be established.

Various factors mean that an increase in profit can only be realized in exactly one of the two proposed cities.

In order to find out which city has the better conditions for an increase in profit, you can buy location evaluations from external consulting companies.

After a thorough assessment of the situation, a consulting firm commissioned in this way will make a recommendation for one of the two cities available for selection, which will have a certain validity (probability that it is correct).

These probability estimates are each based on many years of experience and are therefore very reliable.

To make a location decision, you must first obtain at least one location evaluation, but you can also commission in total three location decisions. The reliability level (probability data) of the second and third site evaluations are identical and are either equal to or lower than the reliability data of the first evaluation.

The consultancies that are being asked for advice always differ and the evaluation procedures are independent of each other!

Depending on the market and order situation, the costs for a site evaluation will vary. Between the location decisions you will be asked which of two lotteries you would prefer to play.

Then the experiment started, and the three types of tasks (investment decision, estimation of probabilities, lottery decisions) were, in randomized order, presented to the subjects.

Additional information (demographics etc.) was elicited at the end of the session.

Fig. D.1-3 show typical screenshots from the three types of tasks in the experiment (translated into English). In Fig. D.1, the subject has to choose between deciding immediately (based on the first cue with given validity), purchasing more information (namely the evaluations 2+3 with their validity shown on screen), or following the first recommendation, but also seeing for a small price what the other cues would have been (“curiosity option”).

FIGURE D.1: Sample Screen from the experiment on the information purchasing decision.

Task No.1

Two locations are available: or

The profit from a correct decision will be

	Evaluation 1	Evaluation 2	Evaluation 3
Recommendation	<input type="text" value="Pingthen"/>	<input type="text"/>	<input type="text"/>
Validity	<input type="text" value="80 %"/>	<input type="text" value="70 %"/>	<input type="text" value="70 %"/>

- I decide for the recommended location Pingthen.
- I decide to obtain two more location evaluations and pay for them:
- I decide for the recommended location Pingthen and take a look at the other location evaluations afterwards. I pay for that:

In Fig. D.2, an estimate for the success probability after receiving two more cues is elicited. These probabilities, as well as the real probabilities from the information purchasing task, are then used for questions as in Fig. D.3, reformulated in a simple lottery choice problem. See Appendix C for an example further illustrating the elicitation methodology.

FIGURE D.2: Estimating the success probability for a hypothetical information purchase.

Task No.2

Two new locations are available:
 or

The profit from a correct decision will be 100'000 \$

	Evaluation 1	Evaluation 2	Evaluation 3
Recommendation	<input type="text" value="Whanbo"/>	<input type="text"/>	<input type="text"/>
Validity	<input type="text" value="90 %"/>	<input type="text" value="80 %"/>	<input type="text" value="80 %"/>

If you decide only based on the first cue, your decision will be correct in 90% of the cases.

How high to you estimate the probability that you would make a correct decision if you get two more cues?

Enter a value in %:

FIGURE D.3: Lottery task based on actual probabilities and outcomes of a purchasing task.

Task No.10

Lotterien:

Please choose one of the following two lotteries:

<p>Variante A</p> <p>Play the following lottery:</p> <ul style="list-style-type: none">■ Win 100,000\$ with probability 65%.■ Win nothing with probability 35%.  <p>Choose: Variante A</p>	<p>Variante B</p> <p>Pay 20,000\$ and play the following lottery:</p> <ul style="list-style-type: none">■ Win 100,000\$ with probability 67.2%.■ Win nothing with probability 32.8%.  <p>Choose: Variante B</p>
--	--

APPENDIX E: Parameters in the experiment

TABLE E.1: Costs (total cost for the two potential additional cues) of tasks with known validity for first and second cue

second cue	validity of first cue			
	0.65	0.7	0.8	0.9
0.6	4000		4000	
	10000		10000	
	20000		20000	
	40000			
0.65		4000		
		10000		
		20000		
		40000		
0.7		10000	4000	
		20000	10000	
		40000	20000	
		80000	40000	
0.8			10000	4000
			20000	10000
			30000	20000
				40000
0.85				2000
				4000
				10000
				20000
				30000
				40000
0.9				10000
				20000
				30000
				40000
				50000

TABLE E.2: Costs for the additional cues in the case of unknown second probability

second cue	validity of first cue		
	0.7	0.8	0.9
0.5-1.0	4000	4000	2000
	10000	10000	4000
	20000	20000	10000
	50000	50000	30000