Turning-on Dimensional Prominence in Decision Making: Experiments and a Model*

Ayala Arad † Amnon Maltz ‡

December 15, 2020

Abstract

Many choices involve a large number of dimensions that ought to be considered before reaching a final decision. It is well accepted that decision makers sometimes place more weight on salient dimensions. But what makes one dimension more salient than another? Kőszegi and Szeidl (2012) and Bordalo et al. (2013) suggest that dimensions are more prominent if, roughly speaking, the variance of their values in the choice set is larger. We provide experimental evidence of another determinant of salience—whether dimensions are *turned-on* or *turned-off*. Intuitively, a dimension of an alternative is turned-on if its value is in the range of its attractive facet. In one study, we show that introducing a small interest rate on checking accounts may actually decrease allocations to checking accounts and increase the share of riskless investments. We provide evidence that the small interest rate highlights or turns-on the safe gains dimension, bumping up its decision weight while shrouding other considerations, such as liquidity. Consequently, choices shift from the checking account to safe investments with superior returns. In another study, social preferences expressed over two unequal allocations are reversed depending on whether a third available allocation is equal or not. In this case, the all-equal split turns-on egalitarian considerations that shift preferences toward equality even when expressed over unequal splits. We present the Turned-on-Dimensions (ToD) model that draws on Kőszegi and Szeidl (2012) and adds a discontinuous channel to the determination of decision weights, which allows to accommodate our findings.

Keywords: Dimension, Experiment, Salience, Social Preferences, Uncertainty.
JEL Codes: D03, C91.

---

*We would like to thank Neta Klein, Eli Mograbi, Dor Shlenger, Yuval Tsuk, Gilad Bar-Levav, Emily Klinger and Eliad Amsalem for their great assistance. We also thank Ariel Rubinstein, Yair Antler, Dotan Persitz, Moti Michaeli and seminar participants at the Hebrew University, Ben-Gurion University, University of Haifa, Tallinn University of Technology, BRIC5 conference, The 2nd Behavior Change Conference in Coller School of Management, Annual meeting of the Israeli Association of Economics, CESS 15th Anniversary meeting at NYU, 2020 ESA World Meeting and BREW India 2020 for helpful comments. We acknowledge support from The Coller Foundation, The Henry Crown Institute of Business Research in Israel and The Solomon Lew Center for Consumer Behavior.

†Tel Aviv University, aradayal@post.tau.ac.il
‡University of Haifa, amaltz@econ.haifa.ac.il

1 Introduction

Imagine that as New Year approaches, your employer tells you that you are about to receive a bonus of $\$2,000$. The bonus will be transferred to one of three options, according to your choice: your checking account that generates no interest, a savings plan that yields 4% yearly interest for sure or a stock that has a $50-50$ chance to go up (and earn 14%) or down (and lose 5%). Which option would you choose?

Now suppose that you are given similar options but your checking account generates a small interest rate, say 2%. Would you choose differently? And what if it yields 0.1%? We suggest that this seemingly minor change of the choice set may have large and counterintuitive effects on choice through the following psychological channel: When the checking account carries no interest, it is mostly evaluated as a liquid tool. A person who highly values liquidity is likely to choose it. When a positive interest rate is introduced, the nature of the checking account changes. Specifically, it now draws attention to another dimension: safe gains. As a result, this dimension becomes more prominent and receives larger weight, at the expense of liquidity, which is now shrouded. As the savings plan performs best on the safe gains dimension, the same person may now prefer the savings plan. Thus, our procedure suggests a non-monotonic response to the introduction of the interest rate on the checking account: It will be less likely to be chosen while the savings plan's likelihood of being chosen will increase.

In this paper we introduce, formally and experimentally, a decision process based on the idea that dimensions of a given option may be *turned-on* in the decision maker's mind, i.e., grab his attention, or *turned-off,* depending on their values and the way they are framed. If dimension k is turned-on in more alternatives than dimension j, then dimension k will be more prominent and receive a larger weight than j when evaluating the alternatives in the choice set. In the above example, the checking account had the safe gains dimension turned-off when it carried no interest and it was turned-on when positive interest was introduced. As a result, the safe gains dimension received a larger weight in the latter scenario.

Our contribution to the literature is twofold: First, we formally add the role of turned-on dimensions into a choice model that is based on the literature on salience and focusing (Bordalo et al., 2013; Kőszegi and Szeidl, 2012). As in that literature, we assume that subjective decision weights depend on context. However, our procedure places a spotlight on turned-on dimensions as the underlying feature that affects decision weights while in the above models weights are determined by the variance of the dimensions' values. As we elaborate below, the existing models are unable to accommodate our findings, which stem from the discontinuous nature of turning-on dimensions. At the same time, we do not attempt to replace these models as they capture important determinants of salience in choice. Rather, we later suggest how the two approaches may be combined into an augmented model with more predictive power. Our second contribution is in providing experimental evidence from three different contexts for the role of turned-on dimensions in determining decision makers' relative weighting.

The model we introduce, dubbed the Turning on Dimensions (ToD) model, formally introduces the intuition behind our checking account example. It consists of three building blocks: (i) Alternatives have some dimensions turned-on and others turned-off, (ii) Dimensions' decision weights are determined by the number of instances in which turned-on dimensions appear in the choice set, and (iii) Decision weights are applied uniformly to all available alternatives as in Kőszegi and Szeidl (2012). The essence of the model lies in the first of these blocks, i.e., in the notion that dimensions may be turned-on or turned-off.

What determines whether a dimension is turned-on in an alternative? According to our approach the answer lies in the value that the alternative has in that specific dimension. Specifically, we say that dimension i is turned on in an alternative x if x_i lies in the range of that dimension's attractive facet. For desirable dimensions, the attractive facet is the range of values that are strictly greater than zero. Thus, the dimension of safe gains is turned-on in the checking account when the account carries an interest rate that is larger than zero and it is turned-off otherwise. On the other hand, undesirable dimensions' attractive facet is highlighted when their value equals zero. For example, imagine that you are searching for an apartment. If one of the apartments has a pool in the same building, it emphasizes the proximity between the apartment and the nearest pool since it is literally right there, i.e., zero meters away. On the other hand, if all apartments you are considering have a pool in walking distance, but not in the building, the distance between each apartment and the nearest pool is less likely to receive much attention. In the next section, we formally define the notion of turned-on dimensions for desirable and undesirable dimensions.

The model's second building block describes the formation of decision weights. The only requirement we impose with respect to this stage is that dimensional weights are monotone with respect to the prevalence of turned-on dimensions. In other words, if dimension i is turned-on in more alternatives than dimension j, the decision weight of dimension i will be larger. Finally, in the last stage of the model, the decision maker settles his choice problem by applying the decision weights to all available alternatives. This 3-step procedure predicts that even small changes to some alternative's dimensional value can generate preference reversals among unchanged alternatives, very much like in the literature on context effects. In this literature, the addition of, say, a dominated or extreme alternative to the choice set, affects the relative subjective ranking of other alternatives in the set (Tversky, 1972; Huber et al., 1982; Simonson, 1989; Tversky and Simonson, 1993). We elaborate on the relation of our findings to these types of context effects in Section 4.

Through the lens of turned-on dimensions and their effect on decision weights, we examine results from three studies that were conducted in three different choice contexts. We show how the psychological procedure underlying the ToD model is able to explain the findings in each context and examine where other leading theories in the existing literature fall short of doing so.

Our first study follows the motivating investment example described above. It shows that turning-on dimensions may be "strong enough" to cause violations of the basic premise of monotonicity in money. Moreover, it does so in a real-life choice scenario which highlights the potential policy implications of this phenomenon. Participants are asked to imagine that they are about to receive a bonus from their employer and are requested to choose one of three payment options, namely whether the money is to be deposited into: their checking account, a savings plan that generates 4% annual interest, or a stock that has a probability of 0.5 of going up (and earning 14%) or down (and losing 5%). In the first treatment the checking account pays no interest, while in the second it generates an annual interest of 2%. The savings plan and stock are unchanged across treatments and the terms of all options are also held fixed.

In both treatments the checking account, which is the most liquid option, is chosen by a non-negligible group of participants. Standard monotonic preferences predict that increasing the interest rate of the checking account from 0% to 2% would (weakly) increase its choice share as it is made objectively better. A similar prediction, which is elaborated upon in Sections 3 and 4, is derived from the focusing and salience models (Kőszegi and Szeidl, 2012; Bordalo et al., 2013).

In contrast to these predictions, we find that a smaller percentage of participants choose the checking account when it pays a 2% interest rate. This drop in choice-share translates into a larger share of participants choosing the savings plan, but does not affect the share of participants who choose the stock. The interest-generating checking account is chosen less frequently, we claim, since it has the safe gains dimension turned-on. As a result, the safe gains dimension receives a larger overall weight in the decision problem and the savings plan, which performs best along this dimension, becomes more attractive. Liquidity, on the other hand, is shrouded and as a result receives lower weight in the decision problem. This underlying psychological mechanism is further supported by an analysis of participants' ex-post explanations, alongside findings from an experiment in which we directly elicit prominent dimensions in this context.

The second study is designed to illustrate the effect of turning-on an undesirable dimension. Participants are asked to rank three monetary allocations that will be paid out to them and to another participant. Using a between-subject design, we examine rankings in two treatments, named *equal* and *unequal*, that differ in the first allocation: In the *equal* treatment it is a split that pays 100 ILS to both participants while in the *unequal* treatment, it pays 100 to the participant himself and 130 to the other participant. The remaining allocations are identical across treatments and offer either a split of 100-140 or 100-160 (in both allocations the smaller amount goes to the participant himself). In this context, we think of inequality as an undesirable dimension that is turned-on in the presence of the all-equal split.

We compare participants' relative rankings of the two unequal allocations that are identical across the two treatments and find significant differences. Specifically, rankings are more in line with inequality minimization in the *equal* treatment compared to the *unequal* treatment. Looking into participants' ex-post explanations, we find that egalitarian considerations are more pronounced while efficiency far less pronounced in the *equal* treatment compared to the *unequal* treatment. Taken together, the findings show that the presence of the all-equal split turns-on the inequality dimension and shifts preferences in the direction of more equal allocations even when expressed over the unequal allocations. Findings from this experiment combined with those from another related experiment are consistent with the ToD model but not with the models of Kőszegi and Szeidl (2012) and Bordalo et al. (2013).

Our third study illustrates that weights can be shifted without actually changing the choice set, i.e., by framing alone. In particular, we show, in the realm of uncertainty, that explicitly mentioning a desirable dimension of a lottery (without actually changing its value) turns that dimension on and increases its relative weight in the decision process. The structure of this study is analogous to that of Study 1: We turn-on a desirable dimension in the first alternative in the choice set and show that choices shift in the direction of the second alternative that has the highest value along that dimension. The choice frequency of the third alternative is almost unchanged. The only difference between this study and the first one lies in the means by which the desirable dimension is turned-on. In Study 1 we turn a dimension on by manipulating the dimensional value, holding the frame fixed. In Study 3, we hold all values fixed and turn-on a dimension by explicitly mentioning it in the alternative's description.

Our findings (especially those from Study 3) may be viewed through the channel of priming. In fact, they demonstrate what may be dubbed as *priming through choice sets*. Priming is an activation of mental processes through subtle situational cues (Bargh and Chartrand, 2000). A large part of the priming literature focuses on prompting participants to think about a specific concept, or recollect past experiences prior to some task, and then measures how participants' behavior is influenced.[1] Our studies provide evidence for the activation of dimensional prominence in a more subtle way: Making a dimension of some alternative explicit (or changing its value to the range of its attractive facet), i.e., turning it on, primes individuals to shift weight from other dimensions over to that dimension when settling their decision problem.

The paper proceeds as follows: In Section 2 we set-up the theoretical framework and outline the main ingredients of the Tod model. Section 3 describes the experimental studies followed by the results. Section 4 discusses related theoretical and experimental literature. Section 5 concludes.

2 A Formal Derivation

We formalize the idea that turning-on a dimension will increase its weight in the evaluation of the choice set. Our model purposely ignores other factors that influence decision weights in order to focus on our suggested mechanism. We also sketch how one may add our channel of turned-on dimensions to those that have already been recognized in the literature on salience and focusing.

[1] The psychological literature on priming is vast. See Cohn and Maréchal (2016) for a recent review of priming in incentivized economic experiments.

Dimensions

Every alternative has a number of characteristics that are relevant for choice. In most choice contexts, characteristics are not spelled out explicitly as part of the alternatives' description. In some specific contexts, as in many behavioral models or lab experiments, they are spelled out and given in the form of a list. In these circumstances they are often referred to as attributes. There are also real-life examples where some non-exhaustive set of relevant dimensions are mentioned explicitly but the rest are not.[2]

In our experiments, we do not use lists of attributes since our studies lie in the domain of choice contexts in which options are normally not described in this manner. We therefore allow participants to shape the dimensions that they deem relevant for choice. To avoid confusion with the existing literature, we use the term 'dimensions' when referring to the characteristics that are relevant for choice from the decision maker's perspective. This allows us to refer to alternatives' characteristics when they are observable and when they are only partially observable, in a unified manner. Thus, if one is to apply our model to a specific context, the first step is to identify the relevant dimensions. The analyst may infer the relevant dimensions through introspection, prior knowledge or more direct elicitation methods. In the context of Study 1, we illustrate in an experiment how one can directly elicit relevant dimensions: Participants are introduced with the same background and the same investment options as in the main experiment (checking account, savings plan and stock) but are not asked to make a choice. Instead, we ask them to list the characteristics that come to their mind when they examine each option. In all experiments, we also use participants' ex-post explanations of their choices as a rough proxy for the list of relevant dimensions.[3]

The ToD Model

We follow Kőszegi and Szeidl (2012) (henceforth KS), and assume that our agent chooses from a finite set $\mathcal{C} \subseteq \mathbb{R}^K$ of K-dimensional objects and maximizes the following context-dependent weighted utility function:

$$\tilde{U}(c, \mathcal{C}) = \sum_{k=1}^{K} g_k(\mathcal{C}) \cdot u_k(c_k).$$

where $u_k(c_k)$ are the 'consumption utilities' assigned to the different dimensions, as in KS, and $g_k(\mathcal{C})$ are the menu-dependent-weights of each dimension. The difference between our ToD model

[2] Consider, for example, shopping for a camera in a department store. While the price, number of megapixels and storage space may be provided by the manufacturer (among other technical attributes), the feel of the camera and its ergonomic design are difficult to quantify and will not appear in the camera's technical description. Nonetheless, these aspects are likely to be taken into consideration by an amateur photographer.

[3] Note that ex-post explanations are much harder to interpret as highlighting noticeable dimensions since participants may discuss dimensions that rationalize their decisions even though they did not consider them when they examined the alternatives. At the same time they may neglect to mention dimensions of non-chosen options. Nevertheless, using our additional elicitation experiment in the context of investments, we claim that ex-post explanations are a reasonable proxy for the collection of relevant dimensions of the choice problem.

and the one proposed by KS comes from the argument of the weighting functions $\{g_k\}_{k=1,...,K}$ that measure the weight given to dimension k in the decision process. In KS, weights of the different dimensions correspond to their variance in the choice set, where higher variance leads to a higher weight. Using the words of KS, "the decision maker focuses more on attributes in which her options generate a greater range of consumption utility" (Kőszegi and Szeidl, 2012, p. 58). Our model suggests a different determinant for these weights, one which we believe is natural and also sheds light on the findings to follow. To formally express these weights, we first need to explain what it means for a dimension to be turned-on in an alternative. We provide two definitions, the first for desirable dimensions and the second for undesirable ones.

Definition 1: Turned-On Desirable Dimensions. We say that a desirable dimension k is *turned-on in alternative* c if $c_k > 0$.

Definition 2: Turned-On Undesirable Dimensions. We say that an undesirable dimension k is *turned-on in alternative* c if $c_k = 0$.

Applying the definitions depends on the context and relevant dimensions. In Study 1, we use the first of the two definitions as the manipulation applied across treatments is made to the interest rate of the checking account which is clearly a desirable dimension. In the context of Study 2, we refer to the second definition since we tweak the undesirable dimension of inequality. Specifically, replacing $(100, 130)$ with the all-equal $(100, 100)$ split pushes its inequality level to zero, the level for which it is turned-on. Separating the definitions into desirable and undesirable dimensions is a convenient way to express our idea formally but it is actually not necessary. We could say that every dimension, desirable or undesirable, has a range of 'attractive values', which corresponds to its attractive facet. This range is $(0, \infty)$ for desirable dimensions and it is $\{0\}$ for undesirable dimensions. If we use this terminology then any dimension $i \in \{1, ..., K\}$ is turned-on in an alternative x if x_i belongs to the range of attractive values of dimension i.

Study 3 suggests that a dimension may be turned-on in an alternative by simply describing it differently. However, a formal derivation of turning-on dimensions by framing requires reference to language rather than to numerical values which is beyond the scope of this paper. In other words, in our model framing is assumed to be held fixed while only numerical values may change. Nonetheless, in a less formal manner, in this study we treat a dimension as turned-on in an alternative if it is explicitly mentioned in that alternative's description, and turned-off otherwise.

Next, we define for every alternative c the K-vector of Turned-on Dimensions c^{ToD} by

$$c_i^{ToD} = \begin{cases} 1, & \text{if } i \text{ is turned-on in } c \\ 0, & \text{otherwise} \end{cases}$$

for every $i \in \{1, ..., K\}$. Following is our assumption on the weights.

Assumption 1 - ToD Weights. The weights g_k^{ToD} are given by

$$g_k^{ToD} = g\left((\sum_{c \in \mathcal{C}} c_k^{ToD}) / (\sum_{j=1}^{K} \sum_{c \in \mathcal{C}} c_j^{ToD}) \right),$$

and the function $g : \mathbb{R} \to \mathbb{R}$ is strictly increasing.

For a given dimension, the ToD weights are calculated by dividing the number of alternatives where that dimension is turned-on by the total number of instances of turned-on dimensions in the choice set. In Study 1, for example, the safe-gain dimension received a larger weight when the checking account's interest rate was raised from 0% to 2% (when it was 0% this dimension was only turned-on in the savings plan and thus carried smaller weight). We do not impose any additional structure on g although it is natural to concentrate on cases where $g'' < 0$ and $g(0) = 0$. The first restriction implies that turning-on a dimension in one more alternative has diminishing effects on the weight of that dimension as the number of alternatives in which that dimension is turned-on grows. The second simply states that when a dimension is turned-off in the entire set, it does not receive any weight in the decision process. We would like to emphasize that these weights are merely one technical formulation that allows us to capture the conceptual idea that turned-on dimensions affect decision weights. Our goal in this section is not to provide the best quantitative fit for actual decision weights.[4] Rather, we offer a simple formulation that captures the directional change behind our suggested mechanism.

The Tod model allows for discontinuities of weights with respect to small changes in the values of dimensions of alternatives. For example, a 0% checking account has the safe-gain dimension turned-off but a 0.1% interest rate will turn it on and increase that dimension's weight. This 'jump' in weight would be the same whenever the interest rate increases to some positive number, no matter how small. This differs from the continuous nature of weights implied by KS. In their model, if the function g is continuous, small changes to a dimension's value lead to small changes in its relative weight.

As is common in the development of theoretical models, our approach is not meant to replace the insights of the existing focusing and salience models, both of which capture important features of human behavior.[5] Moreover, it is quite obvious that the channel of turned-on dimensions is not the only one to affect dimensional weights in a given context. In fact, we believe one has to take into account our insights alongside those from previous work. A model with decision weights that consist of two components—the variance of dimensional values and the number of turned-on dimensions—is one possibility to capture both channels. Such an extension will allow for continuous

[4]In fact, our experiments are also not meant to inform us about actual decision weights. They provide a qualitative assessment of the dimensional weights and how they are affected by turning-on dimensions in the choice set.

[5]For recent experimental support of these models see Dertwinkel-Kalt et al. (2017a,b) and Dertwinkel-Kalt and Köster (2015).

effects of dimensional values based on the variance of each dimension as in KS, without compromising the discontinuities around the 'turning-on' point of these dimensions. In situations in which all dimensions are turned-on, the variance component will dominate. However, when some of the dimensions are turned-on and some are turned-off, the component reflecting the ToD procedure is likely to kick-in and influence the decision weights.

Remarks

- In our model, the overall weights sum up to 1. Thus, an increase in the weight of a specific dimension reduces the weight given to others. This feature of the model highlights the intuition that turning-on a dimension increases that dimension's prominence while it masks the other dimensions at the same time.
- Our model generalizes the standard linear utility model and it reduces to it by imposing $g = 1$. KS refer to this benchmark case as *consumption utility.*
- As in KS, our weights apply to the evaluation of all alternatives in the set. In this sense both models differ from the one proposed by Bordalo et al. (2013) where dimensions' salience, and hence their weights, may differ for different alternatives.

3 Experimental Studies

We illustrate the effect of turned-on dimensions in a wide range of choice contexts. Study 1 deals with investment decisions in which we turn-on a positive dimension (safe gains). In Study 2 we turn-on a negative dimension (inequality) in the context of social preferences. Finally, Study 3 deals with choice under uncertainty and shows how dimensions may be turned-on through framing.

Study 1: Enhancing the Checking Account in Investment Decisions
Our first study is based on the investment choice scenario described in the introduction. We believe that this real life example demonstrates the importance of the turned-on dimensions channel and carries important policy implications. The study consists of a number of experiments and examines the effect of adding an interest rate to the checking account on individuals' investment decisions.

Experiment 1.1
Participants were 201 registered panelists, who regularly participate in online questionnaires, and constitute a representative sample of the Israeli adult population. Their age range was 18 - 65 and roughly 50% were female. A link to the questionnaire was sent out and those who completed it received 5 ILS (roughly \$1.5 at the time of the experiment). It took participants on average 5 minutes to complete 2 questions, each followed by a free text explanation of their answers. Each participant was asked to imagine she/he is an employee in a firm and is about to receive a new

year's bonus of 10,000 ILS (roughly $\$3,000$). They were then asked to choose one of the following options to which the employer will transfer the money:

- Their checking account.
- A savings plan that generates 4% yearly interest.
- A stock that has a 50:50 chance of going up (earning 14%) or down (losing 5%).

Participants were randomly assigned to one of two treatments. In the *2-checking* treatment ($n = 103$), the checking account paid a 2% yearly interest rate. In the *0-checking* treatment ($n = 98$), the checking account earned no interest. All three options were explained in detail, including withdrawal options and renewal terms, and in the most realistic fashion. The savings plan allowed weekly withdrawals while the stock could be sold anytime (online or by phone). It was also stated that they may withdraw any part of the bonus before the end of the year and reap the relative profits (the full questionnaire is available in Appendix C.1). Following their choice and the explanation they provided for it, in the next question participants were asked to imagine the same scenario, except that this time they could choose the proportion of the bonus that they wanted to allocate to each option (so that they summed up to 100%).

We also ran an almost identical experiment (with minor wording changes) with a checking account that had only a tiny yearly interest of 0.1%. That is, in this experiment (207 participants) one treatment had a checking account with no interest, a savings plan and a stock (the exact same options as in the *0-checking* treatment reported above) whereas the other treatment had a checking account with a 0.1% interest rate alongside the same savings plan and the same stock. The results are very similar to those of Experiment 1.1 and therefore only the main findings from this experiment are reported.

Experiment 1.1: Results

First, note that while the checking account has a lower interest rate than the savings plan in both treatments, it has other merits (e.g., highest liquidity and most convenient withdrawal through the ATM) and is therefore not a dominated option. Indeed, a significant amount of participants in both treatments choose this option and their explanations show that they value precisely these merits. Some refer to the urgent need of liquid money (due to overdraft or other types of debt) while others mention the fact that they can invest the bonus later as they see fit because they can access it at any moment in time.

Standard consumer theory predicts a weakly higher share of participants choosing the enhanced checking account compared to the share of choices of the no-interest account. However, counterintuitively, the enhanced checking account is actually chosen less frequently. As shown in Figure I, 23% of the participants choose the checking account with no interest while only 11% do so when it generates a 2% interest (p=0.016 according to a chi-squared test). This reduction translates into

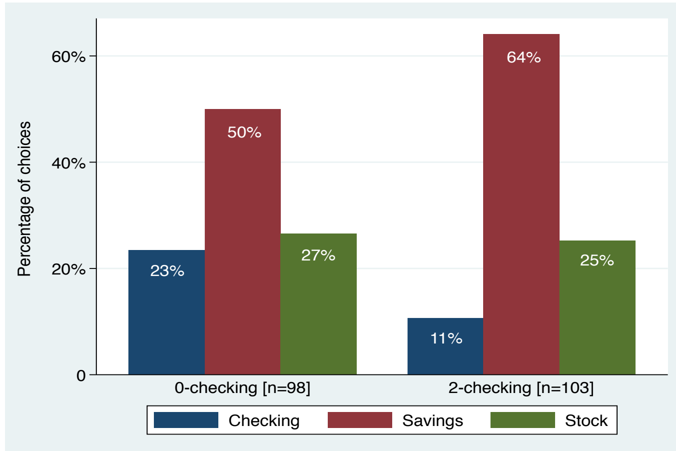


Figure I: Choice percentages of each investment per treatment in Experiment 1.1.

a significant increase in the share of participants who choose the savings plan (an increase of 15%, p=0.044), while the percentage of participants who choose the stock shows no significant change (p=0.835). In the experiment that used the tiny 0.1% interest rate, the percentages choosing the checking account, savings plan and stock in the *0-checking* treatment were 20%, 51% and 28%, respectively, while in the treatment with the 0.1% interest rate, these percentages were 8%, 66% and 26%. The percentage of choices of the checking account was significantly smaller while the percentage of the savings plan significantly larger in the latter compared to the former ($p < 0.05$ for both comparisons).

Interestingly, even the two models of focusing and salience mentioned earlier (Kőszegi and Szeidl, 2012; Bordalo et al., 2013) are unable to explain this choice pattern. Notice that increasing the interest rate of the checking account from 0% to 2% reduces the variance of the safe interest rate in the choice set. According to Kőszegi and Szeidl (2012), decision makers should now focus less on this dimension and it should receive a smaller decision weight. As a result, their model predicts that the savings plan should be chosen less frequently while the other, more liquid options, should gain popularity at its expense.[6]

The results of the second question in Experiment 1.1, where participants were asked to state the proportion of the bonus that they would like to allocate to each option, further support this pattern. Comparing the distribution (and averages) of allocations of each of the options across treatments, we find lower proportions allocated to the checking account in the *2-checking* treatment compared to the *0-checking*. This can be viewed in Figure II which shows the cumulative distribution of allo-

[6]According to Bordalo et al. (2013), increasing the interest rate would reduce the distance of the savings plan's interest rate from the average safe interest rate and hence this dimension would become less salient in the evaluation of the savings plan. It should therefore be chosen (weakly) less. At the same time, the low interest rate of the checking account would be more pronounced when it is 0, hence it should be chosen less in the *0-checking* treatment (once again "pushing" choices in a direction which contradicts our findings).

cations to the checking account across treatments.[7] The Figure shows that the CDF of allocations to the checking account in the *0-checking* treatment first order stochastically dominates the CDF of the allocations to the checking account in the *2-checking* treatment. The two distributions are statistically different from each other. The average allocation to the checking account is 25% in the *0-checking* treatment and 14% in the *2-checking* treatment (p=0.016 according to a two sample t-test). In Figure III we observe higher proportions of the bonus allocated to the savings plan in the enhanced checking treatment (56% of the bonus compared to 46%, p=0.045). Finally, in Figure IV we see no effect on allocations to the stock across the two treatment (29% and 30%, p=0.95).

In an attempt to gain insight into the psychological mechanism behind our participants' choices, we first look into participants' ex-post explanations of their choices in the first question of Experiment 1.1. For this purpose, we asked a research assistant (RA) to read the explanations and to prepare a list of categories of relevant dimensions. We examined the list ourselves and approved it with no changes or adjustments. These categories were exhaustive and reflected the various dimensions that were mentioned by our participants. Then, three RAs, including the one who came up with the list of categories, independently classified explanations into these categories (one explanation could fit into a number of categories). After their initial independent classifications, we determined the final classification by majority rule. While classifications were made separately and independently by each RA, unanimous classifications occurred for the vast majority of cases.[8]

Following the completion of the RAs work, we realized that the most frequently mentioned dimensions were safe gains, liquidity, the possibility of high returns and risk. We concentrate our discussion on the first two as the others were mentioned to a similar extent in the two treatments and hence it seems that their relative weight did not change dramatically. In Figure V we see that participants refer to liquidity more often in the *0-checking* treatment (26%) compared to the *2-checking* treatment (19%) while for safe gains the pattern is reversed (33% compared to 49% respectively). The emerging pattern is well explained by the ToD procedure. When the checking account pays no interest, liquidity receives a higher weight in the evaluation of the entire choice set compared to its weight in the *2-checking* treatment. Since the checking account performs best along this dimension, it is chosen by roughly a quarter of the participants. When it carries a positive interest rate, however, its nature as a riskless investment is apparent and it has the dimension of safe gains turned-on, which increases the weight attached to this dimension at the expense of liquidity. With this weight shift, not much is left for the checking account to show for in this context. After all, along the safe gains dimension, which is now more prominent, it is completely outmatched by the savings plan. Liquidity on the other hand, along which it performs better, is now shrouded and receives a smaller weight. As a consequence, it is chosen less frequently in this

[7] Eight participants were excluded from the calculation of the CDFs since their allocations did not sum up to 100%.

[8] This procedure was held for each of the three studies reported in the paper. In this study, the RAs classifications were aligned along 84% of the explanations. In the second and third studies, unanimous agreement was reached along 91% and 85% of the entries, respectively.

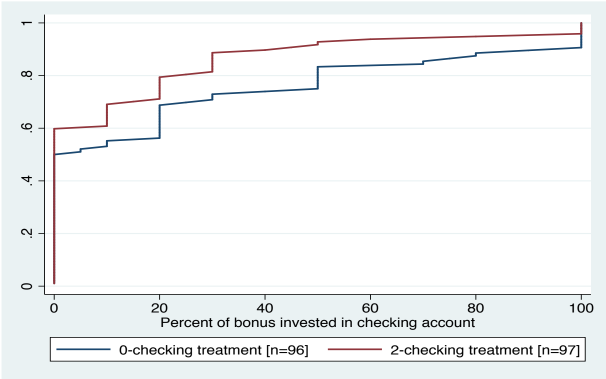


Figure II: CDF of allocation to the checking account per treatment in Experiment 1.1.

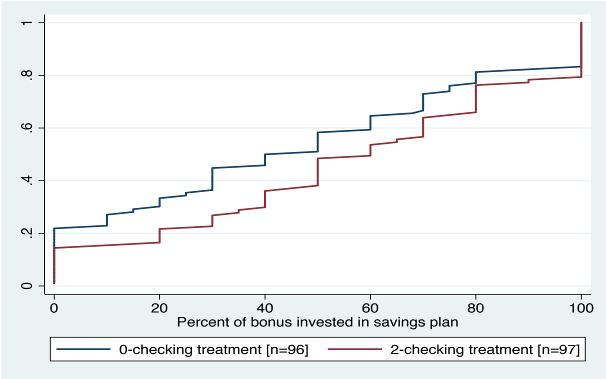


Figure III: CDF of allocation to the savings plan per treatment in Experiment 1.1.

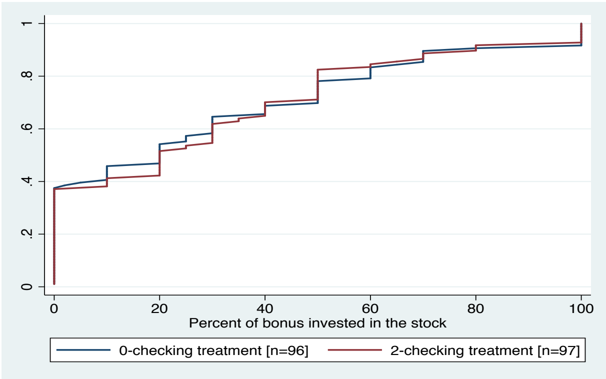


Figure IV: CDF of allocation to the stock per treatment in Experiment 1.1.

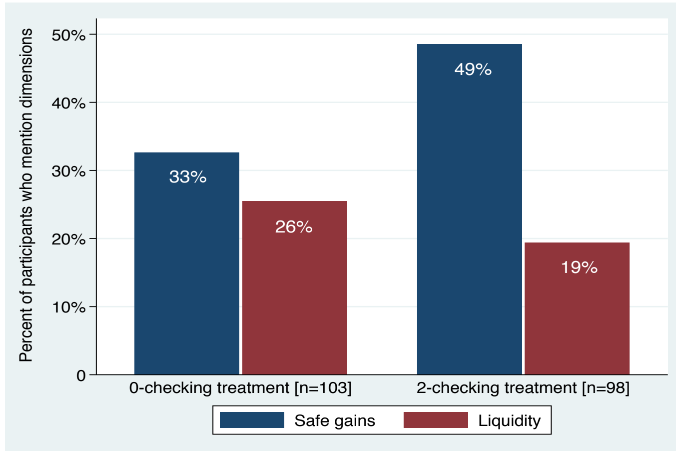


Figure V: Dimensions mentioned per treatment in Experiment 1.1.

treatment. Of course, those who still highly value the liquidity dimension, due to, say, debt or an urgent need for money, may very well choose it even in this case.

In passing, we acknowledge the fact that ex-post explanations are insufficient if one is interested in eliciting *all* relevant dimensions that were noticed by participants at the stage of contemplation of the choice set. Ex-post explanations are naturally concentrated on dimensions of the chosen alternative rather than dimensions of the non-chosen ones. In addition, it is likely that these explanations involve rationalizations rather than first glance dimensional perceptions of the alternatives–the latter being our actual interest.

To partially accommodate these difficulties, we ran another experiment with the goal of directly eliciting the dimensions that come to mind when facing the above investment options. In order to do so, we ran two treatments that are almost identical to those in our main experiment, but with one important difference: In this experiment, participants were not asked to choose. Rather, after seeing the same backstory as in Experiment 1.1, they were asked to write the characteristics that they deem as most prominent for each option (with no limitation regarding the number of characteristics). Two RAs (different than those who analyzed the ex-post explanations) assisted in analyzing the answers. In Appendix A we elaborate on this experiment, the analysis and its results.

The main dimensions elicited in this more direct method are almost identical to those that emerged in the ex-post explanations. The percentages of mentions of each dimension across treatments are also similar to those reported above (although, as one would expect, there are also some differences). Taken as a whole, the findings from this elicitation exercise inform us regarding the limitations of using ex-post explanations, but at the same time they provide support for the earlier conclusions drawn from analyzing them. Given this support and keeping in mind the potential limitations, in the following studies we carry on with ex-post explanation analysis as a proxy for our participants' perceptions of prominent dimensions.

Experiment 1.2

In this experiment, we examined the first question of Experiment 1.1 (choice of investment) with two very similar treatments except that they did not include the savings plan. 214 participants (with similar demographic characteristics to those in Experiment 1.1) were introduced with the same backstory and were asked to choose between transferring the amount into their checking account (with no-interest rate or a 2% interest rate depending on the treatment) or the stock. The purpose of this experiment was to examine whether introducing the interest rate on the checking account in the previous experiment may have actually reduced its attractiveness through some unexpected channel. For example, in Israel most checking accounts do not generate positive interest rates and some individuals may have grown accustomed to it. Having said that, we do not expect the added interest rate to make the checking account worse per-se. The ToD procedure suggests that the added interest rate will shift choices only to alternatives that have a high interest rate, as the savings plan. Therefore, in this experiment, in the absence of the savings plan, we expect to see the enhanced checking account chosen with a similar or higher percentage than the no-interest checking account.

Experiment 1.2: Results

We find that the added interest rate does not harm the checking account per-se. In fact, in both treatments over half of the participants choose the checking account (53.2% when it carries no interest and 52.4% when the 2% interest rate is added to it). In other words, the enhanced checking account (with a 2% interest rate) is not deemed worse than the one with no interest rate.

Study 1: Conclusion

Summing up, this study shows that adding a positive interest rate to the checking account may lead individuals to hold less balances in that account and, instead, allocate balances to other riskless assets. As discussed above, standard theory cannot accommodate these findings, nor can the salience and focusing models of Kőszegi and Szeidl (2012) and Bordalo et al. (2013). In Appendix B.1, we show that the ToD model is able to accommodate these findings and, in fact, generates forces that 'push' in the direction of this behavioral pattern independently of the dimensional utility functions of safe gains and liquidity (as long as they are monotonic and continuous).

Study 2: Social Preferences in the Presence of an Equal Split

Study 1 dealt with turning-on a desirable dimension by changing its value from 0 to a positive value. We now turn to show how an undesirable dimension may be turned-on when its value is shifted in the opposite direction, i.e., from a positive value to 0. For this purpose, we chose the context of social preferences and explored how replacing an unequal allocation with an all-equal split of a pie turns-on the undesirable dimension of inequality. This study consists of two experiments that highlight the role of turned-on dimensions on behavior and point to the ToD procedure, rather than the focusing and salience models, as the mechanism behind the results.

Experiment 2.1

Participants in this experiment were 393 registered panelists. As in Experiment 1.1, they constitute a representative sample of the Israeli adult population. Their age range was 18 - 65 and roughly 50% were female. A link to the questionnaire, which included one question followed by a free text explanation, was sent out and those who completed it, did so in about 3 minutes and received a participation fee of 3 ILS (roughly $0.9 at the time of the experiment). In addition, it was explained in the instructions that 5% of the participants would be randomly selected to receive additional payoffs according to their responses. Participants were randomly assigned to one of two treatments, named *unequal* ($n = 194$) and *equal* ($n = 199$). In both treatments, each participant was presented with a situation in which he, and another anonymous participant, were chosen to receive payment. He was then asked to determine the exact payment for himself and for the other participant. It was clearly stated that the identity of the other participant would not be disclosed. The complete questionnaire appears in Appendix C.2.

Table I shows the different options that were available in each treatment. To control for order effects, each treatment had two opposing orders of the three options.[9] Options b and c are unequal splits that are identical in both treatments, and the first option is different: an unequal split in one treatment (option a) and an equal split in the other (option a'). In each treatment, participants were asked to rank the options from their most preferred to the least preferred. In order to incentivize the full ranking, the instructions explained that if the participant is drawn to receive payment, there is a 60% chance that their most preferred option will be implemented and a 40% chance that it will be their second most preferred option. Finally, participants were asked to provide a brief explanation for their ranking.

Our main interest is in the relative ranking of options b and c across treatments (top ranked options across treatments are also reported). Ranking b above c reflects a stronger emphasis on reducing inequality while the opposite ranking is in line with putting more weight on efficiency considerations. Notice that one does not sacrifice his own payoff by increasing the other (anonymous)

[9] To avoid confusion, we introduced the options in two naturally monotonic orders, i.e., increasing or decreasing in the other participant's payoff.

Options	***Unequal***	***Equal***
a (a')	(100,130)	(100,100)
b	(100,140)	(100,140)
c	(100,160)	(100,160)

Table I: Monetary payments by treatment in Experiment 2.1. A pair (x, y) represents a payment of x ILS to the participant himself and y ILS to the other participant (at the time of the experiment 100 ILS were roughly equal to $30).

person's payoff.[10] It is well-documented that people care both about equity (Fehr and Schmidt, 1999) and efficiency (Charness and Rabin, 2002). In line with the latter, in our experiment we expected most participants in both treatments to rank the outcome with the highest sum of payoffs, $(100, 160)$, on top, which indeed was the case. Nonetheless, we examine the difference in rankings across treatments and its relation to the nature of the first option. In the *unequal* treatment only 18% rank b above c. In the *equal* treatment this percentage rises to 32% (this difference of 14% is significant according to a chi-squared test, p=0.002). In a logistic regression reported in Table II, we control for the order of the alternatives and find a significant positive effect of the *equal* treatment on the probability of ranking b above c; the odds ratio equals 2.1 ($p = 0.002$). In other words, the probability to rank b above c divided by the probability to rank c above b doubles when the $(100, 130)$ allocation is replaced with $(100, 100)$.

Experiment 2.1: Results

In Table III we report the percentages of participants who rank each of the three options on top by treatment. This table reveals the shift of preferences from reflecting efficiency to inequality considerations across treatments, in line with the preference reversal between options b and c. A significantly larger proportion of participants rank option a' on top in the equal treatment (38%) compared to those who rank a on top in the unequal treatment (14%). The difference in proportions is reversed looking at those who rank c on top: 82% in the unequal treatment compared to only 60% in the equal one (both differences are highly significant according to a chi-squared test ($p < 0.001$)).[11]

Next, we wish to gain insight into the underlying psychological procedure using participants' ex-post explanations of their rankings, as we did in Study 1 (and keeping in mind the limitations of

[10] In fact, efficiency considerations in this set-up go hand in hand with altruistic motives. When we refer to efficiency in the discussion and in the participants' ex-post explanation analysis, we include all psychological forces supporting a larger payment to the other participant without hurting one's own payment.

[11] Overall, looking at both treatments together, 92% of the rankings were monotone, i.e., from the most efficient allocation to the least efficient one (70%) or vice versa (22%). Thus the vast majority of participants who ranked the first allocation on top actually ranked (a or a') $\succ b \succ c$ (87 out of 104). Out of the 278 participants who ranked c on top, 276 ranked $c \succ b \succ$ (a or a').

Variable	Coefficient
equal treatment	0.743***
	(0.241)
order	0.059
	(0.236)
cons	-1.542***
	(0.218)
N	393
R^2	0.0224

$^{***}p < 0.01$, $^{*}p < 0.1$

Table II: Experiment 2.1 - Results of a logistic regression in which the dependent variable equals 1 when b is ranked above c and 0 otherwise. Standard errors in parentheses.

these explanations that were discussed in Study 1). The same procedure described in Study 1 was held in this context using the good work of the three RAs. We concentrate on the two categories that were referred to the most: 'inequality' and 'efficiency'. If, as we expect, the inequality dimension is weighted more heavily in the *equal* treatment, it should be mentioned there more often compared to the *unequal* treatment. Similarly, we expect the efficiency dimension to be more prominent in the *unequal* treatment compared to the *equal* treatment because it is not shrouded by the inequality dimension. Figure VI summarizes our analysis of participants' explanations and shows that, indeed, inequality is mentioned more frequently in the *equal* treatment compared to the *unequal* treatment (26% compared to 7%) while the opposite pattern is found for efficiency (55% mention efficiency in treatment *equal* compared to 73% in treatment *unequal*).

Overall, we find that changing the value of the undesirable dimension of inequality to zero, by replacing $(100, 130)$ with $(100, 100)$, turns this dimension on and shifts weights as predicted by the ToD procedure. These findings cannot be explained by any type of stable preferences, i.e., preferences that are context independent. Notice however, that the focusing and salience models

Options	***Unequal***	***Equal***
a (a')	14% (28)	38% (76)
b	4% (7)	2% (4)
c	82% (159)	60% (119)

Table III: Percentage of participants who rank each option on top (numbers of participants in parentheses).

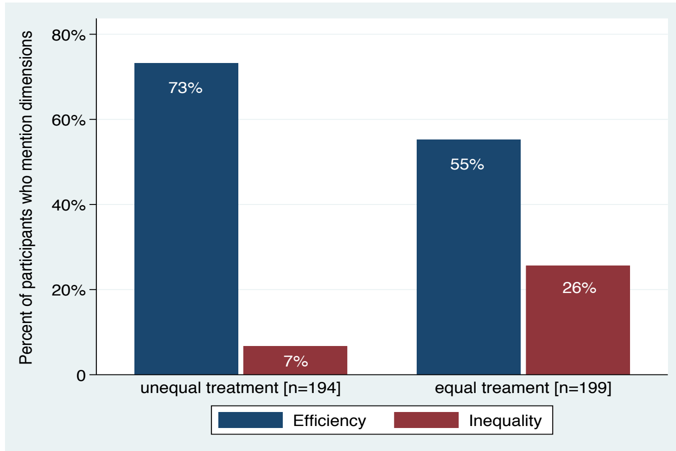


Figure VI: Dimensions mentioned per treatment in Experiment 2.1.

of Bordalo et al. (2013) and Kőszegi and Szeidl (2012) *are* able to predict them given specific parameters. This is due to the fact that replacing $(100, 130)$ with $(100, 100)$ increases the variance of both efficiency and inequality in the choice set. As a result, our findings may emerge from both models with appropriate marginal utilities along the different dimensions. The next experiment, conducted in the same context, allows disentangling the two potential explanations. It points to the discontinuous effect of turned-on dimensions as the psychological mechanism underlying our findings rather than the effect of the change in variance.

Experiment 2.2

Our additional experiment includes two treatments, the options of which are summarized in Table IV. Participants were 221 registered panelists with similar demographic characteristics to those in Experiment 2.1. Notice that the first treatment is identical to the *unequal* treatment of Experiment 2.1. Unlike that experiment, however, the second treatment also consists of unequal splits only. The only difference between treatments lies in the first unequal split. In the *unequal130* treatment the other participant will receive 130 ILS while in *unequal110* he or she will receive 110 ILS (the participant choosing the allocation will receive 100 ILS in all options).

Options	***Unequal130***	***Unequal110***
a (a')	(100,130)	(100,110)
b	(100,140)	(100,140)
c	(100,160)	(100,160)

Table IV: Monetary payments by treatment in Experiment 2.2. A pair (x, y) represents a payment of x ILS to the participant himself and y ILS to the other participant.

This experiment along with the previous one allows us to disentangle our suggested mechanism from the focusing model of Kőszegi and Szeidl (2012). According to the ToD model, in this experiment there should be no significant differences in the relative rankings of options b and c since the same dimensions are turned-on in both treatments. According to Kőszegi and Szeidl (2012), moving from the *unequal130* treatment to *unequal110*, we should expect to see the same directional shift of the relative rankings as in Experiment 2.1, albeit perhaps slightly smaller in magnitude. This is due to the fact that the change of variance of the inequality and efficiency dimensions moving from the *unequal130* treatment to *unequal110* is in the same direction as in Experiment 2.1 (it is moving the payoff of the other participant "two thirds of the way" from 130 to 100).

Experiment 2.2: Results

No significant differences arise between the relative ranking of options b and c in this experiment: 24% rank b over c when the first option is $(100, 130)$ while 17% do so when $(100, 110)$ is the first option ($p = 0.182$). Thus, replacing $(100, 130)$ with $(100, 110)$ does not have the same impact on behavior as replacing it with $(100, 100)$. If anything, there is a slight shift in the opposite direction to the one found in the Experiment 2.1.

Study 2: Conclusion

This study shows that an undesirable dimension may be turned-on by shifting its value to zero. The ToD procedure provides a mechanism through which replacing the unequal split (100,130) with the all-equal split of (100,100) may change the ranking over the other (unaltered) unequal allocations. While Experiment 2.1 may also be explained by the focusing model of Kőszegi and Szeidl (2012) with appropriate parameters for the dimensional utility functions, the combined results of Experiment 2.1 and Experiment 2.2 cannot. In Appendix B.2, we show how the ToD model predicts a preference shift that is in line with our findings from Experiment 2.1, regardless of the specific parameters of the consumption utility functions (as long as these functions are monotonic). The findings from Experiment 2.2 are also in line with the ToD model since in that experiment the same dimensions are turned-on in both treatments and the model predicts no significant difference in rankings.

Study 3: The Framing of a Lottery in the Realm of Uncertainty

In our final study, which included one experiment with four treatments, we demonstrate how framing may be used to turn-on dimensions. We illustrate this in the realm of uncertainty by using different frames for the same lottery in the different treatments. Participants in this study consisted of 243 undergraduate students from various fields in Tel Aviv University, who are registered in the IDMlab of the Coller School of Management. Their age range was 21-30, and roughly 50% were female. The questionnaire consisted of one question followed by a free text explanation and the average completion time was about 5 minutes. Participants were sent a link to the questionnaire and were asked to choose between two or three options, depending on the treatment. They were randomly assigned to one of four treatments (roughly 60 participants in each), named *certain(2)*, *certain(3)*, *lottery(2)*, and *lottery(3)*, and were instructed that 5% of them would be randomly selected to receive a prize according to their choice. Table V summarizes the options in our main treatments: *certain(3)* and *lottery(3)*. The complete questionnaire appears in Appendix C.3.

Treatments *certain(3)* and *lottery(3)*

Participants in *certain(3)* and *lottery(3)* face the exact same choice problems with one difference: In the former, the first option is framed as a certain amount (using the words 'with certainty') plus a potential bonus, whereas in the latter, the first option is framed as a state contingent lottery (probabilities and prizes) just like the framing of option b. Therefore, the certainty dimension is more emphasized in the description of the first option in *certain(3)* while the probability of obtaining the high prize of 95 ILS is more emphasized in the description of the first option in *lottery(3)*. Other lottery features, such as the probability of obtaining the low prize or the expected value are also more explicit in the state contingent frame. According to the ToD procedure, this change of frame is expected to shift weights from the certain amount dimension in *certain(3)* to these lottery features in *lottery(3)*. As a result, we expect option b, which does relatively well along some lottery features—has a high known probability of delivering the large 95 prize and a high expected value—to receive a larger share of choices in *lottery(3)* compared to *certain(3)*. The first option on the other hand, is expected to have a lower choice share in *lottery(3)* compared to *certain(3)* since the certainty dimension is shrouded in the former.

Treatments *certain(2)* and *lottery(2)*

To further investigate the ToD procedure in this context, we turn to the *lottery(2)* and *certain(2)* treatments. These are the same as *lottery(3)* and *certain(3)*, respectively, except for the fact that option b (the 50:50 lottery) is absent. Hence the difference in the weighting of dimensions should be in the same direction as in the main treatments but, in the absence of b, we do not expect

Options	***Certain(3)***	***Lottery(3)***
a (a')	60 with certainty + 35 with prob. 0.14	(0.86,60 ; 0.14,95)
b	(0.5,40 ; 0.5,95)	(0.5,40 ; 0.5,95)
c	Dow-J (30,115)	Dow-J (30,115)

Table V: Options by Treatment in Study 3. A lottery with known probabilities is described by $(p, x; 1-p, y)$, i.e., probability p of winning x ILS and probability $1-p$ of winning y. A bet denoted by Dow-J (x, y) is a bet that pays x ILS if the Dow-Jones index goes up the following day and y if it goes down. (We use the term *lottery* to describe contingent claims where probabilities are objective and known to the decision maker, and *bet* for claims with unspecified probabilities).

the share of the first option to decrease. The reason is that the lottery features, which have been turned-on in option *a'*, are not shared by *c*, the other alternative in the set. Thus, no other option, except for *a'*, is expected to gain from the larger weight given to these features, in contrast to our main treatments where we expect option *b* to do exactly that—gain from the larger weight placed on the lottery features due to the framing of *a'*.

Note that option *c* may benefit from the shift in weights along some other unforeseen channel and the combination of the four treatments in this study allows examining this possibility. This is done in a similar vein to the exploration of adding the interest rate to the checking account in the absence of the savings plan (Study 1, Experiment 1.2). Our complete hypothesis, based on the ToD procedure, states that changing the certain framing to the lottery framing, will lead to a more substantial decrease in the first option's choice share when option *b* is present than when it is absent.

Study 3: Results

A logistic model is estimated to test if the treatment has an effect on the likelihood of the first lottery (presented as *a* or *a'*) to be chosen. The probability that the first lottery is chosen is modeled by $\sigma(\tilde{Y})$ where σ is the CDF of the standard logistic distribution and $\tilde{Y}$ is specified as follows:

$$\tilde{Y}_i \quad = \quad \beta_1 lottery(2)_i + \beta_2 certain(3)_i + \beta_3 lottery(3)_i + \epsilon_i,$$

where $lottery(j)_i,\ j = 2, 3$ is a dummy variable that equals 1 if participant i was assigned to treatment *lottery(j)*, $certain(3)_i$ is a dummy variable that equals 1 if participant i was assigned to treatment $certain(3)$ and ϵ is an error term distributed by the standard logistic distribution. The benchmark treatment is taken to be *certain(2)* where participants choose between option *a*, framed as a certain amount of money plus a possible bonus, and the Dow-Jones bet. Coefficient β_1 measures the net effect of framing the first lottery as *a'*, while β_2 measures the effect of adding option *b* to the choice set without changing the frame, i.e., moving from a doubleton set (without *b*) to a triplet (including *b*). β_3 is the coefficient of the interaction variable which equals 1 when the first lottery is framed as *a'* and option *b* is present. Our main interest lies in the odds ratio implied by this coefficient, i.e., the effect of changing the frame *and* having *b* in the set on top of the two separate main effects. Formally, our main hypothesis is that the odds ratio implied by β_3 is smaller than 1, i.e., the interaction variable will have a negative effect on the probability of choosing the first lottery.

Our full results are summarized in Table VI. Our hypothesis is confirmed by the data as the odds ratio of the interaction variable equals 0.23 (p=0.007). In addition, the coefficient of adding option *b* is not significantly different from 0, while the effect of only changing the frame is actually almost significantly positive ($p = 0.059$). In other words, adding option *b* without changing the frame, or changing the frame without adding option *b*, does not negatively impact the frequency of

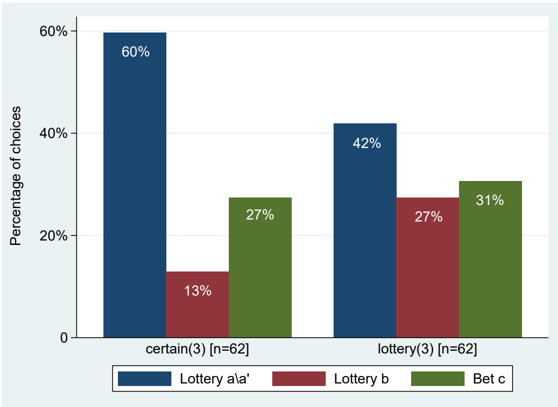
Variable	Coefficient
lottery(2)	0.762*
	(0.404)
certain(3)	-0.013
	(0.369)
lottery(3)	-1.48***
	(0.544)
cons	0.405
	(0.263)
N	243
R^2	0.0461

$^{***}p < 0.01$, $^{*}p < 0.1$

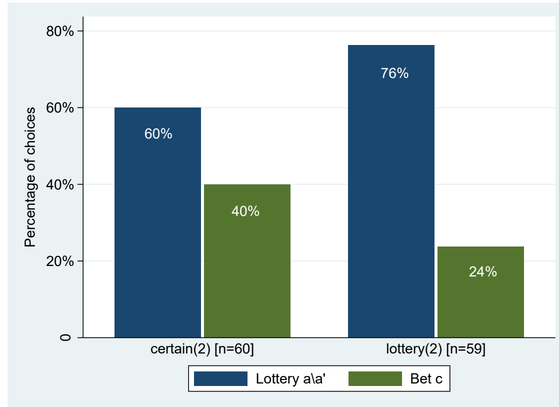
Table VI: Study 3 - Results of a logistic regression in which the dependent variable equals 1 when the first option is chosen and 0 otherwise. Standard errors in parentheses.

choosing the first option. It is only the combination of the two that increases the choice frequency of *b* at the expense of *a'*. Figure VII gives another perspective of the same effect: In panel (a) we can see that 60% of the participants choose the first option in *certain(3)* while only 42% do so in *lottery(3)*. This significant reduction (p=0.048 according to a chi-squared test) translates into an increase in the choice share of Lottery *b* (an increase of 14%, p=0.044) but does not significantly change the percentage of participants who choose to bet on the Dow Jones (p=0.692). This increase in the choice share of *b* arises despite the fact that *a'* is more popular than *a* when compared to *c* alone as shown in panel (b) (76% choose *a'* in *lottery(2)* compared to 60% that choose *a* in *certain(2)*).

Further support is given in Figure VIII. It segments the data by analyzing participants' ex-post explanations in a way that is analogous to our examination of explanations in the previous studies. Once again we focus on the two most common dimensions mentioned in participants' explanations: *certainty* (i.e., a certain amount or a sure gain) and *lottery features*. Lottery features are explanations which refer to expected values and considerations of known probabilities (as opposed to unknown probabilities) to obtain a maximal or a minimal prize. In Figure VIIIa, we see that participants in the *certain(3)* treatment mention certainty far more frequently than participants in the *lottery(3)* treatment (53% compared to 19%), while the prevalence of lottery features in the explanations is reversed (35% compared to 73%). In Figure VIIIb the same pattern is reported for the treatments *certain(2)* and *lottery(2)*. While the two figures show the same pattern of prominence shift due to framing, the first option is chosen less frequently in the lottery frame only in the presence of option *b* but not in its absence.

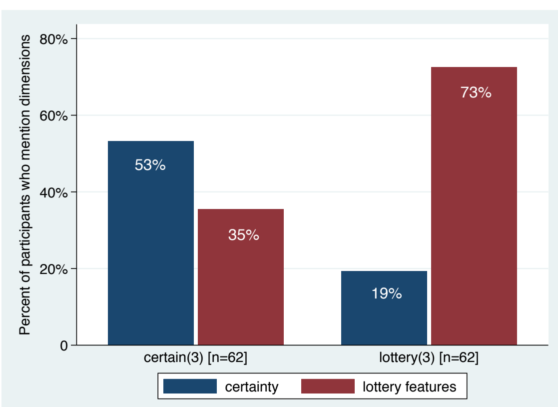


(a) Treatments *certain(3)* and *lottery(3)*

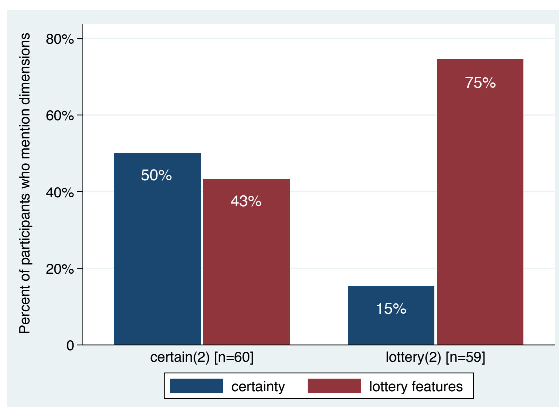


(b) Treatments *certain(2)* and *lottery(2)*

Figure VII: Choice percentages of each option in Study 3. Both panels compare the effect of framing on the choice distributions. Panel (a) does so for the choices from triplets (certain(3) and lottery(3)) and panel (b) compares the choices from binary sets (certain(2) and lottery(2)).



(a) Treatments *certain(3)* and *lottery(3)*



(b) Treatments *certain(2)* and *lottery(2)*

Figure VIII: Dimensions mentioned per treatment in Study 3.

Study 3: Conclusion
This study demonstrates the role of framing in turning-on dimensions: Explicitly mentioning a dimension brings it to the mind of the decision maker and shifts weights in its favor. An illustration of how the ToD procedure accommodates our findings from the four treatments of this study is given in Appendix B.3. While our model does not capture turning-on dimensions through framing, for the purpose of this exercise we assume that dimensions are turned-on when they are explicitly mentioned in the description of the alternative. Note that the focusing and salience models are not set up to deal with framing effects and hence cannot account for the findings from this study.[12]

4 Discussion and Related Literature

4.1 Related Theoretical Models

In this section we briefly discuss our model, alongside other approaches, in light of the behavioral patterns that arise in our studies. The closest models are those of Kőszegi and Szeidl (2012) (KS) and Bordalo et al. (2013). Both have a similar motivation as they deal with how salience, focusing, and weighting of different dimensions affect choice. We draw on the idea, which is common to both models, that some characteristics stand out more than others and receive larger weight in the assessment of goods. In Bordalo et al. (2013), the decision maker examines the dimensions of each alternative, assigning a larger weight to the dimension that is farthest away from the mean level of that dimension in the choice set. Thus, every alternative has its own salient dimension that may differ across alternatives. In KS, the decision maker's focus and weights are determined by the variation of each dimension in the choice set and they are applied uniformly to the assessment of the members of the set. In our model, weights are also applied uniformly to the entire set as in KS and hence, for the purpose of the current discussion we focus on the comparison between their model and ours.[13]

The main difference between our model and KS lies in how weights of different dimensions are determined. In KS, a dimension with a wider range will become more prominent and receive a larger weight. In the ToD model, a dimension's prominence is determined by the number of alternatives in which it is turned-on. In this sense, our model is more 'discontinuous' than KS. For example, slightly decreasing the level of some dimension of one alternative, but keeping that level positive, is likely to have some effect on its prominence according to KS but not according to ToD. By contrast, a tiny dip in the level of some dimension from $\epsilon > 0$ to 0 is likely to generate a larger effect on relative prominence in our model than in theirs.

[12] Kőszegi and Szeidl (2012) discuss potential framing effects in the context of intertemporal choice and suggest that the explicit mention of the time intervals of a payment may change the perspective through which the consumer views his future transactions. In fact, our view of framing effects in the context of turned-on dimensions follows their footsteps and, as Kőszegi and Szeidl (2012), we believe that this is an important venue for future work.

[13] The general discussion in this section would be very similar if we chose to compare our approach to the model of Bordalo et al. (2013) with only small nuances reflecting the different weighting functions.

This difference generates distinct predictions of choice behavior. For example, in Study 1, the range of the safe gains dimension decreased when we increased the interest rate of the checking account from 0% to 2%. Thus, according to KS (and according to Bordalo et al., 2013) the weight placed on this dimension should decrease and, as a result, the savings plan should become less attractive and should be chosen less frequently, in contrast to our findings. The ToD model, on the other hand, will place a larger weight on this dimension since it is now turned-on in the checking account, whereas it was turned-off in that alternative in the *0-checking* treatment.

As another example, consider Study 2. Here the two natural dimensions are inequality and efficiency. Turning from the *unequal* treatment to the *equal* one, the range of both dimensions increases: there is a larger gap in terms of inequality and efficiency between (100,100) and (100,160) than between (100,130) and (100,160). Thus, it is difficult to derive a sharp prediction based on KS as to which dimension becomes more prominent. This will depend on the specifics of the weighting function and marginal utilities along the different dimensions. But notice that replacing (100,100) in the above discussion with (100,110), would maintain the same forces in play according to KS. Thus, KS would predict a shift in ranking that is qualitatively the same in both cases, (100,100) and (100,110), although possibly smaller in magnitude for (100,110). This prediction contradicts our findings from experiment 2.2, in which (100,130) is replaced with (100,110) rather than the all-equal split and the replacement does not generate a significant difference in behavior.[14]

The ToD model, on the other hand, accurately predicts the preference shifts in both experiments (2.1 and 2.2). First, according to the ToD procedure, a larger weight will be placed on egalitarian considerations when (100,130) is replaced with (100,100) as the latter completely eliminates inequality. As a consequence, preferences are expected to shift and express a stronger positive attitude toward equity, regardless of the dimensional utilities, as in experiment 2.1. Second, the model predicts no significant change when (100,110) takes the place of (100,130) due to the fact that in both sets the same dimensions are turned-on. Therefore, decision weights, and consequentially final rankings, are not expected to change.

An approach closely related to focusing and salience, which is interesting to examine in light of our findings, is that of relative thinking. Bushong et al. (2017) derive a model that formally resembles KS but assumes that the decision maker places less weight (rather than more weight as in KS) on dimensions with larger variance of consumption utility.[15] Using the authors' example, the model predicts that the difference between losing 12$ and losing 13$ dollars will loom larger when the range of possible losses is 13$ compared to when the loss range is 25$. While relative thinking, as focusing and salience, is an important phenomenon of human behavior, it is unable to accommodate our findings. As in the case of focusing, we believe that the reason lies in the discontinuous nature of our findings, which is reflected by the ToD procedure, but is not incorporated by the relative

[14]The insignificant shift that does arise, is in contrast to the one we would expect from the model of KS.

[15]Azar (2007) offers another approach to relative thinking. For experimental evidence of relative thinking see, for example, Azar (2011).

thinking model. For example, consider Study 1. As we mentioned earlier, we ran a very similar study which compared choices across the same sets as in Experiment 1.1, where one had a checking account with a tiny interest rate of 0.1% and the other included a checking account with no interest rate. A similar distribution of choices arises when the checking account carries a 0.1% interest rate or 2% and both distributions are different than the one that arises when the checking account has no interest rate. The ToD model suggests that as long as the interest rate is strictly greater than 0, the safe gains dimension is turned-on in the checking account, generating the same dimensional weights across the two experimental versions that are different than those in the corresponding *0-checking* treatments. According to the relative thinking theory of Bushong et al. (2017), we would expect the distributions to look alike when the checking account carries no interest rate and when it generates the tiny interest rate of 0.1% since the variance along the safe gains dimension is very similar across conditions. Our findings contrast this prediction.

In their paper, Bushong et al. (2017) sketch a model which incorporates insights from the focusing model of KS together with their relative thinking approach: Focusing plays a role when choices feature more than two dimensions while relative thinking takes over when there are only two dimensions to consider. In Section 2 we suggested that one could come up with a model that combines our insights alongside those of KS at the stage in which weights are determined. As KS, Bushong et al. (2017) and the ToD procedure seem to complement each other, it would be interesting to consider a model that is general enough to incorporate all of them together. For example, the weight on a specific dimension may combine the number of alternatives in which that dimension is turned-on (as in the ToD procedure) alongside its variance in the choice set. The effect of the variance may depend on the overall number of turned-on dimensions in the choice set a-la the idea raised by Bushong et al. (2017).

Some of our findings may be explained not only through the lens of dimensional weighting. Categories, for example, may be one alternative approach. Models taking this approach describe a decision maker who first forms categories endogenously, and then either chooses the best alternative from the most preferred category (Manzini and Mariotti, 2012) or picks the best option in each category (Furtado et al., 2019).[16] To illustrate, we follow Manzini and Mariotti (2012) and consider the investment example in Study 1. It is plausible that in the *0-checking* treatment, an agent will divide the set into three categories: liquid options, safe options and risky options. Those who care about liquidity may end up choosing the checking account. However, it is also perfectly reasonable that in the *2-checking* treatment the same agent will perceive only two categories: safe options and risky ones. If he is risk averse, he will choose the best option from the first category, which is likely to be the savings plan. Categorization, however, does not seem to apply to the findings from the social preferences study (Study 2) since it does not predict the reversal of ranking between the two unequal splits, which naturally belong to the same category regardless of treatment.

Another channel through which part of our findings may be addressed is choice by iterative

[16] For a different approach involving categories and reference points see Barbos (2010) and Maltz (2019).

search, suggested by Masatlioglu and Nakajima (2013). In their model, the agent starts off with some default option or a reference point in the set. This option generates a consideration set from which the agent picks the best alternative, which replaces his previous reference. The new reference generates another consideration set and the process goes on until the reference point is the best option in the consideration set, at which point it is chosen. The model is a good fit for online search, which often leads to a list of options that need to be skimmed through sequentially. Applying it to our findings, one would naturally treat the first option we introduce as the default. Suppose that when it is the 0% checking account (Study 1), the consideration set includes all perfectly liquid options. In this case, only the checking account is considered and hence it is chosen. However, when the first option is the 2% checking account it consists of all safe options and the agent may end up choosing the savings plan. Once again, as with categories, this approach does not fare well with our findings in Study 2, where preferences over unaltered options, which are likely to be perceived as belonging to the same consideration set, are reversed.[17]

Other models based on reference points, such as loss aversion (Kahneman and Tversky, 1991), may also shed light on our findings but are somewhat harder to apply as they require identifying the reference point from which losses and gains are contemplated. Unlike the iterative search model by Masatlioglu and Nakajima (2013) where the first alternative is a natural and somewhat technical starting point, as in online search, in models based on loss aversion, identifying the reference point is a much more subtle task (Barberis, 2013). Yet, even if we consider the first option as the reference point or the expectation of the participant as he logs in to answer the questionnaire as in Kőszegi and Rabin (2006), our findings are hard to reconcile with the loss aversion approach. Consider once again the investment study in which the checking account is enhanced to include a 2% interest rate and suppose that in the spirit of Kőszegi and Rabin (2006) the reference point's safe gains dimension is taken as the average of the safe interest rates of the checking account and savings plan (2% in the *0-checking* treatment and 3% in the *2-checking* treatment). Under these assumptions, choosing the 0% checking account would generate larger losses compared to choosing the 2% checking account. At the same time, choosing the savings plan would generate larger gains on that dimension in the *0-checking* treatment compared to choosing it in the *2-checking* treatment. As nothing else changes across treatments, no other gain or loss consideration changes either. Thus, the model would predict weakly more choices of the savings plan at the expense of the checking account in the *0-checking* treatment compared to the *2-checking* treatment, in contrast to our findings.

To sum up, the above theoretical models are able to partially explain our findings but none of them is able to predict all patterns. We introduce the ToD procedure that draws on the literature on salience and focusing, while adding the role of 'turned-on' dimensions to relative weighting. The model generates predictions that are in line with the discontinuous nature of our findings in all three studies. The analysis of participants' explanations provides further support for this procedure.

[17] For another approach involving consideration sets formed by an endogenous reference point see Ok et al. (2015).

4.2 Related Experiments

We would now like to relate our findings to experiments reported in the psychology and economics literature. The investment study relates to findings regarding violations of monotonicity. These have been documented in intertemporal choice (Scholten and Read, 2014; Cheng-Ming et al., 2017), in the domain of uncertainty (Gneezy et al., 2006; Bateman et al., 2007) as well as in response to low incentives (Gneezy and Rey-Biel, 2014). These studies argue that an objective improvement (such as a small payment in the future) may actually reduce the attractiveness of an alternative. Our work, on the other hand, focuses on how such improvements may shift dimensional weights and affect the evaluations of other unaltered options as well. For example, we argue that the apparent violation of monotonicity found in Study 1 is not due to the checking account being deemed worse when it generates a positive interest rate but rather because of the increase in the savings plan's evaluation. In fact, it is hard to argue that receiving a 2% annual interest from one's checking account is worse than not receiving any interest (this assessment is supported by Experiment 1.2).

Our studies also share commonalities with experimental work on comparisons along different attributes.[18] The evaluability hypothesis posits that when there are two attributes, one easy to evaluate and the other difficult, the difficult one may have less impact on choice when evaluated separately compared to jointly (Hsee, 1998; Hsee et al., 1999). Slovic and MacPhillamy (1974) show that in binary choices, attributes that are common to both alternatives are weighted more heavily than those that are unique. Building on this early work, Kivetz and Simonson (2000) show that this tendency may lead subjects to choose alternatives that have higher values of the common attributes. In another experiment involving lotteries, Birnbaum (2005) finds that different frames of the same lottery may lead subjects to choose in a manner which violates first order stochastic dominance. Dertwinkel-Kalt and Köster (2015) develop a model in the realm of uncertainty, which is based on the salience model of Bordalo et al. (2013), and incorporates framing effects to account for these violations. The main focus of this development is on how different frames generate different attribute-by-attribute comparisons that may result in anomalies as the ones reported by Birnbaum (2005).

These studies emphasize the role of comparability, whether along common attributes or along attributes that share positive values, on choice. In contrast to this literature, our suggested procedure may place a large weight on attributes that are not common to all alternatives or attributes that equal zero for some alternatives. In fact, if some desirable attribute i is originally turned-off in the entire set but then some alternative z is modified so that $z_i > 0$, then i will receive a positive bump in its weight and become prominent in the decision procedure despite the fact that it is not shared by the other alternatives. Analogously, if all alternatives have some levels (higher than zero) along an undesirable attribute (i.e., it is a common attribute) it will receive no weight in the assessment of goods according to the ToD procedure (as inequality in the *unequal* treatment

[18] Notice that here we use the term attributes as it is most commonly used in this literature where alternatives' characteristics are assumed to be spelled out and observable.

in Study 2). It will receive a positive weight if one alternative carries a value of zero along that attribute (as the $(100, 100)$ allocation in the *equal* treatment in Study 2) even though it is now not a common attribute anymore.

Comparability also allows individuals to provide reasons for their choices which may be at the heart of many context dependent behaviors. Consider, for example, the well known decoy effect (Huber et al., 1982) that refers to the addition of a decoy option to a two-alternative set. When the decoy is dominated by one alternative, but not by the other, preferences have been found to shift in the direction of the dominating alternative.[19] One of the psychological explanations for this phenomenon is that individuals look for reasons to justify their choices (see Lombardi, 2009 and de Clippel and Eliaz, 2012 for two theoretical approaches that hinge on ideas raised in Simonson, 1989, Tversky and Simonson, 1993 and Shafir et al., 1993). Reason-based choice may be one of the underlying forces behind our findings as well. In fact, we may think of turned-on dimensions as providing a basis for reasoning and justifications. However, despite the fact that our work and the decoy effect seem to share an underlying reason-based mechanism, our experiments and suggested procedure are significantly different than those in that literature.

First, our studies do not include a decoy option. In study 1, for example, the checking account is the most liquid option which makes it desirable for quite a few participants. In Study 2 this point of difference is even more pronounced as the all-equal split of (100,100) is actually the highest ranked alternative by almost all participants who prefer option b to option c, let alone a decoy option. Second, in Study 1, the change we introduce to the choice set does not generate a shift of preferences between the two unaltered options (savings plan and stock) as the decoy effect would suggest. Instead, the slight improvement of the checking account shifts preferences away from it and in the direction of the savings plan while the stock's choice share remains the same. Finally, in Study 2, the options that differ between treatments, i.e., (100,100) and (100,130), are either better than the two other options or worse than both of them, for almost all of the participants (according to their own ranking). Thus, the preference shift generated by replacing one of them by the other, is not due to an *asymmetric* dominance relation as in the decoy effect.

Another strand of literature that is related to our work deals with the special effect of the value zero. A number of studies have shown that an attribute with a value of zero may affect choice in a manner which goes way beyond standard cost-benefit analysis. For example, Shampanier et al. (2007) presented students with two chocolates, one of high quality and one of low quality. The price difference between the two chocolates was held constant across treatments (27 cents to 2 cents, 26 cents to 1 cent or 25 cents to 0) but in the treatment in which the low quality chocolate's price hits zero, the proportion of students who chose it peaked dramatically. The authors also provided evidence that the positive affect generated by a free offer is an important psychological factor that

[19] The experimental literature on the decoy effect, also known as the attraction effect, is large and spans a variety of goods, services and even perceptual decision tasks. See, among many others, Simonson (1989); Wedell (1991); Ariely and Wallsten (1995); Dhar and Glazer (1996); Doyle et al. (1999); Scarpi (2008); Hedgcock et al. (2009); Trueblood et al. (2013). See Frederick et al. (2014) for a critical view and Huber et al. (2014) for a response.

drives their results.[20] Palmeira (2010) examines the zero-effect with attributes other than price. He argues that while a free offer, as in Shampanier et al. (2007), generates affect, a value of zero in other attributes does not. For other attributes, he claims, zero "takes the reference away" and hence makes comparisons with other alternatives along that attribute more difficult. In a series of hypothetical experiments he shows that increasing the value of an attribute of one alternative from zero to a small positive value may affect its choice share in a non-monotonic fashion when another alternative outperforms it along that attribute.

Our work differs from these studies in a number of ways. First, our focus is not on the numerical zero value but rather on what turns dimensions on or off in the mind of the decision maker. Second, the Tod model is not confined to one type of dimensions or another. Specifically, it does not require identifying whether some dimension generates affect or not. Predictions may be generated based on whether the dimension is desirable or undesirable, a feature that is normally very easy to identify. Taking price as an example, our model allows to formalize the virtue of affect expressed in Shampanier et al. (2007)—it is the extra weight placed on an undesirable dimension when it carries a value of zero. Finally, in our Study 2, we show that the channel of turned-on dimensions may reverse preferences over two options depending on the characteristics of a third option. Such an 'indirect' effect on choice cannot be accommodated by the psychological procedures suggested in the zero-effect studies.[21]

5 Conclusion

We provide evidence for the effect of turning-on dimensions on individuals' decision processes and choices. In three different contexts, we show that turning-on a dimension shifts participants' dimensional weights when contemplating alternatives and, as a result, choices are affected in a predictable manner. We show that this effect is in some cases strong enough to cause violations of the basic premise of monotonicity in money and may also arise through framing alone. We propose the ToD model that accounts for the discontinuous nature in which turning-on dimensions shifts decision weights in our studies.

As a policy implication we introduce an important, yet unknown, channel through which checking accounts' interest rates may affect investment behavior. Specifically, it suggests that by introducing positive interest rates to checking accounts, banks may increase the subjective weight that investors place on safe gains. As a result, a larger proportion of their assets may be allocated

[20] Recently, Zhang and Slovic (2019) show a similar affect with respect to options that include the possibility of no deaths in the context of life-saving decisions.

[21] Note that even in the binary contexts of the zero-effect findings, the predictions of our model may differ from those in that literature and sometimes even point in opposite directions. For example, according to Palmeira (2010), a bank will be better off by offering student accounts with a low monthly fee rather than zero, if another bank offers student accounts with a large fee. By contrast, The ToD model suggests that the zero offer would turn-on the fee dimension (which is obviously undesirable), increasing the bank's share of student customers as a result.

to safe investments, such as bonds and CDs, at the expense of their checking account balances. Our findings are also relevant to the design of complex contracts and may potentially be taken into account by firms that try to exploit behavioral consumers (e.g., DellaVigna and Malmendier, 2004; Eliaz and Spiegler, 2006; Gabaix and Laibson, 2006) and by the regulator who may try to counteract such exploitation. Consider, for example, a particular health insurance company that does not provide coverage for a relatively common medical condition, which is covered by its competitors. Our findings suggest that by offering even partial coverage for other less probable medical conditions, it would turn them on in the decision makers' minds and consequently decrease the weight assigned to the common medical condition on which it underperforms. This may improve its health plan's evaluation compared to the competing companies' plans at a relatively low-cost. Another platform for exploiting this phenomenon is multi-pricing schemes: Companies that offer services that span a variety of dimensions, such as banks or cellular phone providers, could price many dimensions at zero, understanding that zero payment for a particular dimension of the service will turn it on and mask high prices charged for other dimensions.

These potential applications show the importance of incorporating the role of turned-on dimensions into the decision procedure of different economic agents in the market. A model in which weights are determined by a combination of turned-on dimensions and variance along different dimensions, as in the literature on focusing, salience and relative thinking, may enable us to derive sharper predictions of choice in complex environments.

References

Ariely, D. and T. S. Wallsten (1995): "Seeking Subjective Dominance in Multidimensional Space: An Explanation of the Asymmetric Dominance Effect," *Organizational Behavior and Human Decision Processes*, 63(3), 223–232.

Azar, O. H. (2007): "Relative thinking theory," *The Journal of Socio-Economics*, 36(1), 1–14.

——— (2011): "Do people think about absolute or relative price differences when choosing between substitute goods?" *Journal of Economic Psychology*, 32(3), 450–457.

Barberis, N. C. (2013): "Thirty years of prospect theory in economics: A review and assessment," *Journal of Economic Perspectives*, 27(1), 173–96.

Barbos, A. (2010): "Context effects: A representation of choices from categories," *Journal of Economic Theory*, 145(3), 1224–1243.

Bargh, J. A. and T. L. Chartrand (2000): "The Mind in the Middle," *Handbook of research methods in social and personality psychology*, 253–285.

Bateman, I., S. Dent, E. Peters, P. Slovic, and C. Starmer (2007): "The affect heuristic and the attractiveness of simple gambles," *Journal of Behavioral Decision Making*, 20(4), 365–380.

Birnbaum, M. H. (2005): "A comparison of five models that predict violations of first-order stochastic dominance in risky decision making," *Journal of Risk and Uncertainty*, 31(3), 263–287.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013): "Salience and Consumer Choice," *Journal of Political Economy,*, 121(5), 803–843.

Bushong, B., M. Rabin, and J. Schwartzstein (2017): "A model of relative thinking," *Unpublished manuscript, Harvard University.*

Charness, G. and M. Rabin (2002): "Understanding social preferences with simple tests," *The Quarterly Journal of Economics*, 117(3), 817–869.

Cheng-Ming, J., S. Hong-Mei, Z. Long-Fei, L. Zhao, L. Hong-Zhi, and S. Hong-Yue (2017): "Better is worse, worse is better: Reexamination of violations of dominance in intertemporal choice," *Judgment and Decision Making*, 12(3), 253–259.

Cohn, A. and M. A. Maréchal (2016): "Priming in Economics," *Current Opinion in Psychology*, 12, 17–21.

de Clippel, G. and K. Eliaz (2012): "Reason-based choice: A bargaining rationale for the attraction and compromise effects," *Theoretical Economics*, 7, 125–162.

DellaVigna, S. and U. Malmendier (2004): "Contract design and self-control: Theory and evidence," *The Quarterly Journal of Economics*, 119(2), 353–402.

Dertwinkel-Kalt, M., H. Gerhardt, G. Riener, F. Schwerter, and L. Strang (2017a): "Concentration bias in Intertemporal Choice," *Working Paper*.

Dertwinkel-Kalt, M., K. Köhler, M. R. Lange, and T. Wenzel (2017b): "Demand shifts due to salience effects: Experimental evidence," *Journal of the European Economic Association*, 15(3), 626–653.

Dertwinkel-Kalt, M. and M. Köster (2015): "Violations of first-order stochastic dominance as salience effects," *Journal of Behavioral and Experimental Economics*, 59, 42–46.

Dhar, R. and R. Glazer (1996): "Similarity in Context: Cognitive Representation and Violation of Preference and Perceptual Invariance in Consumer Choice," *Organizational Behavior and Human Decision Processes*, 67(3), 280–293.

Doyle, J. R., D. J. O'Connor, G. M. Reynolds, and P. A. Bottomley (1999): "The robustness of the asymmetrically dominated effect: buying frames, phantom alternatives, and in-store purchases," *Psychology & Marketing*, 16(3), 225–243.

Eliaz, K. and R. Spiegler (2006): "Contracting with diversely naive agents," *The Review of Economic Studies*, 73(3), 689–714.

Fehr, E. and K. M. Schmidt (1999): "A theory of Fairness, Competition, and Cooperation," *The quarterly Journal of Economics*, 114(3), 817–868.

Frederick, S., L. Lee, and E. Baskin (2014): "The Limits of Attraction," *Journal of Marketing Research*, 51(4), 487–507.

Furtado, B. A., L. Nascimento, and G. Riella (2019): "Rational Choice with Full-Comparability Domains," *Working Paper*.

Gabaix, X. and D. Laibson (2006): "Shrouded attributes, consumer myopia, and information suppression in competitive markets," *The Quarterly Journal of Economics*, 121(2), 505–540.

Gneezy, U., J. A. List, and G. Wu (2006): "The uncertainty effect: When a risky prospect is valued less than its worst possible outcome," *The Quarterly Journal of Economics*, 121(4), 1283–1309.

Gneezy, U. and P. Rey-Biel (2014): "On the relative efficiency of performance pay and non-contingent incentives," *Journal of the European Economic Association*, 12(1), 62–72.

Hedgcock, W., A. R. Rao, and H. Chen (2009): "Could Ralph Nader's entrance and exit have helped Al Gore? The impact of decoy dynamics on consumer choice," *Journal of Marketing Research*, 46(3), 330–343.

Hsee, C. K. (1998): "Less is better: When low-value options are valued more highly than high-value options," *Journal of Behavioral Decision Making*, 11(2), 107–121.

Hsee, C. K., G. F. Loewenstein, S. Blount, and M. H. Bazerman (1999): "Preference reversals between joint and separate evaluations of options: a review and theoretical analysis." *Psychological bulletin*, 125(5), 576–590.

Huber, J., J. W. Payne, and C. Puto (1982): "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9(1), 90–98.

——— (2014): "Let's Be Honest About the Attraction Effect," *Journal of Marketing Research*, 51(4), 520–525.

Kahneman, D. and A. Tversky (1991): "Loss Aversion in Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics*, 106(4), 1039–1061.

Kivetz, R. and I. Simonson (2000): "The effects of incomplete information on consumer choice," *Journal of Marketing Research*, 37(4), 427–448.

Kőszegi, B. and M. Rabin (2006): "A Model of Reference-Dependent Preferences," *The Quarterly Journal of Economics*, 121(4), 1133–1165.

Kőszegi, B. and A. Szeidl (2012): "A model of focusing in economic choice," *The Quarterly journal of economics*, 128(1), 53–104.

Lombardi, M. (2009): "Reason-based choice correspondences," *Mathematical Social Sciences*, 57(1), 58–66.

Maltz, A. (2019): "Exogenous Endowment - Endogenous Reference Point," *Economic Journal, Forthcoming*.

Manzini, P. and M. Mariotti (2012): "Categorize Then Choose: Boundedly Rational Choice and Welfare," *Journal of the European Economic Association*, 10(5), 1141–1165.

Masatlioglu, Y. and D. Nakajima (2013): "Choice by iterative search," *Theoretical Economics*, 8(3), 701–728.

Ok, E. A., P. Ortoleva, and G. Riella (2015): "Revealed (P)Reference Theory," *American Economic Review*, 105(1), 299–321.

PALMEIRA, M. M. (2010): "The zero-comparison effect," *Journal of Consumer Research*, 38(1), 16–26.

SCARPI, D. (2008): "The impact of decoys and background information on consumers' preferences and decision making," *The International Review of Retail, Distribution and Consumer Research*, 18(1), 1–15.

SCHOLTEN, M. AND D. READ (2014): "Better is worse, worse is better: Violations of dominance in intertemporal choice." *Decision*, 1(3), 215–222.

SHAFIR, E., I. SIMONSON, AND A. TVERSKY (1993): "Reason Based Choice," *Cognition*, 49(1-2), 11–36.

SHAMPANIER, K., N. MAZAR, AND D. ARIELY (2007): "Zero as a special price: The true value of free products," *Marketing science*, 26(6), 742–757.

SIMONSON, I. (1989): "Choice Based on Reasons: The Case of Attraction and Compromise Effects," *Journal of Consumer Research*, 16(2), 158–174.

SLOVIC, P. AND D. MACPHILLAMY (1974): "Dimensional commensurability and cue utilization in comparative judgment," *Organizational Behavior and Human Performance*, 11(2), 172–194.

TRUEBLOOD, J. S., S. D. BROWN, A. HEATHCOTE, AND J. R. BUSEMEYER (2013): "Not just for consumers: Context effects are fundamental to decision making," *Psychological science*, 24(6), 901–908.

TVERSKY, A. (1972): "Elimination by aspects: A theory of choice," *Psychological review*, 79(4), 281–299.

TVERSKY, A. AND I. SIMONSON (1993): "Context-dependent preferences," *Management science*, 39(10), 1179–1189.

WEDELL, D. H. (1991): "Distinguishing among models of contextually induced preference reversals," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(4), 767–778.

ZHANG, Y. AND P. SLOVIC (2019): "Much ado about nothing: The zero effect in life-saving decisions," *Journal of Behavioral Decision Making*, 32(1), 30–37.

Appendix A: Dimension Elicitation in Study 1

As we explained in the main text, there are difficulties in using participants' ex-post explanations to learn about the dimensions that they deemed as relevant for choice at the time of making the decision. In order to assess whether ex-post explanations may serve as a proxy for the dimensions that were noticed by participants during the first encounter with the choice set, we ran another experiment. In this experiment, which we elaborate upon below, we elicit the dimensions that are relevant for participants without asking them to make a choice. The goals of this experiment are twofold. First, to assess whether the set of dimensions in the ex-post explanations is similar to those elicited directly. Our second goal is to examine whether the frequency with which each of the major dimensions was mentioned is similar across the two methods. We first describe the experiment and then the method used to analyze the results. Finally, we compare the dimensions elicited in this experiment to those that were elicited from ex-post explanations.

The Direct Elicitation Experiment

All participants received the same background story regarding the bonus from the workplace as in Experiment 1.1 and a brief mention of the potential investment options (the full description of the options only came later to prevent participants from thinking about making a choice). Following the background, participants were informed that they are not asked to choose an investment option. Rather, we asked them to write what, in their opinion, are the most prominent dimensions of each option. Then, participants viewed the options (with their full and detailed description) one by one and were provided space to write the prominent dimensions. As in the main experiment of this study, there were two treatments. The first option in the first treatment was a checking account with no interest rate while in the second treatment it was a checking account with a 2% interest rate. The other options in both treatments were a savings plan and a stock with the same characteristics as in the main experiment. We collected data from 223 panelists (different from those who participated in Experiments 1.1 and 1.2) who were randomly assigned to one of the two treatments (111 participants in the *0-checking* treatment and 112 in the *2-checking* treatment).

Method of Analysis

In order to obtain the most objective assessment of dimensions mentioned by our participants we asked for the assistance of two RAs that were not involved in previous text analysis of this project. Both RAs were given the exact same instructions and worked independently.[22] The instructions to the RAs consisted of the text that was observed by the participants in the experiment. In addition they were asked to go through the answers, one by one, and write all dimensions that were mentioned by participants as well as to provide a description for each dimension. We explained what a dimension of an option is and the difference between a dimension and a specific value of

[22] The instructions given to the RAs are available upon request.

Table VII: Dimensions Mentioned by RAs in the Elicitation Experiment.

Dimension	Description	RA-1	RA-2
Risk	Degree of riskiness	✓	✓
Liquidity	Degree of liquidity	✓	✓
Safe Gains	Returns due to interest rate	✓	✓
Potential Gains	Chance for high gains relative to losses (mostly mentioned with risk)	✓	
High Gains	The option involves the possibility for a high gain		✓
Withdrawal Procedure	Requirements for money withdrawal	✓	✓
Trust	Level of confidence with respect to financial institution	✓	✓
Background	Current balances/debt and how it affects decision		✓
Offset Overdraft	Using the bonus to reduce the overdraft in the checking account	✓	
Spending Potential	How likely to wastefully spend money when choosing that option	✓	
Didn't Understand	The participant did not understand	✓	✓
Unclear	The text entry did not make sense	✓	✓

a dimension (e.g., 'blue' is a specific value of the dimension 'car color'). After each RA created her/his list of dimensions, we asked them to go through the answers again and classify each text entry for each option into the categories of dimensions they created (a text entry of a participant referring to one of the options may be classified into more than one dimension). After the RAs completed their work we obtained:

- Two independent lists of dimensions with their descriptions.
- Two independent distributions of percentages of dimensions mentioned for each option in each treatment.

Results

Table VII shows the dimensions listed by the RAs. The first column is the dimension's name (given by the RAs) and the second column consists of a short description. In the last two columns we mark whether that dimension was mentioned by each of the RAs. The first five rows list the dimensions that were frequently mentioned (*frequent dimensions*), followed by those that were mentioned rarely (by less than 7% of participants). The last two rows include those who didn't understand what was asked from them (or didn't understand the option) and those who wrote unclear or meaningless texts (we specifically asked the RAs to include these two categories).

Among the frequent dimensions, risk, liquidity and safe gains were mentioned by both RAs. Potential gains and High gains were mentioned by one but not the other. Looking into these dimensions we find that they overlap although they are clearly distinct. Within the overlapping region one can find texts that refer to high possible gains **and** to the chance to end up losing, which were mostly provided with respect to the stock. Looking at the differences between the two

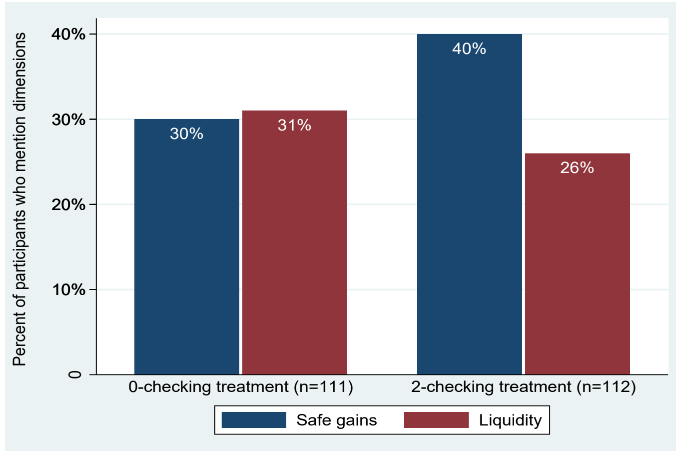


Figure IX: Dimensions mentioned per treatment in the Elicitation Experiment

dimensions, we find that texts that were classified into the potential gains dimension referred not only to the gains but also to the relative gain compared to the potential losses, while the high gains dimension consisted of texts that simply refer to the technical possibility to earn a large sum in that option. Overall, the frequent dimensions are almost identical to those that came up in the ex-post explanations of the main experiment. Among those are safe gains and liquidity, which we focused on in the main text, as well as risk and another dimension that evolves around the large earnings presented by the stock. These are also the dimensions that we use to illustrate the prediction of the ToD model in the next section (excluding risk for simplicity).

We now move on to examine the differences in dimensions mentioned across treatments. Note that this elicitation method may lead participants to mention more dimensions compared to the number of dimensions that come to their minds after choosing. The reason is that we explicitly ask participants to write down the prominent dimensions for *each* option. As a result, participants are made to think through every option and may list dimensions that they wouldn't have noticed when only skimming through the choice set on their way to make a choice. Nevertheless, if the enhanced checking account triggers more thoughts about interest rate and safe gains than the 0% checking account, these should be reflected in our comparisons below.

Figure IX is analogous to Figure V in the main text. We average the number of mentions for each dimension and each option across RAs and show the overall mentions of the safe gains and liquidity dimensions across treatments.[23] Since participants were asked to write relevant dimensions for *each* option (unlike in the main experiment in which only one explanation was provided), the

[23]There were no significant differences across RAs in the percentages of mentions for each dimension of each option (the potential gains of RA-1 was compared with the high gains of RA-2 since these dimensions had a significant overlap as discussed above. These two dimensions were also averaged with each other). The highest difference between the RAs among the frequent dimensions was 7% (in 21 out of 24 instances it was actually less than 4%). Given that these differences were very small, we show their averages in the next tables rather than looking at each RA's distributions at a time.

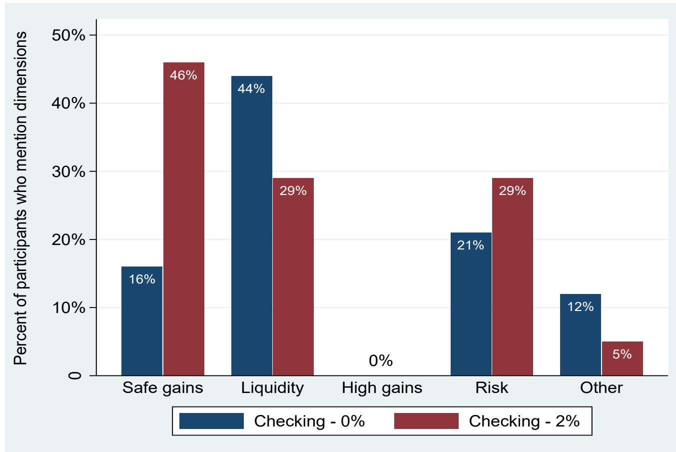


Figure X: Dimensions of checking account per treatment in the Elicitation Experiment

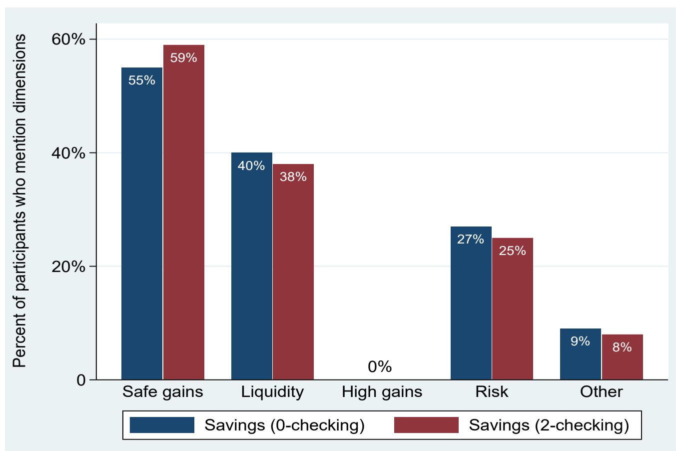


Figure XI: Dimensions of savings plan per treatment in the Elicitation Experiment

percentage was calculated as the overall number of mentions of that dimension in that treatment divided by $3n$, where n is the number of participants in the treatment. The bars in Figure IX are qualitatively similar to those in Figure V and reflect the fact that safe gains is more frequently mentioned in the 0-checking treatment. We view this qualitative similarity as supportive evidence for our usage of ex-post explanations as a proxy for participants' perceived relevant dimensions.

In Figures X, XI and XII we examine the distribution of dimensions mentioned in both treatments for the checking account, savings plan and the stock, respectively.[24] In Figure X, for example, we compare the distribution of dimensions mentioned with respect to the no-interest checking account to the distribution of dimensions mentioned when participants considered the enhanced checking account. This table shows us where the differences in the previous table are coming from: safe gains is mentioned more frequently in the enhanced checking account (46% compared to 16%)

[24]As in IX, we present the average (over RAs) of the percent of mentions of each dimension for each option.

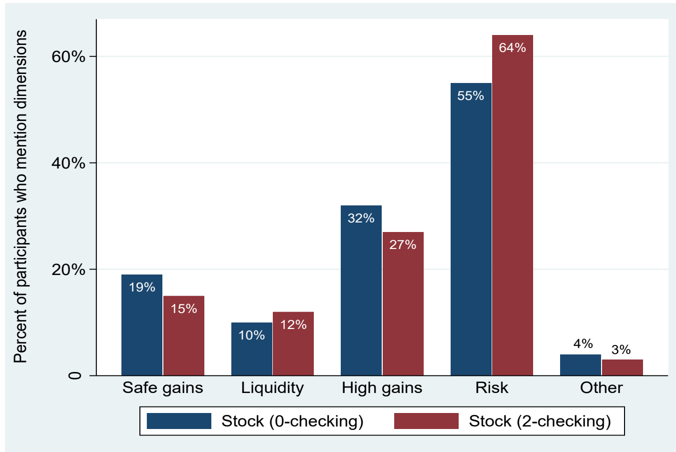


Figure XII: Dimensions of stock per treatment in the Elicitation Experiment

while liquidity is more frequent in the checking account with no interest rate (44% compared to 29%). Other dimensions are mentioned to a similar extent in both options. Finally, Figures XI and XII show that the savings plan and the stock were almost not influenced by the treatment to which they belonged in terms of their perceived prominent dimensions. The largest difference, of 9%, appears in the risk dimension of the stock which was mentioned more frequently in the $2-checking$ treatment. The other dimensions' frequencies per option differed by at most 5% across treatments.[25]

Appendix B: Explaining the Findings with the ToD Model

Our goal in this section is to show that the ToD model predicts changes in the options' evaluations that are in line with the behavioral patterns observed in our experiments, and **that these predictions are independent of the specifics of the consumption utilities**, as long as they are monotonic (and in the case of Study 1 also continuous). For each study, we first derive the general prediction regarding the directional change of preferences and then choose specific parameter values for which this preference change is strong enough to predict the exact choice pattern that we observed.

Appendix B.1: Study 1

There are three available options: *checking*, *savings* and *stock*. ToD weights are simplified by taking g to be the identity function. We consider the following triplet of dimensions, which appeared most frequently in our participants' explanations: safe gains, liquidity and the possibility of high returns

[25]In Figures X - XII all dimensions that were not frequently mentioned were grouped together and named 'other.' These figures also ignore all answers that were unclear or in which the participant did not understand the task.

(higher than 10%).[26] Dimensions are numbered 1, 2, 3 respectively. We assume four 'levels' of these dimensions—(0,L,M,H), where 0 reflects a 0 level of that dimension, L is Low, M is Medium and H is High. The investment options that appear in the study have the following levels in each dimension: *checking*−0%=(0,H,0), *checking*−2%=(L,H,0), *savings*=(H,L,0), *stock*=(0,M,H). In words, both checking accounts have the highest level of liquidity but 0 for the possibility of high returns. The account with a 2% interest rate receives a low level in the safe gains dimension while the one with 0% interest rate naturally receives 0. The savings plan has a high level of safe gains, low level of liquidity and 0 for high returns. The stock has a medium level of liquidity (better than the savings plan but still requiring a visit or a call to withdraw), a high level for the possibility of high returns and 0 for safe gains.[27]

According to these qualitative dimensional values, each investment option has the following vector of turned-on dimensions: $checking - 0\%^{ToD} = (0,1,0)$, $checking - 2\%^{ToD} = (1,1,0)$, $savings^{ToD} = (1,1,0)$, $stock^{ToD} = (0,1,1)$.

Let us now calculate the dimensional weights in each treatment. Denote the choice set in the *0-checking* treatment by $No-Int$ and the choice set in the *2-checking* treatment by $2-Int$. In the *0-checking* treatment, the first dimension, safe gains, is only turned-on in the *savings* since that is the only option which has a value larger than zero in that desirable dimension. The number of overall turned-on dimensions in the choice set is five (the *checking*−0% has only liquidity turned-on, while the *savings* and the *stock* have two turned-on dimensions). Thus the first dimension's weight is:

$$g_1^{ToD}(No - Int) = {}^1\!/_{(1+1+1+1+1)} = {}^1\!/_5.$$

Similarly, we obtain:

$$g_2^{ToD}(No - Int) = {}^3\!/_5,\ g_3^{ToD}(No - Int) = {}^1\!/_5.$$

In the *2-checking* treatment, the weights are different due to the extra turned-on dimension of the checking account:

$$g_1^{ToD}(2 - Int) = {}^2\!/_6,\ g_2^{ToD}(2 - Int) = {}^3\!/_6,\ g_3^{ToD}(2 - Int) = {}^1\!/_6.$$

We now have all the necessary ingredients for the overall evaluation of every alternative in each treatment. The evaluations in the *0-checking* treatment are as follows:

$$\tilde{U}(checking - 0\%, No - Int) = {}^1\!/_5 \cdot u_1(0) + {}^3\!/_5 \cdot u_2(\mathrm{H}) + {}^1\!/_5 \cdot u_3(0).$$

[26] For simplicity, and without loss of generality, we exclude the risk dimension that was also mentioned frequently by our participants.

[27] All options are liquid to some extent as they allow withdrawing the money within, at most, a week. A value of 0 liquidity in our study would fit an option which does not allow withdrawals for a prolonged period of time, say, one year.

Similarly,

$$\tilde{U}(savings, No-Int) = {}^{1}\!/_{5} \cdot u_1(\mathrm{H}) + {}^{3}\!/_{5} \cdot u_2(\mathrm{L}) + {}^{1}\!/_{5} \cdot u_3(0),$$

and

$$\tilde{U}(stock, No-Int) = {}^{1}\!/_{5} \cdot u_1(0) + {}^{3}\!/_{5} \cdot u_2(\mathrm{M}) + {}^{1}\!/_{5} \cdot u_3(\mathrm{H}).$$

Turning to the *2-checking* treatment, we obtain:

$$\tilde{U}(checking-2\%, 2-Int) = {}^{2}\!/_{6} \cdot u_1(\mathrm{L}) + {}^{3}\!/_{6} \cdot u_2(\mathrm{H}) + {}^{1}\!/_{6} \cdot u_3(0),$$

$$\tilde{U}(savings, 2-Int) = {}^{2}\!/_{6} \cdot u_1(\mathrm{H}) + {}^{3}\!/_{6} \cdot u_2(\mathrm{L}) + {}^{1}\!/_{6} \cdot u_3(0),$$

and

$$\tilde{U}(stock, 2-Int) = {}^{2}\!/_{6} \cdot u_1(0) + {}^{3}\!/_{6} \cdot u_2(\mathrm{M}) + {}^{1}\!/_{6} \cdot u_3(\mathrm{H}).$$

We now examine the differences in the evaluations of the checking account and savings plan due to the introduction of the 2% interest rate. For simplicity and without loss of generality, we make one assumption on the consumption utility values, which is: $u_i(0) = 0,\ \forall i$. The increase in the evaluation of the savings plan equals: ${}^{4}\!/_{30} \cdot u_1(\mathrm{H}) - {}^{3}\!/_{30} \cdot u_2(\mathrm{L})$. The first term is the added value due to the increase in the weight of the safe gains dimension, the second term is due to the decrease in the weight of the liquidity dimension. A similar calculation shows that the increase in the evaluation of the checking account amounts to ${}^{2}\!/_{6} \cdot u_1(\mathrm{L}) - {}^{3}\!/_{30} \cdot u_2(\mathrm{H})$. Finally, the evaluation of the stock is increased by: $-{}^{3}\!/_{30} \cdot u_2(\mathrm{M}) - {}^{1}\!/_{30} \cdot u_3(\mathrm{H})$. Thus, if the interest rate is low enough (and u_1 continuous as we assumed) the increase in the evaluation of the savings plan outweighs that of the checking account (and the stock) and pushes in the direction of our observed preference reversal. Reflecting on Study 1 and the participants' frequent mention of safe gains in the enhanced *2-checking* treatment, we argue that this describes the actual weight shift of prominent dimensions for at least some participants.

Moving on to our numerical example, we further assume that the decision maker appreciates high safe gains and does not need the money right now so that a high level of the first dimension is more valuable to him than a high level in one of the other dimensions. Thus, for Dimension 1 : $u_1(0) = 0,\ u_1(\mathrm{L}) = 1, u_1(\mathrm{H}) = 5$. For Dimension 2 we have $u_2(L) = 1,\ u_2(\mathrm{M}) = 2, u_2(\mathrm{H}) = 3$, and for Dimension 3 : $u_3(0) = 0,\ u_3(\mathrm{H}) = 2$.

Given the dimensional weights that we calculated above, the evaluations in the *0-checking* treatment are as follows:

$$\tilde{U}(checking-0\%, No-Int) = {}^{1}\!/_{5} \cdot u_1(0) + {}^{3}\!/_{5} \cdot u_2(\mathrm{H}) + {}^{1}\!/_{5} \cdot u_3(0) = {}^{1}\!/_{5} \cdot 0 + {}^{3}\!/_{5} \cdot 3 + {}^{1}\!/_{4} \cdot 0 = {}^{9}\!/_{5}.$$

Similarly,

$$\tilde{U}(savings, No-Int) = {}^{1}\!/_{5} \cdot 5 + {}^{3}\!/_{5} \cdot 1 + {}^{1}\!/_{5} \cdot 0 = {}^{8}\!/_{5},$$

and

$$\tilde{U}(stock, No-Int) = {}^{1}\!/_{5} \cdot 0 + {}^{3}\!/_{5} \cdot 2 + {}^{1}\!/_{5} \cdot 2 = {}^{8}\!/_{5}.$$

Thus, an agent described by the ToD procedure with the above consumption utilities' values will choose the checking account in the *0-checking* treatment. Turning to the *2-checking* treatment, we obtain:

$$\tilde{U}(checking-2\%, 2-Int) = {}^{11}\!/_{6},\ \tilde{U}(savings, 2-Int) = {}^{13}\!/_{6},\ \tilde{U}(stock, 2-Int) = {}^{8}\!/_{6}$$

and we observe a choice reversal that is an apparent violation of monotonicity. Looking at the numbers, it is evident that the checking account is not made worse due to its additional interest rate. In fact, its overall utility goes up from ${}^{9}\!/_{5}$ to ${}^{11}\!/_{6}$. However, the shift of weights also leads to an increase in the overall utility of the savings plan. These forces pull the relative attractiveness of the two options in opposite directions and according to our utility specification the latter prevails. As shown earlier, the relative change in utilities operates in the direction of our observed behavioral pattern for *any choice of consumption utility values*, as long as the interest rate added to the checking account is small enough and consumption utilities are monotonic and continuous in every dimension.

Appendix B.2: Study 2

ToD weights are simplified by taking g to be the identity function. We naturally consider the undesirable inequality dimension (Dimension 1) alongside the desirable efficiency dimension (Dimension 2), which were the two dimensions that participants referred to most frequently in their explanations.[28] We assume five possible 'levels' (0,VL,L,M,H) of these dimensions where VL reflects a very low level of that dimension, L is Low, M is medium and H is High. Here are the levels along each dimension of the options that appeared in the study: $(100, 100)$=(0,VL), that is 0 in Dimension 1 (inequality) and VL in Dimension 2 (efficiency), $(100, 130)$=(L,L), $(100, 140)$=(M,M), $(100, 160)$=(H,H). In words, the level of both inequality and efficiency is lowest for $(100, 100)$ and increases with the payoff for the other participant. Notice that the level of the desirable dimension of efficiency is above 0 in every alternative (as they all allocate an overall substantial amount to the participants) and hence turned-on in each alternative, while the undesirable dimension of inequality is only turned-on in the $(100, 100)$ split that has a 0 level along that dimension. Thus, for the above qualitative dimensional values, each option has the following vector of turned-on dimensions:

$$(100, 100)^{ToD} = (1, 1),\ (100, 130)^{ToD} = (100, 140)^{ToD} = (100, 160)^{ToD} = (0, 1).$$

[28]In our study, efficiency is equivalent to the sum of payoffs since increasing the other participant's payoff does not entail reducing one's own payoff.

Let us now calculate the dimensional ToD weights. Denote the choice set in the *unequal* treatment by U and in the *equal* treatment by E. In the *equal* treatment, Dimension 1 (inequality) is only turned-on in one option while there are overall four instances of turned-on dimensions in the set (both dimensions are turned-on in (100,100) while only efficiency is turned-on in the other allocations). Hence the dimensional weights are:

$$g_1^{ToD}(\mathrm{E}) = {}^1\!/_{(1+1+1+1)} = {}^1\!/_4,\ g_2^{ToD}(\mathrm{E}) = {}^3\!/_4.$$

In the *unequal* treatment, the weights are different due to the fact that the inequality dimension is completely turned-off. The weights are:

$$g_1^{ToD}(\mathrm{U}) = {}^0\!/_3,\ g_2^{ToD}(\mathrm{U}) = {}^3\!/_3.$$

We now have all the necessary ingredients for the overall evaluation of every alternative in each treatment. The evaluations in the *equal* treatment are as follows:

$$\tilde{U}((100,100),E) = {}^1\!/_4 \cdot u_1(0) + {}^3\!/_4 \cdot u_2(\mathrm{VL})$$

Similarly,

$$\tilde{U}((100,140),E) = {}^1\!/_4 \cdot u_1(\mathrm{M}) + {}^3\!/_4 \cdot u_2(\mathrm{M})$$

and

$$\tilde{U}((100,160),E) = {}^1\!/_4 \cdot u_1(\mathrm{H}) + {}^3\!/_4 \cdot u_2(\mathrm{H}),$$

while given the ToD weights in the *unequal* treatment, we obtain:

$$\tilde{U}((100,130),U) = {}^3\!/_3 \cdot u_2(\mathrm{L}),\ \tilde{U}((100,140),U) = u_2(\mathrm{M}),\ \tilde{U}((100,160),U) = u_2(\mathrm{H}).$$

Thus, moving from the *unequal* treatment to the *equal* treatment, the difference in the evaluation of $(100,160)$ amounts to:

$$\Delta(\tilde{U}) = {}^1\!/_4 \cdot u_1(\mathrm{H}) - {}^1\!/_4 \cdot u_2(\mathrm{H}),$$

while the difference in the evaluation of $(100,140)$ equals:

$$\Delta(\tilde{U}) = {}^1\!/_4 \cdot u_1(M) - {}^1\!/_4 \cdot u_2(\mathrm{M}).$$

Given that Dimension 1 is undesirable and Dimension 2 is desirable, it is evident that the second expression is larger than the first for any choice of monotonic consumption utilities. Thus, the evaluation of (100,140) increases by more than the evaluation (100,160). In other words, the model qualitatively 'pushes' the relative ranking between (100,140) and (100,160) in favor of the former when the (100,130) allocation is replaced with the all-equal (100,100) split. Highlighting the inequality dimension by replacing $(100, 130)$ with the all-equal split alongside the shrouding of the efficiency dimension is the driving force behind this qualitative effect.

We now provide a numerical example that generates an actual reversal between the two unequal allocations. We assume that the decision maker cares about inequality more than he cares about efficiency in terms of their intrinsic influence on his well-being. Thus $u_1(\mathrm{H}) = 0$, $u_1(\mathrm{M}) = 4, u_1(\mathrm{L}) = 8$, $u_1(0) = 12$, and $u_2(VL) = 1$, $u_2(\mathrm{L}) = 2, u_2(\mathrm{M}) = 3$, $u_2(\mathrm{H}) = 4$. Given the ToD weights that we have already calculated above, the evaluations in the *unequal* treatment are as follows:

$$\tilde{U}((100, 130), U) = 2,\ \tilde{U}((100, 140), U) = 3,\ \tilde{U}((100, 160), U) = 4.$$

Hence, an agent in the *unequal* treatment who abides to the ToD procedure and has the above consumption utility values will rank the option $(100, 160)$ first, followed by $(100, 140)$ and $(100, 100)$. Moving on to the *equal* treatment, the evaluations are as follows:

$$\tilde{U}((100, 100), E) = 1/4 \cdot u_1(0) + 3/4 \cdot u_2(\mathrm{VL}) = 1/4 \cdot 12 + 3/4 \cdot 1 = 15/4.$$

Similarly,

$$\tilde{U}((100, 140), E) = 1/4 \cdot 4 + 3/4 \cdot 3 = 13/4,$$

and

$$\tilde{U}((100, 160), E) = 1/4 \cdot 0 + 3/4 \cdot 4 = 12/4.$$

We see that in the equal treatment the rankings are reversed, in line with our findings for a significant percent of participants.

Appendix B.3: Study 3

As in the case of the previous studies, ToD weights are simplified by taking g to be the identity function. We consider three dimensions: The known probability of receiving a prize of 95 ILS (Dimension 1), receiving at least 50 ILS with certainty (Dimension 2) and the possibility to win a prize above 100 ILS (Dimension 3).[29] The study focuses on the first two dimensions: The high prize of 95 ILS is explicitly mentioned in *a'* but not in *a* while certainty is mentioned in the description of option *a* but not in *a'*. We assume three *levels* (0,L,H) of the first dimension and two (0,H) for the other discrete dimensions, where 0 reflects a 0 level of that dimension, L is Low and H is High. The following are the options' levels along the different dimensions: *a*=(L, H, 0), *a'*=(L,H,0), *b*=(H,0,0), and *c*=(0,0,H).

Here is an explanation for the choices of different levels for each option: *a* and *a'* are exactly the same so they have identical levels in all dimensions. Specifically, they have a low probability

[29] For simplicity, we use only these dimensions although others, such as expectations and risk were also referred to by our participants.

(14%) of winning the prize of 95 ILS, a prize larger than 50 ILS with certainty and no chance of obtaining a prize higher than 100 ILS. Option b has a high probability (50%) of winning the prize of 95 ILS, but a certain prize of only 40 ILS and, as a and *a'*, does not offer any prize above 100 ILS. Option c is a bet with unknown probabilities hence it receives a level of 0 in the first dimension. Its minimal prize is smaller than 50 ILS but it does offer a prize that exceeds 100 ILS if the Dow-Jones Index goes up. Keep in mind that this study deals with framing so that an alternative may have a positive level in some dimension which is still not noticed by the decision maker since it is not explicitly mentioned in the description of the alternative. Given the manner in which alternatives are described in the study, each option has the following vector of turned-on dimensions:

$$a^{ToD} = (0,1,0),\ a'^{\ ToD} = (1,0,0),\ b^{ToD} = (1,0,0),\ c^{ToD} = (0,0,1).$$

In other words, Dimension 1 is turned-on when the prize of 95 ILS is **explicitly mentioned** alongside its probabilities, i.e., in options *a'* and b (it is turned-off in a despite its positive value since the decision maker is likely not to think about a prize of 95 ILS given the framing of a). Dimension 2, the prize of at least 50 ILS with certainty, is turned-on only in a since it is the only alternative that is described using the words 'with certainty.' Alternative c is the only one in the set that has Dimension 3 turned-on and that is its only turned-on dimension.

Given these vectors, ToD weights in the *certain(3)* treatment are the following:

$$g_1^{ToD} = {}^{(1)}/_{(1+1+1)} = {}^{1}/_{3}, g_2^{ToD} = {}^{1}/_{3},\ g_3^{ToD} = {}^{1}/_{3}.$$

In the *lottery(3)* treatment the second dimension is turned-off in all alternatives. The dimensional weights are therefore equal to:

$$g_1^{ToD} = {}^{2}/_{3},\ g_2^{ToD} = {}^{0}/_{3},\ g_3^{ToD} = {}^{1}/_{3}.$$

We now have all the necessary ingredients for the overall evaluation of every alternative in each treatment. In *certain(3)*:

$$\tilde{U}(a, \{a, b, c\}) = {}^{1}/_{3} \cdot u_1(\mathrm{L}) + {}^{1}/_{3} \cdot u_2(\mathrm{H}) + {}^{1}/_{3} \cdot u_3(0).$$

Similarly,

$$\tilde{U}(b, \{a, b, c\}) = {}^{1}/_{3} \cdot u_1(\mathrm{H}) + {}^{1}/_{3} \cdot u_2(0) + {}^{1}/_{3} \cdot u_3(0)$$

and

$$\tilde{U}(c, \{a, b, c\}) = {}^{1}/_{3} \cdot u_1(0) + {}^{1}/_{3} \cdot u_2(0) + {}^{1}/_{3} \cdot u_3(\mathrm{H}).$$

Turning to the *lottery(3)* treatment, we obtain:

$$\tilde{U}(a', \{a', b, c\}) = {}^2\!/\!_3 \cdot u_1(\text{L}) + 0 \cdot u_2(\text{H}) + {}^1\!/\!_3 \cdot u_3(0),$$

$$\tilde{U}(b, \{a', b, c\}) = {}^2\!/\!_3 \cdot u_1(\text{H}) + 0 \cdot u_2(0) + {}^1\!/\!_3 \cdot u_3(0),$$

and

$$\tilde{U}(c, \{a', b, c\}) = {}^2\!/\!_3 \cdot u_1(0) + 0 \cdot u_2(0) + {}^1\!/\!_3 \cdot u_3(\text{H}).$$

Moving from *certain(3)* to *lottery(3)*, the difference in the evaluation of b equals ${}^1\!/\!_3 \cdot u_1(\text{H})$, which is strictly positive regardless of the choice of utility values.[30] Thus, the ToD procedure predicts that it will have a higher evaluation due to the change of frame of the first option. The change in the evaluation of the first option, on the other hand, equals: ${}^1\!/\!_3 \cdot u_1(\text{L}) - {}^1\!/\!_3 \cdot u_2(\text{H})$, which a-priori may be positive or negative. However, if the known probability of obtaining the high prize of 95 ILS (Dimension 1) is small enough and given our continuity assumption, the overall evaluation of the first alternative will not increase and the model's prediction is in line with our reported choice reversal.

For the purpose of the numerical example, we assume that the decision maker has the following evaluations along the three dimensions: $u_1(0) = 0,\ u_1(\text{L}) = 7, u_1(\text{H}) = 9,\ \ u_2(0) = 0, u_2(\text{H}) = 3$, and $u_3(0) = 0, u_3(\text{H}) = 5$. Given the dimensional weights we calculated above, we may calculate the evaluations of every alternative in each treatment. In *certain(3)*:

$$\tilde{U}(a, \{a, b, c\}) = {}^1\!/\!_3 \cdot u_1(\text{L}) + {}^1\!/\!_3 \cdot u_2(\text{H}) + {}^1\!/\!_3 \cdot u_3(0) = {}^1\!/\!_3 \cdot 7 + {}^1\!/\!_3 \cdot 3 + {}^1\!/\!_3 \cdot 0 = {}^{10}\!/\!_3.$$

Similarly,

$$\tilde{U}(b, \{a, b, c\}) = {}^1\!/\!_3 \cdot 9 + {}^1\!/\!_3 \cdot 0 + {}^1\!/\!_3 \cdot 0 = {}^9\!/\!_3,$$

and

$$\tilde{U}(c, \{a, b, c\}) = {}^1\!/\!_3 \cdot 0 + {}^1\!/\!_3 \cdot 0 + {}^1\!/\!_3 \cdot 5 = {}^5\!/\!_3.$$

Such an agent would choose a in the *certain(3)* treatment. Turning to the *lottery(3)* treatment, we obtain:

$$\tilde{U}(a', \{a', b, c\}) = {}^2\!/\!_3 \cdot 7 = {}^{14}\!/\!_3,\ \tilde{U}(b, \{a', b, c\}) = {}^{18}\!/\!_3,\ \tilde{U}(c, \{a', b, c\}) = {}^5\!/\!_3.$$

[30] We assume once again that $u_i(0) = 0,\ \forall i$. In this exercise this assumption does entail some loss of generality. Without it, we would need to require that for a small enough known probability of obtaining the high prize of 95 ILS (Dimension 1), the term $[{}^1\!/\!_3 \cdot u_1(\text{H}) - {}^1\!/\!_3 \cdot u_2(0)]$ is greater than the term $[{}^1\!/\!_3 \cdot u_1(\text{L}) - {}^1\!/\!_3 \cdot u_2(\text{H})]$.

Thus, the change of frame shifts an individual described by the ToD model with the above utility values from choosing a in the *certain(3)* treatment to choosing b in treatment *lottery(3).* While the first option does not change per se, the lottery framing with its explicit mention of the prize of 95 ILS turns-on the first dimension that was turned-off in the certain payment framing. At the same time, the certain payoff is no longer mentioned in *lottery(3)* and as a result the dimension on which the first option performs well—Dimension 2—receives no weight. Overall, a higher weight is given to the first dimension and a lower weight to the second dimension. Given our choices of utility values, option b benefits the most from this shift in weights since it performs best along the dimension with the bumped up weight. The first option gains from the increased weight of the first dimension but is hurt from the reduced weight of the second dimension. Overall, its evaluation increases but to a lesser extent than the evaluation of b which is now the highest in the set.

To complete the picture we show how the model, with these specific utility values, explains the findings from treatment *certain(2)* and *lottery(2).* In the former, weights are given by:

$$g_1^{ToD} = 0,\ g_2^{ToD} = g_3^{ToD} = {}^{(1)}/_{(1+1)} = {}^1/_2,$$

while in treatment *lottery(2)*:

$$g_2^{ToD} = 0,\ g_1^{ToD} = g_3^{ToD} = {}^1/_2.$$

With these weights, we obtain the following evaluations. In *certain(2)*:

$$\tilde{U}(a, \{a, c\}) = {}^1/_2 \cdot u_2(\mathrm{H}) + {}^1/_2 \cdot 0 = {}^3/_2$$

and

$$\tilde{U}(c, \{a, c\}) = {}^5/_2.$$

On the other hand, in treatment *lottery(2)* we obtain:

$$\tilde{U}(a', \{a', c\}) = {}^7/_2,\ \tilde{U}(c, \{a', c\}) = {}^5/_2.$$

In the absence of b, Dimension 1 receives 0 weight in treatment *certain(2)* while Dimension 2 receives 0 weight in treatment *lottery(2).* According to our numerical example, When a is replaced by *a'* and turns on the first dimension this leads to a relatively large shift in the evaluation of the first option. At the same time, Dimension 3 receives the same weight across treatments and hence the evaluation of option c is unchanged. This leads to the pattern we observe across these binary choice treatments—a higher proportion of participants choosing the first option in the *lottery(2)* treatment.

Appendix C: Questionnaires

Below are the English translations for the instructions of the main experiments of all studies (the instructions were originally written in Hebrew as the experiment was run in Israel). The wording of the parallel treatment is reported in square brackets.

Appendix C.1. Study 1: Instructions of the *2-checking* [*0-checking*] treatment

Decision Making Questionnaire - General Instructions

1. Thank you for agreeing to participate in a brief decision making experiment. The experiment includes just a few questions and is expected to take a few minutes to complete.
2. The questions are phrased in masculine form but are addressed to women and men alike.
3. The questionnaire deals with your preferences and therefore there are no right or wrong answers.
4. The questions describe hypothetical situations in which you are asked to choose between several options. For the success of the experiment we ask that you answer the questions sincerely.[31]
5. The experiment is completely anonymous.

Question 1

Imagine that you are an employee in a firm. At the beginning of the new year your employer informs you that you, as well as the other employees, are about to receive a bonus of 10,000 ILS. This bonus will be deposited for you by your employer in one of three options. Which one would you choose?

a. In your checking account which generates a 2% yearly interest rate with certainty. [which does not generate any interest.]
   * Some checking accounts in Israel have interest and some do not. Please assume for this questionnaire that your account has a 2% interest [no interest] even if this is not the case in reality.

b. In a savings plan which generates a 4% yearly interest rate with certainty.
   * The account has weekly exit options, in which you can withdraw the money by making a request online or by phone.

[31] Participants received a flat rate of 5 ILS for completing the questionnaire but that was not iterated in the instructions as it was communicated through their user account in the panel company.

c. In stocks that can gain or lose with a 50-50 chance. If it goes up, it earns 14% a year, if it goes down it loses 5% a year.
   * The stocks can be sold any time by making a request online or by phone.

Note: If the amount (or part of it) is withdrawn before an entire year has passed, you will receive the proportional share of the annual profits. At the end of each year, the remaining balance on your chosen track will remain on the same track under the same conditions unless you specify otherwise.

Question 2

Please briefly explain your choice: ☐

Question 3

Now imagine that the situation is the same as described in Question 1, only that now the employer asks you to choose the percentage of the amount of 10,000 ILS that you would like to deposit in each option. Note that the sum of the percentages must equal 100. What is the percentage you would like to allocate to each option?

a. In your checking account which generates a 2% yearly interest rate with certainty. [which does not generate any interest.] ☐

b. In a savings plan which generates a 4% yearly interest rate with certainty. ☐

c. In stocks that can gain or lose with a 50-50 chance. If it goes up, it earns 14% a year, if it goes down it loses 5% a year. ☐

Please briefly explain your choice: ☐

Appendix C.2. Study 2: Instructions of the equal [unequal] treatment

Decision Making Questionnaire - General Instructions

1. Thank you for agreeing to participate in a brief decision-making experiment. The experiment includes two questions and is expected to take a few minutes to complete.

2. The questions are phrased in masculine form but are addressed to women and men alike.

3. The questionnaire deals with your preferences and therefore there are no right or wrong answers.

4. **In this questionnaire there is a possibility of winning a significant amount of money. At the end of the experiment (in about two days) 5% of those who complete the entire questionnaire will be randomly drawn to receive prizes according to their choices. Please note that this payment is on top of the participation fee which you will receive for filling out the questionnaire.[32] At the moment it is impossible to know which of the participants will be drawn for payment and therefore it is recommended to answer according to your true preferences. Those who will be drawn to receive the additional payment will be notified of their prize via email.**

5. The experiment is completely anonymous.

[32]Participants received a flat rate of 3 ILS for completing the questionnaire but the exact compensation was not iterated in the instructions as it was communicated through their user account in the panel company.

Question 1

Assume that you have been selected for payment. Chosen alongside you is another participant that you do now know (which will also complete the questionnaire). You are asked to determine the payment for both of you. There are three options:

a. 100 ILS for you and 100 ILS for the other participant. [100 ILS for you and 130 ILS for the other participant.]

b. 100 ILS for you and 140 ILS for the other participant.

c. 100 ILS for you and 160 NIS for the other participant.

Please rank the options according to your preferences: **1 - the option you prefer the most, 2 - the option that is ranked 2nd according to your preferences, 3 - the option that you prefer the least.**

You and the other participant will not know anything about each other's identity.

Note: For payment purposes, the option you rank highest will be selected with a 60% chance and the option you rank second will be chosen with a 40% chance. Therefore, it is recommended that you rank all three options according to your true preferences.

a. 100 ILS for you and 100 ILS for the other participant. [100 ILS for you and 130 ILS for the other participant.] ☐

b. 100 ILS for you and 140 ILS for the other participant. ☐

c. 100 ILS for you and 160 NIS for the other participant. ☐

Question 2

Please briefly explain your choice: ☐

Appendix C.3. Study 3: Instructions of the *certain(3)* [*lottery(3)*] treatment

Below are the instructions for treatments *certain(3)* and *lottery(3)*. The instructions for treatment *certain(2)* and *lottery(2)* are identical except for the fact that option (b) is excluded.

<u>Decision Making Questionnaire - General Instructions</u>

1. Thank you for agreeing to participate in a short experiment that includes two questions and is expected to take a few minutes.

2. The questions are phrased in masculine form but are addressed to women and men alike.

3. The experiment is anonymous. You are only requested to specify your gender, your major, and age range. In addition, we ask you to type your email address which will be used only to update you if you won a prize.

4. The questionnaire deals with your preferences and therefore there are no right or wrong answers.

5. If you have any questions or comments, please send an email to Ayala Arad from Tel Aviv University (aradayal@post.tau.ac.il).

6. As you will shortly see, the experiment describes a choice between several options that entitle you to significant amounts of money. As soon as the experiment ends (it will end in a couple of days), 5% of those who fill out the entire questionnaire will be randomly drawn **to receive the money amount according to their choice.** We will send an email to the winners and explain where they can receive their payment. Payment can also be received through Bit and Pepper Pay payment applications.

7. At the moment it is impossible to know which of the participants will be drawn for payment and therefore it is recommended to address the question as if you will really receive your chosen option.

Email (to be used only to notify you if you won a prize): ________

Gender:

- Male
- Female

Age:

- 18-25
- 26-35
- 36-45
- 46+

Major:

Question 1

You are facing the following three options. Which one would you like to choose?

a. Receive 60 ILS with certainty. On top of this amount, you will receive an additional 35 ILS if you win in a lottery that will be performed by the computer (a 14% chance). [Participate in the following computer lottery: A 14% chance to receive 95 ILS and an 86% chance to receive 60 ILS.]

b. Participate in the following computer lottery: A 50% chance to receive 95 ILS and a 50% chance to receive 40 ILS.

c. Participate in the following gamble on the stock market: If the Dow Jones Industrial Average Index at the end of the next trading day is higher than at the beginning of that day you will receive 115 ILS. If it drops, you will receive 30 ILS (the probability that the index will increase / decrease is not known).
Note: The Dow Jones Industrial Average Index is a stock market index that shows how 30 large publicly owned companies based in the United States have recently traded.

Question 2

Please briefly explain your choice: