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

Ayala Arad † Amnon Maltz ‡

September 25, 2018

Abstract

Could introducing a tiny interest rate on positive balances of checking accounts affect investment decisions? We suggest, counterintuitively, that it might decrease allocations to checking accounts while increasing riskless investments with higher returns. This violation of monotonicity is a potential outcome of a novel behavioral phenomenon that we formalize and investigate experimentally in different environments. It posits that even a 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 may shift from the most liquid option, the checking account, to safe investments with superior returns.

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

*We would like to thank Neta Klein, Eli Mograbi, Dor Shlenger, Yuval Tsuk and Gilad Bar-Levav for their great assistance. We also thank Ariel Rubinstein, Yair Antler, Dotan Persitz and seminar participants at the Hebrew University, Ben-Gurion University, University of Haifa, BRIC5 conference, the 2nd Behavior Change Conference in Coller School of Management, Annual meeting of the Israeli Association of Economics, and CESS 15th Anniversary meeting at NYU for helpful comments.

†Coller School of Management, Tel Aviv University, aradayal@post.tau.ac.il
‡Department of Economics, 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 the same 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-monotnoic 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*, i.e., explicit and obvious to the decision maker, 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.

Our contribution to the literature is twofold: First, we design three experiments that provide evidence for the effect of turning-on dimensions on choice in different environments. Second, we propose a choice model which takes the role of turned-on dimensions into account. This model hinges on ideas raised in the literature on salience and focusing (Bordalo et al., 2013; Kőszegi and Szeidl, 2012) and, as in that literature, it assumes that subjective decision weights depend on the context. However, our procedure places a spotlight on turned-on dimensions as the underlying feature that determines decision weights while in the above models the variance of the dimensions' values is the underlying feature that affects weights. As we elaborate below, the existing models are unable to accommodate some of our experimental findings. After we present our formal model, we discuss how the two approaches may be combined to derive predictions in different circumstances.

1.1 Psychological Channel in Brief

In order to describe our proposed procedure, we first need to explain what it means for a dimension to be turned-on in an alternative. Generally speaking, a dimension is turned-on in some alternative if it is explicitly present in that alternative's description and therefore stands out (a formal definition of turned-on and turned-off dimensions and a complete description of the psychological procedure appear in Section 3). We explore two channels for turning-on dimensions: Changing a dimension's value and changing the description of an alternative without altering dimensional values.

Turning-on and off by changing dimensional value

We suggest that the value of dimension i in an alternative c determines whether the dimension is turned-on or off. Specifically, it depends on whether this value is greater than or equal to zero. To describe this channel we make a distinction between what we call *desirable dimensions*, along which more is better, and *undesirable dimensions* where more is worse. For example, the annual interest rate is a desirable dimension of checking and savings accounts while the temporal distance of obtaining a prize is an undesirable dimension of payment schemes. A desirable dimension is turned-on in an alternative if its value is greater than zero and turned-off if it equals zero. Going back to our investment example, when the checking account generated no interest, its safe gains dimension was turned-off and so it was unlikely to be thought of as an investment tool. By contrast, if it generates positive interest, thoughts about annual returns naturally arise. In this case, the checking account has the safe gains dimension turned-on. As another example, one can think of the warranty on a new product. If there is no warranty, you may not think about this aspect at all, while even a very short period of warranty is likely to make you take this aspect into account.

For undesirable dimensions we employ a definition that mirrors the one for desirable dimensions. We say that an undesirable dimension is turned-on in an alternative if its value equals 0 and turned-off otherwise. For example, imagine that you are searching for an apartment. If one has a pool in the same building it emphasizes the distance 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. As another example consider temporal choice: If two prizes are to be paid to you in the future, say one in a week and the other in 10 days, the temporal distance until the prize is paid will probably be less pronounced than the prizes themselves. However, if the first prize is to be paid out today, the temporal distance becomes a much more salient consideration. In other words, an undesirable dimension is turned-on in an alternative when its absence highlights the attractive facet of that dimension (e.g., right here compared to x meters away, immediately vs. y days from now, or completely equal vs. some inequality).

Thus, both desirable and undesirable dimensions are turned-on in an alternative when their attractive facet is explicitly present, i.e., when its level is greater than zero for desirable dimensions and equals zero for undesirable ones.

Turning on and off by the description of alternatives

Framing can also make dimensions more or less explicit and hence determine whether they are turned-off or turned-on. For example, a $50-50$ lottery that pays \$50 or \$140 may also be described as \$50 with certainty and a 50% chance to win an additional \$90. In the former more standard lottery description, the high prize of 140 is clearly stated and therefore draws attention. In the latter frame, on the other hand, the certainty of winning "50" stands out and is likely to draw more attention than the high prize. Similarly, when shopping for a product, the attributes that are listed next to each product are the ones that draw the largest amount of attention when the product is being examined.

Decision Procedure

Our agent starts off by recognizing which dimensions are turned-on in each alternative. Then he determines weights for each dimension. Specifically, the weight of dimension i is proportional to the number of alternatives in the choice set in which i is turned-on divided by the overall number of instances of turned-on dimensions in the set. In other words, the more frequently a given dimension is turned-on in the choice set, the more salient it is and the higher weight it carries. These weights are used to evaluate the final utility value of all available alternatives. We show that this procedure predicts that even small changes to some alternative's dimension 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).

Our suggested decision making process is dubbed the ToD (Turned-on Dimensions) procedure and it is formalized in Section 3. This procedure strengthens the knot between salience and context effects. The notion of salience has recently been introduced into economic models of decision making. Bordalo et al. (2012) discuss salience under risk and later expand to riskless consumer environments (Bordalo et al., 2013) where context effects are also explored. Kőszegi and Szeidl (2012) develop a model of consumer choice that is formally closest to the procedure we suggest in this work (in fact, we build on their approach when we lay out the model). The main difference between our approach and the above models lies in the feature that underlies salience. In Kőszegi and Szeidl (2012), roughly speaking, a dimension's salience depends on its variance in the choice set. In Bordalo et al. (2012) each alternative may have its own salient dimension depending on the distance of that dimension's value from its mean in the set. In the model we suggest, a dimension's salience is determined by the share of options in the set in which it is turned-on compared to other turned-on dimensions. We view our turning-on channel as complementary to this literature and discuss how it may be integrated with the existing models in Section 5.

1.2 Experimental Studies in Brief

In the design common to all of our studies, a decision maker has to choose among a set of alternatives that have values along different dimensions. In laboratory studies and behavioral models alternatives are often treated as if fully described by checklists of attributes. However, in many real life scenarios alternatives consist not only of such a checklist but also of additional features perceived by the decision maker. For example, a laptop has technical objective attributes, such as storage capacity or battery power, but also subjective features, such as fineness of design or the feel of the keyboard. This is also true for the decision environment that we set up in our studies. Therefore, we use the term dimensions (and sometimes criteria or features) rather than attributes when referring to aspects of the alternatives.

Our experiments span three different choice contexts: social preferences, investments and choice under uncertainty. We modify one alternative in the choice set in a manner that turns-on one of its dimensions by slightly altering its dimensional values or by changing the way it is framed. To gain deeper insight into the decision making considerations, we not only examine final choices, but also detect evidence of dimensional prominence by analyzing participants' explanations.

In our first study, 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 only in the first allocation. In the *unequal* treatment, participants face the following allocations:

$$a = (100, 130),\ \ b = (100, 140),\ \ c = (100, 160),$$

where a pair (x, y) stands for x Israeli Shekels (ILS) for the participant and y ILS for another anonymous participant. In the *equal* treatment, allocation a becomes an equal $(100, 100)$ split while allocations b and c remain unchanged. In this context, we think of inequality as an undesirable dimension and the option $(100, 100)$, which has a level of zero in inequality, turns this dimension on.

The change we introduce to the choice set should carry no consequences on the relative ranking of b and c if participants hold stable preferences over alternatives. However, we find that a significantly higher proportion of participants rank b over c in the *equal* treatment compared to the *unequal* treatment. To gain deeper insight into the forces behind this choice pattern, we first identify the most common dimensions mentioned in our participants' explanations, which are "inequality" and "efficiency". We examine the prevalence with which these dimensions are mentioned in the explanations and find that inequality is 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 $(100, 100)$ allocation turns-on the undesirable inequality criteria and shifts preferences in the direction of more equal allocations. Notice that replacing $(100, 130)$ with $(100, 100)$ increases the variance of both inequality and efficiency in the choice set. Thus, predictions of the models by Kőszegi and Szeidl (2012) and Bordalo et al. (2013), in which salience is determined by variance, depend on the exact shapes of the utility functions and specifically on the relative diminishing sensitivity along these criteria. While some functions may generate predictions in line with our findings, others will generate the opposite predictions. By contrast, as we show in Section 4, the ToD procedure has a unique prediction in line with our findings.

In our second study we turn-on a desirable dimension by shifting the value of one alternative's dimension from 0 to a positive level. The study follows our earlier investment example. 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 investments are held fixed across treatments: The checking account is entirely liquid and accessible anytime as in real life. The savings plan has weekly withdrawal options whereas the stock allows withdrawal anytime, but both require a phone call or a visit to the bank in order to do so.

Following participants' explanations, the most common dimensions in this decision problem are: liquidity, safe gains and the possibility of high returns. In both treatments the checking account is chosen by a non-negligible group of participants because, according to their explanations, it is entirely liquid. 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 is derived from the focusing and salience models (Kőszegi and Szeidl, 2012; Bordalo et al., 2013): The range of the safe gains dimension in the choice set shrinks due to the increase of the checking account's interest from 0% to 2%, making this dimension less salient and hurting the attractiveness of the savings plan in the process. Alongside the objective enhancement of the checking account both models predict a (weakly) smaller proportion of participants choosing the savings plan and a (weakly) larger proportion choosing the checking account.

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. Analyzing participants explanations, we find support for our hypothesized ToD procedure: when the checking carries no interest, liquidity receives substantial weight in the decision process, leading about a quarter of our participants to choose the checking account. By contrast, when the checking account generates positive interest, its dimension of safe gains is turned-on and therefore this dimension receives a larger weight at the expense of liquidity, which is now shrouded. This change in weights leads to a higher evaluation of the savings plan, as it entails the highest safe gain in the set. As a result, more participants choose the savings plan while only about 10% opt for the checking account. This study sheds light on 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 safe investments, such as bonds and CDs, and lead, counterintuitively, to a reduction in checking account balances.

Last but not least, in our third study we show that weights can be shifted without actually changing the choice set, i.e., by framing alone. In particular, we show that explicitly mentioning a dimension of a lottery (without actually changing its value) turns it on and increases its relative weight in the decision process. In the first treatment, participants are asked to choose one of the following three alternatives:

- Lottery A: 60 ILS with certainty + an additional 35 ILS with a 14% chance.
- Lottery B: 50% chance of winning 40 ILS and 50% chance of winning 95 ILS.
- Bet C: If the Dow- Jones Index drops tomorrow, you win 30 ILS; otherwise, you win 115 ILS.

In the second treatment, participants face the exact same alternatives except for the framing of the first lottery:

- Lottery A': 86% chance of winning 60 ILS and 14% chance of winning 95 ILS.

Lottery A includes the phrase "with certainty" thereby drawing attention to the sure gain of 60 ILS. Lottery A' on the other hand, does not mention certainty. Instead, it explicitly mentions the highest prize one can win in this lottery, which is 95 ILS. The certainty dimension is therefore expected to be turned-on in Lottery A but turned-off in A' and the opposite for the dimension reflecting the probability of winning 95 ILS. Following our ToD procedure, in the second treatment a larger weight will be placed on the probability of winning the high prize, a dimension in which Lottery B outperforms Lottery A'. At the same time, certainty is shrouded in this treatment, meaning that the relative advantage of Lottery A' over B receives lower weight. As a result, Lottery B will be chosen more often in the second treatment compared to the first.

Indeed, we find that a large share of participants in the first treatment choose Lottery A. By contrast, in the second treatment, framed as Lottery A', its share is significantly lower while the share of Lottery B increases by the same magnitude. The explanations provided by our participants support the shift in decision weights in the direction of the turned-on dimensions. Importantly, the share of participants who choose Bet C (which has non-specified probabilities) is almost unchanged across treatments. This suggests that the main effect is indeed due to weight shifting between the two different dimensions of the first lottery that are alternately made explicit across treatments. In two additional treatments we rule out an alternative explanation for this choice pattern according to which Lottery A is simply better than A' in the eyes of our participants, regardless of the context.

These findings 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). In the priming literature, different types of cues are manipulated in order to measure their influence on participants' behavior.[1] A large part of the literature in priming focuses on prompting participants to think about a specific concept or recollect past experiences prior to some task. Our third study provides evidence for the activation of dimensional prominence in a more subtle way: Making a dimension of some alternative explicit, i.e., turning it on, primes individuals to give more weight to that dimension when settling their decision problem.

The paper proceeds as follows: In Section 2, we describe the experimental studies in detail followed by the results. Section 3 outlines the ToD model while Section 4 illustrates how it accommodates the experimental findings. In Section 5, we discuss other modeling approaches and related experimental evidence. Section 6 concludes.

2 Experimental Studies

Study 1: Social Preferences in the Presence of an Equal Split

This study aims at illustrating how an undesirable dimension is turned-on when its value equals 0. We deal with social preferences and explore how replacing an unequal allocation with an all-equal split of a pie turns the undesirable inequality dimension on. Participants in this study were 393 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, which included only two simple questions (actually 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 $0.9). In addition, it was explained in the instructions that 5% of the participants would be randomly selected to receive additional payoffs according to their responses. Participants were randomly assigned to one of two treatments, named *unequal* ($n = 194$) and *equal* ($n = 199$). Both treatments described a situation in which the participant was chosen, alongside another anonymous participant, to receive payment and was asked to determine the exact payment each of them will receive. It was explicitly mentioned that the identity of the other participant would not be disclosed. The complete

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

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

Table I: Monetary payments by treatment in Study 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 study 100 ILS were roughly equal to $30).

questionnaire appears in Appendix B.1.

Table I shows the different options that were available in each treatment.[2] Options b and c are unequal splits that are identical in both treatments, and option a is different: an equal split in one treatment and an unequal split in the other. 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. Upon completion of the study, 20 participants were randomly drawn and received payments accordingly. Finally, participants were asked to provide a brief explanation for their ranking.

Study 1: Results

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 efficiency considerations. Notice that one does not sacrifice his own payoff by increasing the other (anonymous) person's payoff.[3] In line with previous findings in the social preference literature (see Charness and Rabin, 2002), 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 option a. In the *unequal* treatment only 18% rank b over c. In the *equal* treatment this percentage rises to 32% (this difference of 14% is significant according to Pearson's chi-squared test, p=0.002). In a probit regression reported in Table II we control for the order of the alternatives and find a significant effect of the *unequal* treatment on ranking b over c and no effect for the order of presentation. The treatment effect amounts to a decrease of 14% in the likelihood of ranking b over c when the $(100, 100)$ allocation is

[2] To control for order effects, each treatment had two opposing orders of the three options. To avoid confusion, we kept an increasing or decreasing order (in the other participant's payoff).

[3] 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' explanation analysis, we include all psychological forces supporting a larger payment to the other participant without hurting one's own payment.

Variable	**Marginal Effect**
Unequal treatment	-0.14***
	(0.002)
reverse order	0.01
	(0.831)
cons	-0.928***
	(0.000)
N	393
R^2	0.024

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

Table II: Marginal effects on the probability of ranking b over c.

replaced with $(100, 130)$.

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 $(100, 100)$ on top in the equal treatment (38%) compared to those who rank $(100, 130)$ on top in the unequal treatment (18%). The difference in proportions is reversed looking at those who rank $(100, 160)$ on top: 82% in the unequal treatment compared to only 60% in the equal one (both differences are highly significant according to Pearson's chi-squared test ($p < 0.001$)).[4]

Next, we wish to gain insight into the underlying psychological procedure leading to these marked differences. In order to do so, we analyze the explanations that were provided by the participants for the ranking they chose. For this purpose, we prepared a list of categories of relevant criteria after reading the explanations ourselves. These categories were exhaustive and reflected the various dimensions that were mentioned by our participants. Then, three research assistants 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

[4]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 a on top actually ranked $a \succ b \succ c$ (87 out of 104). Out of the 278 participants who ranked c on top, 276 ranked $c \succ b \succ a$.

Options	***Equal***	***Unequal***
a (100,100)/(100,130)	38% (76)	14% (28)
b (100,140)	2% (4)	4% (7)
c (100,160)	60% (119)	82% (159)

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

for the vast majority of cases.[5] We concentrate on the two categories that were referred to the most: "inequality" and "efficiency". If, as we expect, the inequality criterion is weighted more heavily in the *equal* treatment, it should be mentioned more often in the explanations compared to the *unequal* treatment. Similarly, we expect the efficiency criterion to be more prominent in the *unequal* treatment compared to the *equal* treatment because it is not shrouded by the inequality criterion. Figure I 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*).

Taking stock, in this study we find that moving the value along 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. Our findings cannot be explained by any type of stable preferences, i.e., preferences that are context independent. Moreover, since the replacement of $(100, 130)$ with $(100, 100)$ increases the variance of efficiency and inequality in the set, predictions of the models of salience and focusing (Bordalo et al., 2013; Kőszegi and Szeidl, 2012) depend on the shape of the utility functions, specifically on the marginal utilities along different dimensions. In general, they may predict a shift in line with our findings but may also predict the opposite shift. By contrast, as we show in Section 4, the model based on the ToD procedure, which we formulate in the next section, predicts a preference shift that is in line with our findings regardless of marginal utilities along dimensions (as long as monotonicity along every dimension is assumed).

Study 2: Enhancing the Checking Account in Investment Decisions

Our first study dealt with turning-on undesirable dimensions. Next, we turn-on a desirable dimension by increasing its value from 0 to some level greater than 0. Our set up is

[5]This procedure was held for each of the three studies. In this study, their classifications were aligned along 91% of possible entries. In the second and third studies, unanimous agreement was reached along 84% and 85% of the entries, respectively.

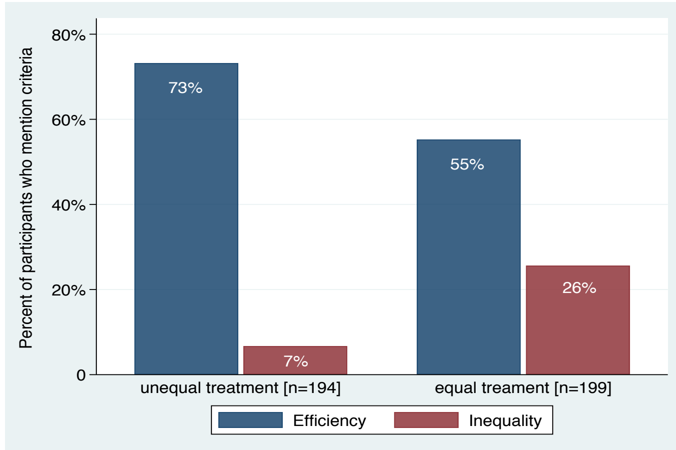


Figure I: Criteria mentioned per treatment in Study 1.

a hypothetical scenario which we believe is not uncommon in real life and therefore carries important policy implications. Specifically, we examine the effect of adding a positive interest rate to the checking account on individuals' investment decisions. Participants were 201 registered panelists who received 5 ILS for completing the questionnaire (the demographic details are similar to Study 1). 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. 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 (and earn 14%) or down (and lose 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 withdrawal options while the stock could be sold anytime. It was also stated that early withdrawal from the savings plan or the stock required a phone call or a visit to the bank and that they may withdraw part of the money with the relative expected gains (the full questionnaire is available in Appendix B.2). 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 the same study (with minor wording changes) with an enhanced checking account that had only a "tiny" yearly interest of 0.1%. That is, in that study one treatment had a 0% checking account, a savings plan and a stock (the exact same options as in the *0-checking* treatment reported above) whereas the other treatment had a 0.1% checking account alongside the same savings plan and the same stock. The results are very similar to those reported below and are therefore omitted.

Study 2: Results

First, note that despite the fact that the checking account is dominated by the savings plan along the interest rate dimension in both treatments, it has other merits and is therefore not a completely dominated option. It is the most liquid of all alternatives and has the most convenient withdrawal requirements (simply using the ATM). A significant amount of participants choose this option (in both treatments) 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 it later as they see fit (because they can access it at any moment in time).

Standard consumer theory would predict a weakly higher share of participants choosing the checking account when it earns positive interest compared to the share that choose it when it earns no interest rate due to monotonicity of preferences in money. However, counterintuitively, the checking account is actually chosen less often when presented with an interest rate. As shown in Figure II, 23% of the participants choose the checking account with zero interest while only 11% do so when it generates a 2% interest (p=0.016, chi-squared test). This significant reduction translates into an increase in the share of participants who choose the savings plan (an increase of 15%, p=0.044), but does not change the percentage of participants who choose the stock (p=0.835). Interestingly, even the two models of 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), this dimension now becomes less salient and receives smaller decision weights. As a result, their model predicts the savings plan to be chosen less frequently while the other, more liquid options, should gain popularity at its expense.[6]

The results of the second question, where participants were asked to state the proportion of the bonus for each option, further support this pattern. Comparing the distribution (and averages) of allocations of each of the options across treatments, we find lower proportions

[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).

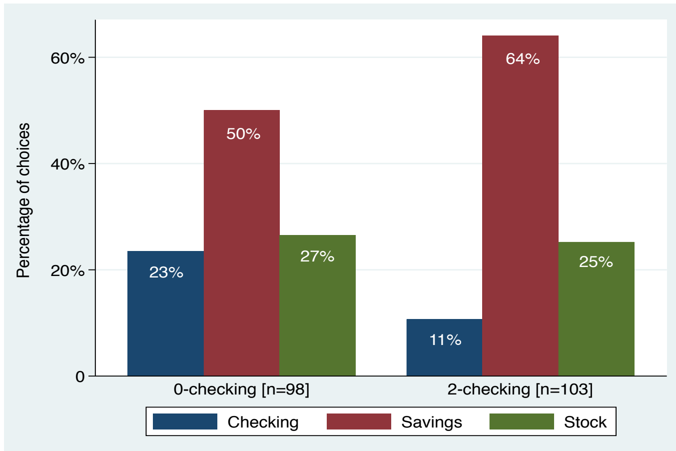


Figure II: Choice percentages of each investment per treatment in Study 2.

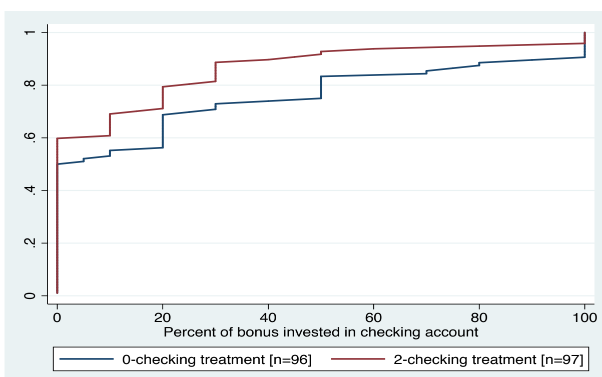


Figure III: CDF of allocation to the checking account per treatment in Study 2.

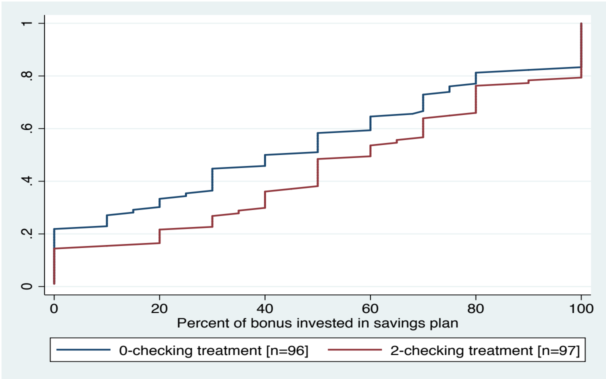


Figure IV: CDF of allocation to the savings plan per treatment in Study 2.

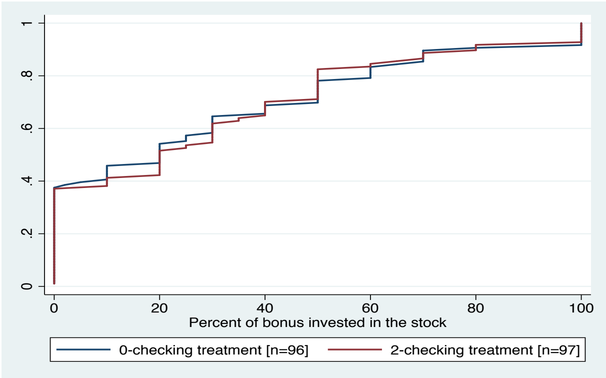


Figure V: CDF of allocation to the stock per treatment in Study 2.

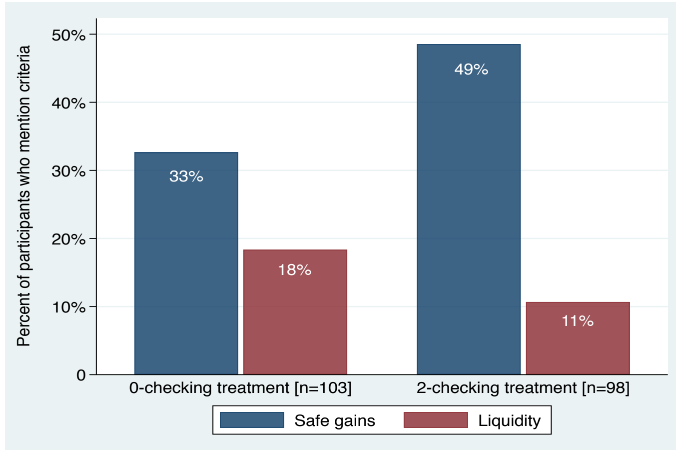


Figure VI: Criteria mentioned per treatment in Study 2.

allocated to the checking account in the *2-checking* treatment compared to the *0-checking*. This can be viewed in Figure III which shows the cumulative distribution of allocations to the checking account across treatments. 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 contribution 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 IV 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 V we see no effect on allocations to the stock across the two treatment (29% and 30%, p=0.95).[7]

Looking into participants' explanations of their choices in the first question gives a more complete picture of the decision-making process.[8] In Figure VI we see that participants refer to liquidity more often in the *0-checking* treatment (18%) compared to the *2-checking* treatment (11%) 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

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

[8]In this study, the most frequently mentioned dimensions were safe gains, liquidity and the possibility of high returns. We concentrate our discussion on the first two as the last one was mentioned to a similar extent in the two treatments and hence it seems that its relative weight did not change dramatically.

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. After all, along the safe gains dimension, which is now more prominent, it is completely dominated by the savings plan and liquidity, along which it performs better, is now shrouded and receives a lower weight. As a consequence, it is chosen less frequently in this treatment. Of course, those who still value the liquidity dimension, due to, say, debt or an urgent need for money, may very well choose it even in this case. In Appendix A, we show that the ToD model formulated in Section 3 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 values of safe gains and liquidity (as long as they are monotonic and continuous along these dimensions).

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

In our final study, we demonstrate how framing may be used to turn-on dimensions. We illustrate this in the realm of uncertainty where in two different treatments we use two different frames for the same lottery: In the first it is framed as a certain amount plus a possibility to obtain a small bonus while in the second it is described as state contingent outcomes with their respective probabilities. In the former, the certain amount that the lottery delivers is emphasized using the words "with certainty" while in the latter, the highest prize one can win in this lottery is explicitly spelled out.

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 two straightforward questions and the average completion time was about 5 minutes. As in Study 1, participants were sent a link to the questionnaire and were asked to choose between two or three options, depending on the treatment, and provide a brief explanation of their choice. Participants 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

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 IV: 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).

their choice. Table IV summarizes the options in our main treatments: *certain(3)* and *lottery(3)*. The complete questionnaire appears in Appendix B.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 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 probability of obtaining the high prize of 95 ILS that is explicitly mentioned now in two options is more emphasized 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 in the evaluation of the entire set 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)*.

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 criteria weighting should be in the same direction as in the main treatments but, in the absence of B, we do not expect the share of the first option to necessarily decrease. The reason is that the lottery features, which have been turned-on in option A', are not shared by other alternatives in the set. Thus, no other option, except for A', will gain from the larger weight given to these features, in contrast to our main treatments where option B does exactly that: it gains from the larger weight placed on the lottery features due to the framing of A'. This leads to our complete hypothesis, which states that the first option will lose more share moving from the certain framing to the lottery framing when option B is present than when it is absent.

Study 3: Results

A probit 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 as $\Phi(\tilde{Y})$ where Φ is the CDF of the standard normal distribution and $\tilde{Y}$ is specified as follows:

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

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. The benchmark treatment is taken to be *certain(2)* where participants choose between option A, framed as a certain amount

Variable	Marginal Effect
lottery(2)	0.177*
	(0.053)
certain(3)	-0.003
	(0.97)
lottery(3)	-0.35***
	(0.004)
cons	0.25
	(0.122)
N	243
R^2	0.046

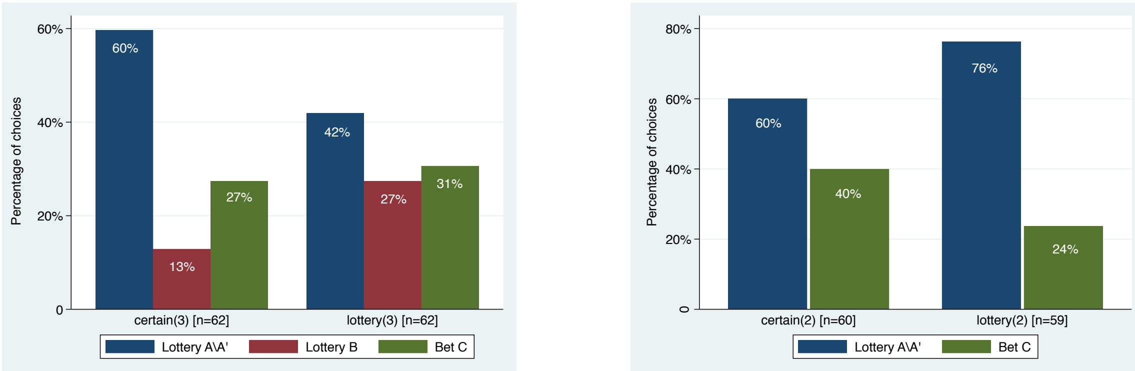
$^{***}p < 0.01$, $^{*}p < 0.1$

Table V: Marginal effects on the probability of choosing the first option in Study 3.

of money plus a possible bonus, and the Dow-Jones bet. Coefficient β_1 measures the net effect of framing option A 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). Coefficient β_3 is our main coefficient of interest - the interaction coefficient. It measures the effect of changing the frame, and adding B to the set on top of the main effects. Formally, our main hypothesis is that $\beta_3 < 0$.

Our full results are summarized in Table V. Our hypothesis is confirmed by the data as $\beta_3 = -0.35$ (p=0.004). In addition, β_2 is not significantly different from 0, and β_1 is positive, evidence of the fact that adding option B without changing the frame, or changing the frame without adding option B, does not negatively impact the frequency of 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, 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' 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*



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

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

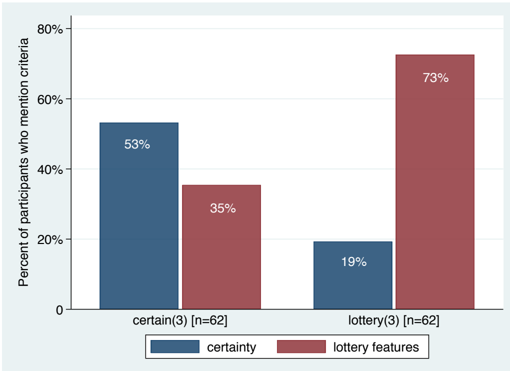
Figure VII: Choice percentages of each option in Study 3. Both panels compare the effect of framing on the choice distributions. Panel (a) does so for the choices from triplets [certain(3) and lottery(3)] and panel (b) compares the choices from binary sets [certain(2) and lottery(2)].

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, they lead the first option to be chosen less only in the presence of option B but not in its absence. To sum up, 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.[9]

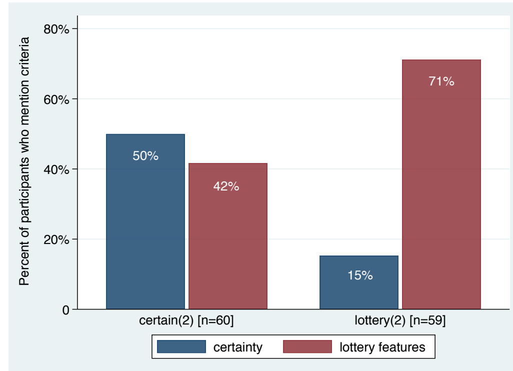
3 The ToD Model

We started by examining the idea that turning-on a dimension by making it explicit will increase its weight in the evaluation of the set. Here we formalize this idea. 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:

[9] A formal illustration of how the ToD procedure explains our findings from the four treatments in this study is given in Appendix A.



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



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

Figure VIII: Criteria mentioned per treatment in Study 3.

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

where $u_k(c_k)$ are the "classical utilities" as in KS assigned to the different dimensions and $g_k(\mathcal{C})$ are the menu-dependent-weights of each dimension.[10] The difference between our ToD model and the one proposed by KS comes from the argument of the weighting functions g_k, which measure the weight given to dimension k in the decision process. KS define these weights as follows:

Assumption 1 in KS. The weights g_k are given by $g_k = g(\Delta_k(\mathcal{C}))$, where $\Delta_k(\mathcal{C}) = \max_{c' \in \mathcal{C}} u_k(c'_k) - \min_{c' \in \mathcal{C}} u_k(c'_k)$ and the function g is strictly increasing in Δ.

This assumption implies that the weights of the different dimensions correspond to their variance in the choice set. Using the words of KS,"the decision maker focuses more on attributes in which her options generate a greater range of consumption utility." From now on, we will refer to these weights as g_k^{KS}. We would like to suggest a different determinant for these weights, one which is motivated by our studies. In order to do so, we need to define 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 $c_k > 0$.

[10]KS use the term attributes while we use dimensions, as discussed earlier.

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 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. In the context of Study 2, 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. 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 (desirable or undesirable) is turned-on if its level belongs to its set of *attractive values*.

For the case of framing we consider a dimension as turned-on if it is explicitly expressed in the description of the alternative and turned-off if it is not. Such a definition refers to language rather than numerical values and in this paper we choose not to formally define what is considered to be explicitly expressed in natural language. Nonetheless, Study 3 illustrates this channel for turning-on dimensions and Appendix A shows how the model can accommodate its findings.

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

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

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 (i.e., if some dimension is turned-on in two alternatives it will be counted twice in the denominator). In Study 2, 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.

The Tod model allows for discontinuities of weights with respect to small changes in the levels 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 an attribute's level 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 salience models, both of which capture important features of human behavior.[11] In fact, we believe one has to take into account their insights as well as ours. For example, one can imagine a more general model which takes our approach towards "turned-on" vs. "turned-off" dimensions but acts as suggested by KS when all dimensions are "turned-on" (where ToD is silent with respect to small changes in the dimension levels). Combining the models in this manner allows for continuous effects of dimension levels based on the variance of each dimension as in KS without compromising the discontinuities around the "turning-on" point of these dimensions which will be accounted for by the ToD procedure.

Remarks.

- 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, and hence their weights, may differ for different alternatives.

[11]For recent experimental support of these models see Dertwinkel-Kalt et al. (2017a), Dertwinkel-Kalt et al. (2017b) and Dertwinkel-Kalt and Köster (2018).

4 Explaining our Findings with the ToD Model

In this section we illustrate how the model can explain the findings of Study 1. A similar exercise is carried out in Appendix A to explain the findings from studies 2 and 3. Our goal in this section is to show that the ToD model *is able* to accommodate our findings rather than find the range of dimensional utility values for which it would do so. In addition, we show formally that although not every choice of dimensional utility values would lead the model to predict a preference reversal, the directional change in evaluations is independent of the choice of the consumption utility function (as long as it satisfies monotonicity in every dimension) and is in line with the behavioral pattern we observe.

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. 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 and VL in Dimension 2, $(100,130)$=(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 all alternatives and hence turned-on, while the undesirable dimension of inequality is only turned-on in the $(100,100)$ split that has a 0 level along that dimension.

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 $u_2(VL) = 1$, $u_2(\mathrm{L}) = 2, u_2(\mathrm{M}) = 3$, $u_2(\mathrm{H}) = 4$. In addition, each option has the following vector of turned-on dimensions (we use the second definition for the undesirable inequality dimension and the first definition for efficiency as it is a desirable dimension):

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

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:

$$g_1^{ToD}(\mathrm{E}) = {}^{1}\!/_{(1+1+1+1)} = {}^{1}\!/_{4},\ g_2^{ToD}(\mathrm{E}) = {}^{3}\!/_{4}.$$

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:

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

Similarly,

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

and

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

Thus, an agent in the *equal* treatment described by our utility function and abiding to the ToD procedure will rank the option $(100,100)$ first, followed by $(100,140)$ and $(100,160)$. Turning to the *unequal* treatment, we obtain:

$$\tilde{U}((100,130),U) = {}^{3}\!/_{3} \cdot 2 = 2,\ \tilde{U}((100,140),U) = 3,\ \tilde{U}((100,160),U) = 4.$$

In the *unequal* treatment the ordering is reversed, in line with our findings for a significant percent of participants. Shrouding the inequality dimension by replacing the all-equal split with $(100,130)$ alongside the enhancement of the efficiency dimension is the driving force behind the observed preference reversal. It is important to note that while our choice of utility values leads to the observed reversal, **any** choice of values would push preferences in the same direction. Thus, while some values may not lead to an actual reversal in our study they would all increase the relative evaluation of $(100,160)$ compared to $(100,140)$ when replacing the equal split with the unequal one. Being more precise, the change in value of $(100,160)$ amounts to:

$${}^{1}\!/_{4} \cdot u_2(\text{H}) - {}^{1}\!/_{4} \cdot u_1(\text{H}),$$

while the change in value of $(100,140)$ equals:

$${}^{1}\!/_{4} \cdot u_2(\text{M}) - {}^{1}\!/_{4} \cdot u_1(M).$$

Given that Dimension 1 is undesirable and Dimension 2 is desirable, it is evident that the former expression is larger than the latter for any choice of intrinsic utility values.

5 Discussion and Related Literature

5.1 Our Model and Related Theories

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 criteria stand out more than others and receive larger weights 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, salience is determined by the variation of each dimension in the choice set and it applies uniformly to the assessment of the members of the set. In our model, salience is also determined by the choice set and applies to the entire set as in KS and hence, for purpose of the current discussion we focus on the comparison between their model and ours.[12]

The main difference between our model and KS lies in how weights of different dimensions are determined. In KS, a dimension with a larger range will become more prominent and receive larger weights. In the ToD model, a dimension's prominence is determined by the number of alternatives that explicitly express that dimension. In this sense, our model is more discontinuous than KS. For example, slightly decreasing the level of some dimension of one alternative is likely to affect 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.

This difference generates different predictions of choice behavior. For example, in Study 2, the "safe gain" dimension's range 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, in contradiction 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 the social preferences of Study 1. 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. By contrast, the ToD model predicts a larger weight placed on egalitarian considerations when [100,100] is present due to its explicit reflection of equality. As a consequence, preferences are expected to shift and express a stronger positive attitude toward egalitarianism.

A closely related approach, 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

[12] 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.

in KS) on dimensions with larger variance of consumption utility.[13] 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 thinking model. For example, consider Study 2. As we mentioned earlier, we ran a very similar study which compared choices across the same sets as in Study 2, where one had a checking account with a tiny interest rate of 0.1% and another with a checking account with no interest. A similar distribution of choices arises when the checking account carries a 0.1% interest rate or 2%. We suggest that as long as the interest rate is strictly greater than 0 the safe gain dimension is turned-on in the checking account, generating the same dimensional weights across the two experimental versions. By contrast, according to the relative thinking theory of Bushong et al. (2017) increasing the interest rate from 0.1% to 2% decreases the utility variance along the safe gain dimension, and therefore this dimension should receive a higher weight in the *2-checking* treatment. Thus, their model would predict a higher share of choices of the savings plan when the checking account has an interest rate of 2% than when it has 0.1%.

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 3 we suggested that one could come up with a model which combines our insights alongside those of KS at the stage in which weights are determined. As these approaches 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, following the sketch of Bushong et al. (2017), one may consider a model in which facing multiple dimensions, variance and turned-on dimensions considerations lead the agent to concentrate on two dimensions for which he applies relative thinking to reach his final choice.

Our findings may be explained, at least to some extent, not only through the lens of dimensional weighting. Categories 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., 2017).[14] To illustrate, we follow Manzini and Mariotti (2012) and consider the investment example in Study 2. 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

[13]Other approaches to relative thinking have been suggested by Azar (2007) and Cunningham (2013). For experimental evidence of relative thinking see, for example, Azar (2011).

[14]For a different approach involving categories and reference points see Barbos (2010) and Maltz (2017).

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 the savings plan. Categorization, however, does not seem to apply to the findings from the social preferences study 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 our findings may be addressed is choice by iterative search, suggested by Masatlioglu and Nakajima (2013). In their model, the agent starts off with some default option or 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 2), 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 1, where preferences are actually reversed, a phenomenon that is hard to reconcile through the channel of consideration sets or categories.

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 gain 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 with 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 three patterns. We suggest the ToD procedure that draws on the literature on salience and focusing, while adding the role of "turned-on" dimensions to relative weighting and accounts for the discontinuous nature of our findings. This novel aspect of the model generates predictions that are in line with the findings of all three studies. In addition, the analysis of participants' explanations provides further support for this procedure.

5.2 Experiments

We would now like to relate our findings to experiments reported in the psychology and economics literature. For example, 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) as well as in the domain of uncertainty (Gneezy et al., 2006; Bateman et al., 2007). These studies focus on the intrinsic valuation of goods and argue that sometimes 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, is not focused on intrinsic values of alternatives. We argue that the apparent violation of monotonicity found in Study 2 is due to the shift of dimensional weights and its effect on the other options in the choice set, rather than the checking account being deemed worse when it generates a positive interest rate. 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.

Violations of monotonicity may also arise when options are evaluated separately, as shown by Hsee (1998). These violations, which vanish when options are evaluated jointly, have been explained by the evaluability hypothesis (Hsee et al., 1999). It posits that when there are two dimensions, one easy to evaluate and the other difficult, the difficult one may have little impact on choice in the separate evaluation. When options are presented together, the difficult dimension becomes somewhat easier to evaluate and monotonicity of preferences is restored. By contrast, all of our findings (including the violation of monotonicty in Study 2) appear in joint evaluations of three-alternative choice sets and therefore cannot be explained by the evaluability hypothesis.

Our studies also share commonalities with experimental work on comparisons along different attributes.[15] 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 a similar vein, Palmeira (2010) suggests that subjects find it easier to compare two positive values of a given attribute than a positive value and a zero value of that same attribute. He claims that compared to zero, any number is infinitely larger, and so it becomes meaningless to make a comparison between them. He provides evidence of apparent violations

[15] Notice that here we use the term attributes as it is the term most commonly used in this literature.

of monotonicity in binary choices by manipulating attribute values from 0 to small positive levels, findings that are similar in spirit to those we report in Study 2. 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 is based on the salience model of Bordalo et al. (2013), and incorporates framing effects to account for these findings. 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 equal zero for some alternatives (or not common to all alternatives). In fact, if some positive attribute i is originally turned-off in the entire set and later on 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 a negative attribute (i.e., it is a common attribute) it will receive no weight in the assessment of goods accroding to the ToD procedure (as inequality in the *uneqaul* treatment in Study 1). 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 1) even though it is now not a "common attribute" anymore.

6 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' prominent criteria 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 consider the possibility of increasing safe investments through the introduction of positive interest rates in checking accounts that currently carry no interest. 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). 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

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 (2012): "Salience theory of choice under risk," *The Quarterly journal of economics*, 127(3), 1243–1285.

——— (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.

Cunningham, T. (2013): "Comparisons and choice," *Unpublished Manuscript, Harvard University.*

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.

——— (2018): "Violations of first-order stochastic dominance as salience effects," *Working Paper*.

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

Furtado, B. A., L. Nascimento, and G. Riella (2017): "Rational Choice with Categories," *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.

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.

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.

Maltz, A. (2017): "Exogenous Endowment - Endogenous Reference Point," *Working Paper*.

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.

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

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

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.

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.

Appendix A

Explaining Study 2 with the ToD Model

We perform a similar exercise to the one we held in Section 4 to show that the ToD model can accommodate our findings from Study 2. We also show that while not every choice of utility values would lead the model to predict our observed violation of monotonicity, adding a small enough interest rate to the checking account *would* indeed pull evaluations in the direction of our observed pattern of behavior.

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 (higher than 10%).[16] Dimensions are numbered 1, 2, 3 respectively. We assume four *levels* (0,L,M,H) of these dimensions 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 gain 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.[17]

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 others. 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 : $u_3(0) = 0$, $u_3(\mathrm{H}) = 2$. In addition, each investment option has the following vector of turned-on dimensions (the dimensions that have some positive level, according to definition 1): $checking - 0\%^{ToD} = (0, 1, 0)$, $checking - 2\%^{ToD} = (1, 1, 0)$, $savings^{ToD} = (1, 1, 0)$, $stock^{ToD} = (0, 1, 1)$.

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:

$$g_1^{ToD}(No - Int) = {}^{1}/_{(1+1+1+1+1)} = {}^{1}/_{5}.$$

Similarly, we obtain:

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

[17] 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.

$$g_2^{ToD}(No-Int) = 3/5,\ g_3^{ToD}(No-Int) = 1/5.$$

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:

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

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 dimensional valuations 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 in our cardinal exercise 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. While for some utility values the model would predict no reversal, the relative change in utilities operates in the direction of our observed behavioral pattern *for any choice of values*, as long as the interest rate added to the checking account is small enough and utilities are monotonic and continuous in every dimension.

To see this, we examine the changes in evaluations without specifying utility values for each dimension. The increase in the evaluation of the savings plan due to the introduction of the 2% checking account 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 gain 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 $10/30 \cdot u_1(\mathrm{L}) - 3/30 \cdot u_2(\mathrm{H})$. 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) 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 (which may or may not take place depending on initial utility evaluations in the *0-checking* treatment). Reflecting on Study 2 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.

Explaining Study 3 with the ToD Model

Here we show how the model is able to predict the findings from Study 3. As in the previous exercises, following the numerical example, we show that these predictions are general enough, in the sense that they pull in the direction of our findings regardless of the exact choice of utility values along the relevant dimensions (as long as continuity is maintained). Once again 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).[18] The study focuses on the explicit mention of Dimension 1 in option A' compared to option A in which it is not mentioned. 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. Each option has the following levels along the different dimensions: Option A=(L, H, 0), option A'=(L,H,0), option B=(H,0,0), and option C=(0,0,H).

Here is an explanation for the choices of different levels for each option: Options A and A' are the same so they receive the same levels in all dimensions. Specifically, they have a low probability (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 options 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.

We assume that the decision maker has the following evaluations along dimensions: $u_1(0) = 0$, $u_1(\mathrm{L}) = 7, u_1(\mathrm{H}) = 9$, $u_2(0) = 0, u_2(\mathrm{H}) = 3$, and $u_3(0) = 0, u_3(\mathrm{H}) = 9$. These reflect monotonicity in each dimension with the first dimension having marginal decreasing effects. In addition, a higher value is attached to the possibility of earning over 100 ILS than for the minimal prize being greater than 50 ILS. 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

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

the alternative. Specifically, each option has the following vector of turned-on dimensions:

$$A^{ToD} = (0,1,0),\; A'^{\,ToD} = (1,1,0),\; B^{ToD} = (1,0,0),\; C^{ToD} = (0,0,1).$$

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 and A'. Alternative C is the only one in the set that has Dimension 3 turned-on.

ToD weights in the *certain(3)* treatment:

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

In the *lottery(3)* treatment, the weights are different due to the different framing:

$$g_1^{ToD} = {}^{2}/_{4},\; g_2^{ToD} = {}^{1}/_{4},\; g_3^{ToD} = {}^{1}/_{4}.$$

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

$$\tilde{U}(A, \{A, B, C\}) = {}^{1}/_{3} \cdot u_1(\mathrm{L}) + {}^{1}/_{3} \cdot u_2(\mathrm{H}) + {}^{1}/_{3} \cdot u_3(0) = {}^{1}/_{3} \cdot 7 + {}^{1}/_{3} \cdot 3 + {}^{1}/_{3} \cdot 0 = {}^{10}/_{3}.$$

Similarly,

$$\tilde{U}(B, \{A, B, C\}) = {}^{1}/_{3} \cdot 9 + {}^{1}/_{3} \cdot 0 + {}^{1}/_{3} \cdot 0 = {}^{9}/_{3},$$

and

$$\tilde{U}(C, \{A, B, C\}) = {}^{1}/_{3} \cdot 0 + {}^{1}/_{3} \cdot 0 + {}^{1}/_{3} \cdot 9 = {}^{9}/_{3}.$$

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

$$\tilde{U}(A', \{A', B, C\}) = {}^{2}/_{4} \cdot 7 + {}^{1}/_{4} \cdot 3 = {}^{17}/_{4},\; \tilde{U}(B, \{A', B, C\}) = {}^{18}/_{4},\; \tilde{U}(C, \{A', B, C\}) = {}^{9}/_{4}.$$

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 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 in the first alternative that was turned-off in the certain payment framing. Thus, a higher weight is now given to this dimension, which benefits options A' and B but the effect on B is larger since it performs best along that dimension (at the same time, the evaluation of option A' is also hurt to some extent since the high minimal prize dimension is now shrouded). Overall, the evaluation of A' increases but by a lesser amount than the evaluation of B which is now the highest in the set.

Notice that moving from *certain(3)* to *lottery(3)*, the change in evaluation of B equals $1/6 \cdot u_1(\mathrm{H})$, which is strictly positive regardless of the choice of utility values. Thus, the ToD procedure predicts 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: $1/6 \cdot u_1(\mathrm{L}) - 1/12 \cdot u_2(\mathrm{H})$, 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 the utility function continuous, the overall evaluation of the first alternative will not increase. Thus, while for some lotteries (those with a high enough probability of the high prize) the model will not generate the effect we find as a prediction, if we make our grid finer and choose lotteries with low enough probabilities for obtaining the 95 prize, we are bound to generate a prediction in line with our reported choice reversal.

To complete the picture we show how the model with the above 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)*:

$$g_1^{ToD} = g_2^{ToD} = g_3^{ToD} = 1/3.$$

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

$$\tilde{U}(C, \{A, C\}) = 9/2.$$

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

$$\tilde{U}(A', \{A', C\}) = 10/3, \ \tilde{U}(C, \{A', C\}) = 9/3.$$

Thus, we observe that when option B is absent, the change of frame highlights the first dimension in a way that shifts choices from C in treatment *certain(2)* to A' in treatment *lottery(2)*. Notice that in the absence of B, Dimension 1 receives 0 weight in treatment *certain(2)* since it is turned-off in both alternatives. When A it replaced by A' this dimension is turned-on and leads to a relatively large shift in weight from 0 to $1/3$ leading to the pattern we observe across these binary choice treatments.

Appendix B

Below are the English translations for the instructions 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 B.1. Study 1: 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.[19] 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.

[19] 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 others 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 B.2. Study 2: 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.[20]
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.
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.

[20] 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.

Note: If the amount (or part of it) is withdrawn before an entire year has passed, you will receive the proportional share of the expected 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 B.3. Study 3: Instrctions 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.

Decision Making Questionnaire - General Instructions

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: [ ]

<u>Question 1</u>

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.

<u>Question 2</u>

Please briefly explain your choice: [ ]