האוניברסיטה העברית בירושלים

THE HEBREW UNIVERSITY OF JERUSALEM

ADVICE-GIVING UNDER CONFLICT OF INTEREST: CONTEXT ENHANCES SELF-SERVING BEHAVIOR

By

MEIR BARNERON AND ILAN YANIV

Discussion Paper # 750 (April 2023)

מרכז פדרמן לחקר הרציונליות

THE FEDERMANN CENTER FOR THE STUDY OF RATIONALITY

Feldman Building, Edmond J. Safra Campus,
Jerusalem 91904, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Advice-Giving under Conflict of Interest:

Context Enhances Self-Serving Behavior

Meir Barneron[1,2]

Ilan Yaniv[1,3]

[1]Department of Psychology, Hebrew University of Jerusalem

[2]National Institute for Testing and Evaluation, Jerusalem, Israel

[3]Federmann Center for the Study of Rationality, Hebrew University of Jerusalem

This research was supported by a grant from the Israel Science Foundation (327/10) to the second author and a doctoral fellowship from the Federmann Center for the Study of Rationality to the first author. Correspondence concerning this article should be addressed to either author email: meir.barneron@gmail.com, ilan.yaniv@mail.huji.ac.il.

Abstract

Advisors face a conflict of interest when their interests and those of the recipients of their advice are misaligned. Conflicted advisors need to resolve the tension between two competing motives, the need to provide sincere advice that fulfills the recipient's goals and the temptation to give advice that caters to their self-interest. We theorized that the choice context should affect selfish advice-giving. Our basic experimental condition presented the advisor with two alternative recommendations, one optimal for the recipient, and one preferable for the advisor. We hypothesized that introducing a third (inferior) alternative (in the context condition) should increase the advisor's tendency to give selfish advice. In Study 1, advisors who were instructed to transmit a recommendation to an anonymous recipient, were more selfish in the context than in the basic condition. Study 2 further found that the effect was obtained when the third alternative was strictly dominated by the selfish recommendation. Studies 3-4 tested the idea that forewarning advisors about the need to explain their choices should moderate the effect. Study 5 tested the advisors' awareness of the context effect. Studies 6-7 investigated the reactions of advice recipients and social observers to selfish advice-giving and found them also biased by context. Our theoretical account posits a reference-based evaluation process. This mechanism explains the advisors' tendency to give selfish advice as well as the social actors' reactions to the transmission of such advice. We discuss the context effect in relation to the asymmetric dominance effect, social preferences, and ethical decision making.

Keywords: conflict of interest, advice-giving, social preference, asymmetric dominance effect, self-serving bias; recommendations

Advice Giving under Conflict of Interest:

Context Enhances Self-Serving Behavior

Decision-makers often seek the advice of others – friends, colleagues and even professionals – who could help them make better decisions. Studies indeed show that the recipients of advice stand to gain from the utilization of sincere advice (Yaniv, 2004; Yaniv & Choshen-Hillel, 2012). The recipients can benefit, in particular, from the advice of laypeople or experts who keep their recipients' best interests in mind and deliver truthful information.

Advice however may not always be sincere. Advisors do not always adhere to the cooperative norm of social interaction (Grice, 1975). In shopping, getting a car fixed, banking, buying real estate, or seeking medical help, the advice received from experts could be tainted by personal interests of the advisors, interests that are not aligned with those of the advice seekers. For example, in the interest of generating revenue, a banker might recommend a loan that is not optimal for the client, or a physician might recommend a test that is unnecessary for the patient. Friends and relatives as well could give advice that is tainted by personal agendas. In general, a conflict of interest arises when an advisor faces two alternative courses of action, one of which is favorable for the recipient, while the other is preferable for the advisor.

An important task for psychological research is to investigate the factors that enhance (or diminish) an advisor's tendency to give selfish advice, which is also suboptimal for the recipient. A number of factors have been found to affect an advisor's tendency to give self-serving advice, including identifiability (Sah & Loewenstein, 2012), disclosure (Cain, Loewenstein, & Moore, 2005), and role-taking biases (Moore & Loewenstein, 2004; Moore, Tanlu, & Bazerman, 2010).

The present research was inspired by informal observations of conflicted advisors who, in attempt to rationalize selfish advice-giving, introduced a reference against which the selfish advice appeared attractive. How might a newly-introduced reference alternative cause such a

shift in choice? We build on findings from the area of consumer decision-making (Simonson, 1989). Suppose A and B are two similar computers, except that computer A is better than B in terms of screen quality, but computer B is better than A in terms of price. Thus, a consumer choosing between these two computers should resolve the tradeoff between screen quality and price.

Past findings suggest that consumers resolve this tradeoff in a predictable way when a third alternative (C), inferior to one of the existing alternatives, is added to the consideration set (Huber, Payne, & Puto, 1982; Simonson & Tversky, 1992). Suppose computer C is inferior to B because its screen is worse, while it costs as much as B does. Since C is inferior to B, the consumer should drop it from consideration, and then proceed to choose between A and B only. Studies find however, that while C, being inferior, is rarely chosen, adding it the set increases the consumers' tendency to choose B over A, compared with the situation where C is not included. Thus, B is evaluated more highly in the presence of C, the dominated alternative (Highhouse, 1996; Simonson, 1989).

We build on this insight in our study of selfish advice-giving under conflict of interest. Consider an advisor who faces two alternative recommendations (A and B), where A is optimal for the recipient, and B is preferable for the advisor. This basic conflict-of-interest setting is designed to pit in the advisor's mind two opposing motives, self-interest and caring for the recipient's interest. Self-interested behavior should come at the emotional cost of violating social norms, such as honesty and helping a person in need (Gneezy, 2005).

Suppose now the advisor's basic consideration set is augmented to include a third alternative, C, which is worse than B for the recipient, while it is not any better for the advisor. In this setting, C is inferior to the selfish recommendation B, suggesting the advisor should delete it from consideration, and proceed to choose between A and B only. We predict however, that rather than ignoring C, the advisor will use it as a reference in evaluating B, increasing B's favorability, and hence, the chances of selfish advice-giving.

This effect ties in with a more general psychological principle, whereby the perception of a target stimulus varies as a function of the context in which it appears. For example, the physical properties of a target stimulus (e.g., its size or brightness) are enhanced or diminished, depending on comparison stimuli that precede the target in time or conjoin it in space (Roberts, Harris, & Yates, 2005). Such context effects occur also with hedonic stimuli. The evaluation of a music piece is affected by the quality the music piece played before (Parker, Bascom, Rabinovitz, & Zellner, 2008) and judgments of attractiveness of faces are made in relation to other faces presented (Cogan, Parker, & Zellner, 2013). In sum, we posit that the reference-based evaluation mechanism should bias conflicted advisors to view selfish advice-giving more favorably, given a contextual alternative of the sort described.

The present setup extends the asymmetric dominance paradigm (e.g., Simonson, 1989). While in past work, the decision makers focused their attention on product characteristics, here advisors are supposed to weigh two fundamental social motives, self-interest and caring for the other. The advisors presumably go beyond directing their attention to the particular option characteristics, as they are presumably motivated to find signals that would make the selfish option appear more acceptable.

Social Perception of Selfish Advice-giving

According to our theory, context should affect not only the behavior of advisors, but also the perceptions of the recipients. Imagine a person who had received an advisor's recommendation and acted on it. Suppose this person then finds out that the advisor had given her poor advice in order to advance her selfish goals. The recipient now pauses to evaluate the advisor's behavior.

While the recipient's reaction is surly expected to be negative, it should also depend on her understanding of the advisor's set of options. The posited reference-based evaluation process implies that the recipient should view the advisor's selfish actions less negatively in the

context condition (where the third, contextual option is included) than in the basic condition (where the contextual option is not included). In other words, the recipient should be more forgiving towards the receipt of selfish advice in the context condition.

We also investigated the perceptions of outside audience. The impressions formed by social observers about an advisors' behavior should matter to advisors. We hypothesize that observers should also be more forgiving towards advisors who transmit selfish advice in the context than in the basic condition. Our theoretical framework yields the seemingly paradoxical result, whereby recipients and social observers should judge selfish advisors less harshly in the very same condition where they are more likely to witness selfish advice, or even suffer from receipt of such advice.

Description of the Paradigm

In order to test these hypotheses, we have developed a game paradigm, involving two roles, advisor and recipient of the advice. The advisor was told that she or he was paired with an anonymous recipient whose task was to pick a card from a set, reveal its value, and collect this value in cash. Furthermore, the advisor was informed about the value of each card in the set (of two or three cards) and also that the recipient was oblivious regarding the cards values, prior to making a choice among them. The sole source of information to be had by the recipient about the card values was the recommendation delivered by the advisor paired with her/him.

This setting simulates common, everyday situations in which advisors are in possession of information that others might need in order to make optimal decisions. In the present setup, a conflict of interest is built into the game via the payoff structure for the advisors and recipients. Thus, in addition to knowing the card values, the advisors know also that they will be remunerated contingent on their recommendation.

The basic experimental paradigm compares two conditions. The basic condition involves two cards, A and B, such that A is the optimal recommendation for the recipient and

B is the self-serving choice for the advisor. The context condition includes three cards, A, B, and a third card C, which is dominated by B. We conducted seven studies testing the context effect in various advisor-recipient settings. In five studies, we investigated advisors' behavior and, in two, we investigated social actors' perceptions of selfish advice-giving.

Study 1 tested the hypothesis that adding the third alternative C should increase the advisors' tendency to recommend the self-serving card B. Study 2 tested the boundary conditions for obtaining the context effect. Studies 3-4 tested the hypothesis that forewarning advisors that they would be asked to justify their choices should moderate the strength of the context effect. In Study 5, we tested the advisors' awareness of the context effect. In Study 6, we tested the reactions of recipients who had been given suboptimal, selfish advice. In Study 7, we tested the reactions of social observers who had been exposed to selfish advice-giving. In these studies, we report all measures, manipulations and exclusions. Sample sizes were determined before all data analyses reported in this paper.

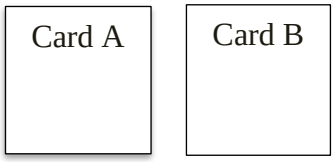
Study 1

Method

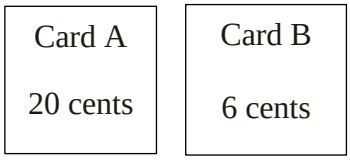
Participants. Participants (N = 90) were recruited via the Amazon Mechanical Turk website. They were offered a base fee of $0.30 and an opportunity to receive an additional payment depending on their decisions.

Procedure and design. The experimental procedure was administered on a computer online. The instructions on the first screen informed the participants that the experiment would involve a communication game with two roles: advisor and advice recipient. Every advisor was supposed to be paired with an anonymous recipient of the advice. The participants were told that they would be assigned one of the two roles, at random. In fact, they were all assigned the advisor role.

In the basic condition, two cards were presented on the screen, face down, as shown below:



The participants (in the role of advisors) were asked to click on each card to reveal its monetary value for the recipient. Clicking on A and B revealed the values of 20 cents and 6 cents, respectively.

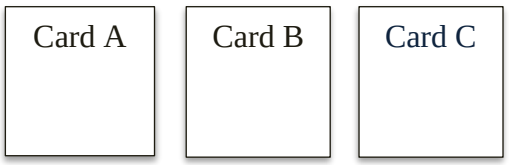


The next screen informed the advisors that they could receive additional payment, depending on which recommendation they choose to transmit to the recipients. The screen read:

- Should you choose to recommend card A, you will receive 5 cents, in addition to the payment for your participation.
- Should you choose to recommend card B, you will receive 10 cents, in addition to the payment for your participation.

They were further told, "The recipient will *not* know the value of the cards prior to making a choice. Thus, the only information available to the recipient will be your recommendation." The final screen displayed the card values for the recipient and the advisor's payment for recommending each card. The participants were then prompted to select the card that they would recommend to the recipient. For their final payment, they received the base fee plus the payment that was based on their recommendation.

In the context condition, three facedown cards were displayed on the screen:



Here too the participants were instructed to click on each card to reveal its value. The values of A, B and C (for the recipients) were 20 cents, 6 cents, and 1 cent, respectively.

The participants were then informed that they could get an extra payment depending on their recommendations. They were told:

- Should you choose to recommend card A, you will receive 5 cents in addition to the payment for your participation.
- Should you choose to recommend card B, you will receive 10 cents in addition to the payment for your participation.
- Should you choose to recommend card C, you will receive 10 cents in addition to the payment for your participation.

As in the basic condition, they were further told, "The recipient will *not* know the value of the cards prior to making a choice. Thus, the only information available to the recipient will be your recommendation."

In sum, the design included two between-participants conditions, basic and context. The procedure involved only one round of the game, thus advisors made only one recommendation. The payoffs were such that Card A was the optimal recommendation for the recipient, and card B was suboptimal for the client, but produced a higher payoff for the advisor. Card C was dominated by card B, as it offered the same payment to the advisor and had a lower value for the recipient, compared to card B.

After giving their recommendations, all the participants were asked to recall the value of Card A for the recipient. Two participants (2.22% of the sample), who failed this check, were excluded, leaving 88 participants for the analysis.

Results and Discussion

Table 1 displays the distribution of the various recommendations in each condition. Our dependent variable was the advisor's recommendation. We conducted two related analyses of the data. In one analysis, the responses were coded in a binary format. Recommending the optimal card (A) for the recipient was coded 1 and recommending either of the nonoptimal cards (B or C) was coded 0. As predicted, the participants chose the optimal recommendation (card A) less often in the context than in the basic condition, 45.2 vs 69.6%, respectively, χ^2 (1, N = 88) = 5.33, p = .021, d = 0.56. (The effect size was derived from proportions, based on Cooper, Hedges, and Valentine, 2009.)

In a second (related) analysis, we tested the increase in the proportion of self-serving recommendations (B) in the context condition. To allow proper comparison with the basic condition, we excluded from the data the C responses (and thus calculated an adjusted estimate of the proportion of B responses out of the A or B responses). The proportion of B responses was higher in the context than in the basic condition, 53.7 vs 30.4%, χ^2 (1, N = 87) = 4.82, p = .028, d = 0.54. These findings are consistent with our hypothesis. The presence of the reference alternative (card C) enhanced the appeal of the self-serving recommendation (card B) and, thus, reduced the proportion of optimal recommendations for the clients.

Table 1

The Distribution (%) of Recommendations in Study 1

Condition	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	69.6	30.4	-
Context	45.2	52.4	2.4

Study 2

Study 1 found that context affected the advisor's tendency to give selfish advice. Specifically, the context condition was created by adding a third (dominated) alternative to the basic consideration set. The third alternative was dominated by the self-serving alternative, thus, technically, it could be deleted or ignored. Indeed, the third alternative (C) was almost never chosen, but it changed the choice context. Specifically, it provided a new reference in the evaluation of the remaining alternatives which served to enhance the evaluation of the selfish alternative, and thereby, the chances of choosing it.

In the present study, we explore some of the boundary conditions for obtaining this context effect, and in particular, the characteristics of the third (reference) alternative that play a critical role in creating the context effect. Study 2 included one basic condition and four context conditions in which we varied the third alternative. In particular, we manipulated the advisor's incentive for recommending the reference alternative, while keeping the other payoffs the same across conditions.

As shown in Table 2, the advisor's payoff was varied from 1 to 19 cents. The specific pay levels were chosen such that in the first two context conditions (C-10 and C-1), the reference alternative remained dominated. In the two other context conditions (C-11 and C-19), the advisor incentives were such that the reference alternative was no longer dominated by the

self-serving option. Moreover, the incentives for selfish behavior increased. This allowed us to test not only the role of dominance in creating the effect, but the effect of increasing the incentives and also making them either near or far apart from the advisor's payoff for giving the selfish option (B).

Method

The study was preregistered at: https://doi.org/10.17605/OSF.IO/FMV9J.

Participants. We recruited 350 participants (49.8% males, mean age 40.3) via the Amazon Mechanical Turk website. The participants were offered a participation fee of $0.30 and an opportunity to receive additional payment, depending on their recommendations to the recipients. Power analysis was used to determine the sample size based on the results of Study 1. It suggested that 63 participants per condition were needed to detect the effect with power of .80, and $p = .05$. We aimed for 70 participants per condition, anticipating some attrition.

To ensure the quality of the data, we presented two probes at the start of the study, a captcha question and a comprehension question, "How many cards are involved in the game?" which participants had to answer correctly (2 or 3, depending on the condition) in order to continue the study. At the end of the study, we included a memory question, "What was the value of card A?" Of the participants who completed the study 11 (3.14%) failed the memory question and were excluded, thus, the final sample included 339 participants. The *N*s for the five conditions ranged from 66 to 70.

Procedure and design. The experimental procedure used the same game as in Study 1. The study involved five between-participant conditions (see Table 2). The basic condition involved two cards, A and B, identical to those used in Study 1. The four context conditions involved three cards, A, B and C. The conditions varied only in terms of the advisor's payoff for recommending C, from 1 to 19 cents. The payoffs associated with cards A and B were held the same as in the basic condition.

Table 2

The Values of the Cards (in Cents) for the Recipients and the Payoffs for the Advisors in the Basic and Four Context Conditions of Study 2

Condition		Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	Recipients	20	6	-
	Advisors	5	10	-
Context: C-10 (dominated)	Recipients	20	6	1
	Advisors	5	10	**10**
Context: C-1 (dominated)	Recipients	20	6	1
	Advisors	5	10	**1**
Context: C-11 (not-dominated)	Recipients	20	6	1
	Advisors	5	10	**11**
Context: C-19 (not-dominated)	Recipients	20	6	1
	Advisors	5	10	**19**

Note: Only the numbers indicated in bold were varied across the four context conditions. No part of this table was presented to the participants of the studies. The participants discovered the values for the recipients by clicking on the cards (as in Study 1) and then the advisor payoffs provided (see methods of Study 1).

As shown in Table 2, in the first two context conditions (C-1, C-10), the third alternative was dominated by the self-serving one. The advisor's payoff for recommending the third card (C) was either the same as for recommending B (10 cents), or much lower (1 cent) than the payoff for recommending B, the self-serving card. Note that the C-10 condition is identical to the context condition in Study 1. In the two other context conditions (C-11, C-19), the third alternative was not dominated, as the advisor's payoff was *larger* than that given for recommending the selfish alternative (B). These two conditions differed in terms of the size of

the advisor's payoff. In C-11, the incentive for recommending C (11 cents) involved a minimal increase from C-10. In C-19, the incentive for recommending C was substantially larger.

Results

Table 3 displays the distribution of the recommendations made in each condition. In the basic condition, the majority of the advisors (66.7%) provided the optimal advice. In the four context conditions, the percentage of optimal advice-giving was significantly lower (41.8, 42.4, 48.6 and 38.8%, respectively). Four chi-square tests affirmed this conclusion, $\chi^2 > 3.95$, *N*s > 135, *p*s < .047.

Inspection of the percentages of B responses (selfish advice) in the four conditions sheds light on the origins of the context effect. An increase in the percentage of B responses (relative to the basic condition) was obtained *only* when the third alternative (C) was dominated. To allow for proper comparison of the context conditions with the basic condition, we calculated the percentage of the B choices out of the set of A or B choices only (i.e., excluding the C choices), as was done in the previous study. The adjusted percentages differed in some conditions from the raw percentages shown in Table 3. As in Study 1, we found more B choices in C-10 than in the basic condition, 54.1 vs 33.3%, χ^2 (1, N = 130) = 4.88, p = .027, d = 0.47. Similarly, there were more B choices in C-1 than in the basic condition, 57.6 vs 33.3%, χ^2 (1, N = 135) = 7.06, p = .008, d = 0.55. Conditions C-10 and C-1 did not differ from one another in terms of B choices, 54.1 vs 57.6%, respectively, χ^2 (1, N = 127) = 0.05, p = .829, d = 0.08. These results are in line with the idea that the third alternative (C) provided a reference in the evaluation of the selfish alternative, that although it was barely chosen, it affected the tendency to choose B as a recommendation.

No context effect was obtained in C-11 and C-19, where the third alternative (C) was not dominated. The percentage of B choices in C-11 did not differ from that in the basic conditions, 39.3 vs 33.3%, respectively, χ^2 (1, N = 125) = 0.25, p = .616, d = 0.14. The

percentage of B choices in C-19 did not differ from that in the basic condition, 36.6 vs 33.3%, respectively, χ^2 (1, N = 110) = 0.02, p = .889, d = 0.08. Finally, conditions C-11 and C-19 did not differ from one another, 39.3 vs 36.6%, respectively, χ^2 (1, N = 97) = 0.003, p = .953, d = 0.06.

Table 3

The Distribution (%) of Recommendations in the Basic and Four Context Conditions of Study 2

Conditions	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	66.7	33.3	-
Context: C-10 (dominated)	41.8	49.3	8.9
Context: C-1 (dominated)	42.4	57.6	0.0
Context: C-11 (non-dominated)	48.6	31.4	20.0
Context: C-19 (non-dominated)	38.8	22.4	36.8

Discussion

Study 2 explored the boundary conditions for obtaining the context effect. The results establish the idea that only dominated alternatives produce the context effect. The context effect was held strong when the third recommendation was dominated, regardless of the level payoff for the advisor. The context effect was not found when the reference was not dominated. In condition C-11 (see Table 2), the advisor's payoff for the C response was very close to that for the B response as it differed by just one cent (11 vs 10), thus, C was no longer dominated by B. This sufficed to eliminate the context effect, suggesting the critical role of dominance. In sum, in conditions C-1 and C-10, card C was inferior and, thus, not a viable

alternative for choice. In these two conditions, however, card C provided a reference for evaluating the selfish alternative B and thus it created the context effect.

We now turn to the remaining results in Table 3. In conditions C-11 and C-19, the third possible recommendation, card C, was a viable alternative. Specifically, cards A and C represented extreme choices, where advising A was a prosocial move, and advising C involved extreme selfishness (highest possible payoff for the advisor and lowest payoff for the recipient). The percentage of participants advising C increased from 20 to 37%, as the incentive for the advisor grew from 11 to 19 cents. Thus, selfish advising increased when it paid more to do so. Importantly, advising B in these conditions is a compromise between the two extreme alternatives. A tendency to seek a compromise was documented in previous studies of consumer choice (Simonson, 1989) as well as in studies involving economic (dictator-like) games (cf. List, 2007). Advising B (between the two extreme options) enabled the advisors to cater to two opposite social orientations, self-interest and the pro-sociality. While the compromise effect was not the topic of our present research, it is of interest in its own right, and thus should be the topic of further studies.

Study 3

Do advisors give selfish advice more often when they are better able to justify it to themselves? Do manipulations that prime justifications further enhance the context effect? Simonson's work (1989) implies that accountable decision makers (i.e., ones who expect having to justify their choices) display a stronger context effect (see also Simonson & Tversky, 1992). Relatedly, Lerner and Tetlock's review (1999) suggests that accountability could either amplify, attenuate or maintain cognitive biases. According to their theory, accountable individuals faced with a choice task tend to pick an alternative they could readily justify to an imagined audience. The literature on ethical decisions also points out that the tendency to

display unethical behavior is bolstered by the availability of reasons or excuses for such behavior (Aquino & Reed, 2002; Dana, Weber, & Kuang, 2007).

The goal of Study 3 was to investigate whether manipulating accountability – that is, telling advisors (vs not telling them) that they would be asked to explain their choices – can moderate the context effect. The design included, therefore, two factors, consideration set (basic vs context) and forewarning (present vs absent) about the need to give justification. Half the advisors were forewarned that, after choosing what recommendation to give to the recipient, they would be asked to explain this choice. The other half were not forewarned about the impending request to explain their choice. We expected this factor to moderate the context effect.

As in the previous studies, the advisors made consequential choices and received payments depending on their recommendations. The participants were undergraduates recruited on campus. To protect their anonymity, we have developed a protocol in which the experimenter remained blind to the participants' responses throughout the procedure (thus analysis was possible only at the group level).

Method

Participants. We recruited 120 participants at the campus library. They were promised a show-up fee of 5 shekels for their participation. (One shekel was worth $0.30 at the time of the study.)

Procedure and design. The participants were run individually in a private room in the library. They were seated at a desk and given an instruction booklet. The procedure was self-paced, with no involvement on the part of the experimenter. They were told that they would participate in one round of an interactive game.

The first page of the instructions described a game with two roles, advisor and recipient. Participants were told that they would be assigned a role. On the second page,

participants found that they were assigned the role of advisor. The next two pages, described either the two-card or the three-card game, depending on the condition (basic or context). The relevant cards for each game were stapled to the page, face down. The back of the cards were labeled A and B in the basic condition and A, B and C in context condition. Advisors were instructed to take a peek at the faces of the cards and find out the payoff values for the recipients, which were eight, five, and one shekel(s), for A, B and C, respectively. Then they were supposed to select one card to be sent to their assigned recipients. The advisors were told that the recipients would be shown a parallel set of cards, except that the card values would not appear on them, and that the only information available to them would be the advisor's recommendation.

The advisors were then informed that they would receive additional payment that would depend on the recommendation they delivered to the recipient. The payments for recommending cards A, B and C, were two, four, and four shekels, respectively. The additional payments were displayed on the page under the relevant cards, as in Study 1.

After reading the instructions, the participants were asked to indicate their recommendation, using a procedure that aimed to maintain their anonymity. Depending on the condition, either two or three envelopes were attached to an instruction sheet; they were marked A and B in the basic condition, and A, B and C in the context condition. The advisors were asked to choose a recommendation by opening the corresponding envelope. Each envelope contained a recommendation card and the advisor's payment – the show-up fee (five shekels) plus some payment that depended on the recommendation. Thus, envelopes A, B, and C contained seven, nine, and nine shekels, respectively. The advisors were instructed to take the money from the envelope they chose and place the recommendation card in another envelope addressed to the recipient. Finally, they were supposed to insert the envelope (through a narrow slit) into a card box labeled "recommendations." The contents of this box

were collected only at the end of the experiment. This procedure guaranteed anonymity such that choices could only be analyzed at the group level and participants could not be identified.

Two factors were manipulated orthogonally, between participants, choice set (basic vs context condition) and forewarning (present vs absent). In the forewarning-present condition, the participants read a sentence telling them that after making a recommendation they would be asked to explain their choice in writing. In the forewarning-absent condition, this sentence did not appear. There were 30 participants per condition.

Results

Table 4 displays the distribution of choices as a function of the consideration set and the forewarning instructions. Our goal was to test the context effect and whether forewarning moderated this effect. We compared the percentages of self-serving (B) recommendations in the various conditions. As before, the dependent variable was created using binary coding, such that a self-serving recommendation was coded 1 and optimal recommendation was coded 0. The data of the one participant who recommended the reference alternative were excluded from the analyses. A logistic regression was conducted using advisor recommendation as the dependent variable. The predictors were the consideration set, the forewarning factor, and the interaction between them (the two main predictors were centered prior to computing the interaction term).

As predicted, participants made more self-serving recommendations in the context than in the basic condition (62.7 vs 23.3%), odds ratio = 5.77, $z = 4.22$, $p < .001$. Forewarning did not have a significant effect (38.3 vs 47.5%), odds ratio = 1.65, $z = 1.20$, $p = .229$, nor did the two-way interaction, odds ratio = 0.59, $z = -0.63$, $p = .527$.[1]

[1] We obtained virtually the same results in further analyses using all data, in which we recoded the B recommendations as 1, and the remaining recommendations as 0.

Table 4

The Distribution (%) of Recommendations in Study 3

Forewarning absent

Condition	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	83.3	16.7	-
Context	40.0	60.0	0.0

Forewarning present

Condition	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	70.0	30.0	-
Context	33.3	63.3	3.3

Discussion

Study 3 provided further evidence for the effect of context. The presence of a reference recommendation (card C) increased the advisors' propensity to recommend to the recipient the self-serving option (B). The results replicated the findings of Studies 1 and 2, this time using a paper-and-pencil procedure. It is notable that the data patterns generalized to this setting, even though the advisors in this study were undergraduates who might have felt more affinity to the recipients who were their peers in the same college (compared with the M-Turk respondents).

Forewarning advisors prior to making a choice that they would need to provide written justifications of their choices did not affect their tendency to give selfish recommendations. Moreover, the forewarning factor did not moderate the context effect, in contrast with an

earlier study (Simonson, 1989) that found such moderation. In order to reach a more definitive conclusion, we conducted additional study.

Study 4

In this study, as in the previous one, we tested forewarning as a moderator of the context effect. We recruited a larger sample, using the Amazon M-Turk panel. The study involved as before, two between-participants factors, consideration set (basic vs context) and forewarning (present vs absent). Half the advisors were forewarned that, after giving their recommendation to the recipient, they would be asked to explain it, while the other half did not receive such forewarning. In fact, all participants, regardless of whether or not they were forewarned, received a prompt to write an explanation, after making their choice. This could be readily accomplished in this computer-controlled study. In Study 3, we did not elicit explanations in the forewarning-absent condition, since we were concerned that participants might inadvertently take note of the impending request for justification, simply by leafing through the pages of the booklet they were handed.

Study 4 allowed us to assess the effects of forewarning more broadly. As discussed previously, the moderation hypothesis implies that we should observe a larger context effect on choices when forewarning is present than when it is absent. Furthermore, we will assess the contents of the explanations, focusing on certain motives, such as self-interest and pro-sociality. Thus in addition to testing the hypothesized choice patterns, we could also assess if they are accompanied by a surge in the proportion of explanations relating to self-interest in the context condition, relative to the basic condition, and whether the surge in the frequency of such explanations is larger when forewarning is present than when it is absent.

Method

The study was preregistered at: https://doi.org/10.17605/OSF.IO/MX65G.

Participants. We recruited 598 participants via the Amazon Mechanical Turk website. We offered them a base fee of $0.30 plus additional payment, depending on their responses.

Procedure and design. The procedure was implemented on a computer, using the same general instructions as in Studies 1-2. Following the instructions phase, the participants (in the role of advisors) were presented with either two or three cards, depending on condition. They were asked to click on the cards in order to reveal their monetary values for the recipient. In the basic condition, the values of cards A and B were 20 and 6 cents, respectively. In the context condition, the values of cards A, B and C were 20, 6, and 1 cent(s), respectively. The participants were then informed that they could earn additional payment, depending on the recommendation they chose to transmit to the recipient. The payoffs were 5 and 10 cents for transmitting cards A and B (and 10 cents for C). The participants were then prompted to select their recommendation. Note that card A was the optimal for the recipient, while Card B was suboptimal for the recipient, but most profitable for the advisor. In terms of the payoffs to the advisor and recipient, card C was inferior to B. As before, the procedure involved one round of the game.

As in Study 3, forewarning was crossed with choice set, between-participants. In the forewarning-present condition (only), the participants were told, “After making your recommendation to the recipient, you will be asked to explain it. Your explanation will not be sent to the recipient. Only the researchers will read it.” In the final screen, participants in both conditions were told, “Please explain your considerations in choosing what recommendation to send. Your explanation will be read only by the researchers. It will not be sent to the recipient.” For their final payments, the participants received the base fee plus additional payment, based on their recommendations.

The procedure included screening questions. The instructions phase involved one captcha question and then one attention check, as in Study 2. After making their recommendation, the participants were presented with a memory question identical to the one

used in Study 2. Twenty participants (3.38% of the sample) who failed this probe were excluded, leaving 572 participants for the analyses. The *N*s for the four conditions ranged from 139 to 148.

Results

Table 5 displays the distribution of choices as a function of the experimental factors. The proportions of self-serving (B) recommendations were compared in the various conditions, with the goal of testing the effects of context and forewarning. The dependent variable was binary, such that the self-serving recommendation was coded 1 and the optimal recommendation (for the recipient) was coded 0. The data of the six (1%) participants who recommended the reference alternative were excluded from the analyses. A logistic regression was conducted using advisor recommendation as the dependent variable, with choice set, forewarning, and the interaction between them, as predictors. The two main predictors were centered prior to computing the interaction term.

As predicted, the participants made more self-serving recommendations in the context than in the basic condition (52.0 vs 39.6%), odds ratio = 1.65, $z = 2.96$, $p = .003$. Forewarning had no effect (46.6 vs 44.5%), odds ratio = 0.92, $z = -0.49$, $p = .621$, nor did it have an interactive effect with choice set, odds ratio = 1.29, $z = 0.74$, $p = .461$.

Table 5

The Distribution (%) of Recommendations in Study 4

Forewarning absent

Condition	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	57.9	42.1	-
Context	48.2	51.1	0.7

Forewarning present

Condition	Card A: Optimal for recipient	Card B: Self-serving for advisor	Card C: Reference alternative
Basic	62.8	37.2	-
Context	45.7	50.7	3.6

Content analysis. We analyzed the advisors' explanations with the goal of assessing the effects of forewarning. We coded the occurrence in the data of any of the following four social motives. The first was self-interest (e.g., "My main consideration was how much it would benefit me"). The second was regard for the other person (e.g., "I would like my partner to receive a bonus, even though I get a slightly lower bonus"). The third was fairness / equality (e.g., "Recommending B feels slightly fairer to both parties"). The fourth was ethical behavior (e.g., "I tried to be honest. I did not want to lie.")

Two independent judges (who were blind to the goals of the study) coded the presence (1) or absence (0) of each of these four motives in each explanation. Therefore, each explanation included, in principle, zero, one, two, three or four motives. The judges saw only the texts of the explanations, and received no other information. Their inter-judge agreement was 83%.

The frequencies of the various types of motives were computed by averaging the (0-1) codes (made by the two coders) for each explanation (see Table 6). Note that the numbers in each row do not add up to 100%, since a single explanation could include multiple motives.

We ran a linear regression analysis with the frequency of only one motive, self-serving motive, as a dependent variable. The factors of choice set, forewarning, and interaction were predictors. The results indicated that the self-serving motive was no more frequent in the context than in the basic condition, (40.8 vs 41.5%, respectively), b = -0.006, t = -0.17, p = .869, and no more frequent in the forewarning-present than in forewarning-absent condition (40.6 vs 41.7%, respectively), b = -0.01, t = -0.28, p = .782. There was no interaction effect, b = -0.04, t = -0.52, p = .605. Thus, we found no evidence for the hypothesized moderating effect of forewarning on the context effect.

Table 6

Study 4: Frequencies (%) of the Motives Mentioned in the Participants' Explanations

Forewarning absent

Condition	Self-interest	Other-regarding	Fairness	Ethics
Basic	41.0	36.9	11.4	8.6
Context	42.4	35.1	23.6	25.0

Forewarning present

Condition	Self-interest	Other regarding	Fairness	Ethics
Basic	41.9	33.4	19.6	16.6
Context	39.3	44.4	24.4	23.3

Our final analysis of the explanations, aimed to evaluate the complexity of the explanations as a function of choice set, forewarning and the interaction between them. We conducted a linear regression analysis with the *number* of motives mentioned in each

individual explanation as the dependent variable. We found that more motives were mentioned in the context than in the basic condition (1.29 vs 1.05, respectively), $b = 0.24$, $t = 5.32$, $p < .001$, and more motives were mentioned when forewarning was present than when it was absent (1.21 vs 1.12, respectively), $b = 0.10$, $t = 2.13$, $p = .034$. There was no interaction effect, $b = -0.08$, $t = -0.90$, $p = .368$. Thus, the addition of context drew more considerations and so did forewarning. The hypothesized moderating effect of forewarning on the context effect was not found.

Discussion

Consistent with the previous studies, Study 4 revealed an overall effect of context, whereby the addition of the third, reference recommendation (card C) increased the advisors' propensity to select the self-serving recommendation (B). The forewarning factor – that is, informing versus not informing participants that they would need to explain their choices – did not affect their tendency to give selfish recommendations. Moreover, this factor did not moderate the documented context effect.

We also analyzed participants' explanations of their choices, for the types of social motives mentioned in them. The context condition elicited richer explanations than the basic condition, consistent with the idea that the addition of a third alternative produced for the participants a new reference for evaluation of the options. Moreover, advisors offered richer explanations in terms of the social motives, when forewarning was present than when it was absent. This finding is consistent with Tetlock's (1983) theory that accountable individuals tend more to engage in integrative and complex thinking. That is, they tend to take into consideration multiple perspectives in support of their choices.

Our studies differ from Simonson's (1989) in that they did not reveal a moderation effect. Forewarning did not amplify the context effect. Several aspects of our studies may account for why forewarning did not yield the expected moderation of the context effect. Our

advisors made consequential choices, whereas Simonson's participants made hypothetical consumer choices, such as about nameless beers. Moreover, our participants engaged in choices that involved a tradeoff between self-interest and the welfare of another person. Such choices reflect on one's self-image (Aquino & Reed, 2002; Dunning, 2007; Sachdeva, Iliev, & Medin, 2009) and seem more value laden (and emotionally charged) than hypothetical consumer choices (Simonson, 1989).

Another possible account for the lack of moderation effect is that our forewarning manipulation produced opposite influences that cancel each other out. While we had initially expected accountability to highlight the relative merit of the dominating alternative (i.e., the self-serving card B), it might have instead reminded the participants of social norms, such as concern for others. We do not have direct evidence for this account, save that it is in accord with the finding that the forewarning manipulation did not moderate the context effect.

There is however evidence that our recommendation task was value laden, so much that the advisors felt morally accountable, regardless of the forewarning manipulation. The participants' written explanations showed that they negotiated (internally) various ethical concerns, pitting prosocial concerns (helping others) and fairness against the lure of self-interest. Forewarning therefore, might have not made a difference, simply because the participants were negotiated the merits of each option internally, irrespective of the instructions. The evidence here comes from the finding of similar frequencies of motives across the different forewarning conditions (present vs absent).

Study 5

The studies so far showed that advisors were prone to make selfish recommendations when a third (dominated) alternative was included in the choice set. Are participants aware of this effect? Is the effect of adding a third (dominated) alternative transparent to them? Awareness of this effect is of interest and possible importance, since advisors could

strategically exploit it, for example, by adding to the choice set a dominated alternative that would help justify giving a selfish recommendation.

In order to test awareness, we presented to all our participants two game settings, one involving two cards and another involving three cards. We then asked them to consider anonymous advisors and assess "how comfortable these persons would feel" making self-serving recommendations. If participants are aware of the effect of context on choice, then they should indicate that the advisor placed in the three-card setting would be the one more comfortable giving a selfish advice.

Method

Participants. We recruited 60 undergraduate participants (71.7% women; mean age: 24.0). They participated in the study in exchange for a voucher for coffee plus pastry at the campus cafeteria (value of 10 shekels).

Materials, procedure and design. Participants received a booklet with descriptions of two games entitled "the two-card communication game" and "the three-card communication game" (the basic and context conditions from Study 3, respectively). The participants who read the two-card game first were told that in a past study, 30 students were recruited to play the advisor role and they each made a choice between A and B. The participants of Study 5 were asked to guess the results of this past study, that is, to indicate which card (A or B) was preferred by the majority of the advisors. The same participants then read the instructions for the three-card game. They were told that another study had been conducted, with 30 new students in the role of advisors, who were supposed to choose one card of three. The participants of the current study (5) were asked again to guess which of the three cards (A, B, or C) was the most popular choice in this past study. The order of presentation of the two games was counterbalanced between participants.

In sum, the participants guessed the outcomes of two studies. This procedure encouraged careful consideration of the game scenarios and prepared the participants for the main question that provided the primary measure of the study: "Consider two advisors. One advisor played the two-card game and chose after some deliberation card B. The other played the three-card game and chose after some deliberation card B as well. In your opinion, who felt more comfortable recommending B, the one who played the two-card game or the one who played the three-card game?" The participants had to indicate one option.

Results

We started by analyzing the participants' predictions of the majority choice in each game. In the two-card game, 70.0% of the participants indicated the selfish recommendation (B) as the majority choice. In the three-card game, 86.7% of the participants indicated selfish recommendation (B) as the majority choice. This difference was significant, McNemar's $\chi^2 = 8.01$, $p = .004$. The pattern here corresponds to what we found in Study 3, in terms of direction, though not in terms of the exact percentages, suggesting that participants had a good grasp of the two games described to them, lending credence to the primary measurement of the study.

Almost all participants (95%) indicated that the advisor who played the three-card game must have felt more comfortable giving the selfish recommendation B. This proportion was significantly different from the value of 50%, $\chi^2 = 46.82$, $p < .001$.

Discussion

First, our participants made sensible predictions regarding the majority choice in the two- and three-card communication games described to them. Second, the vast majority of the participants (95%) indicated the advisor who played the three-card game must have been more comfortable making a selfish recommendation. This result implies that the nature of the mechanism of reference-dependent evaluation was transparent to them. Simply put, they

understood that the presence of the third (dominated) alternative changed the perception of the selfish alternative for the better.

An intriguing implication of findings such awareness is that conflicted advisors could use this insight strategically. Advisors who are tempted to give self-serving advice could introduce dominated alternatives into consideration, in order to justify giving self-serving advice. This way, they could obtain both the psychological benefit afforded by the context and the material gain associated with the choice to give self-serving advice. This intriguing possibility deserves separate investigation in the future.

Study 6

We have posited an account of the context effect, whereby advisors engage in a reference-dependent evaluation process of the alternatives. Conflicted advisors behave more selfishly, if an alternative recommendation is introduced, which is dominated by the self-serving recommendation, thereby making the self-serving one appear more favorable and hence more justifiable. Our hypothesized evaluation process though makes predictions with respect to the givers as well as the recipients of the advice.

Consider a scenario whereby the recipients of selfish advice are informed (after the fact) about the advisors' incentives. Given this information, how would recipients respond to the receipt of selfish advice, under the context and basic conditions? Our participants were placed in such settings and then asked to rate the advisor's behavior on several scales, indicating the extent to which they found it selfish versus other-regarding. Based on our hypothesis, we expect the recipients to evaluate the advisor's selfish choice as a function of the context. Specifically, if the recipients' perceptions are subject to the comparison-based evaluation mechanism, then they should rate selfish advice-giving less negatively when the (dominated) reference alternative is introduced to the set. In other words, the recipients should display the same bias as the advisors.

Method

Participants. Eighty participants (60% women, mean age 25.1) were recruited on campus, in exchange for a chance of 1 in 20 to win a lottery for 20 shekels.

Procedure and design. The participants were presented with a scenario describing an interaction between an advisor and a recipient. Specifically, they were instructed to imagine taking part in a game involving two roles, advisors and recipients, and further, that they had been assigned the recipient role. They were then asked to imagine that their task was to choose one envelope from a set of two, labeled A and B (or three, labeled A, B and C) and, then, open it and get the money inside that envelope. They were further asked to imagine that their sole information about the contents of the various envelopes was the recommendation of an advisor who was randomly paired with them, and that the recommendation was to open envelope B. Finally, they were asked to imagine they had followed the recommendation and thus opened envelope B, where they found five shekels, their award in the experiment. At this point, they were informed about the full structure of the game, namely, the values of the two (three) cards and the advisor's respective payoff for each of the cards.

Once the participants finished reading all the information, they were asked to rate the advisor's behavior. Specifically, they rated how strongly they felt the advisor's behavior was (a) disappointing, (b) fair, (c) egoistic, (d) moral, and (e) considerate. The ratings were made on five 7-point scales, anchored at 1 (not at all) and 7 (very much). The items were intended to tap the same underlying construct, namely, the extent to which the advisor's behavior was selfish. These ratings were the prime measurements of the study.

Participants were randomly assigned to one of the two versions of the advisor-recipient scenario. In the basic version, the participants were asked to imagine that they had to select one of two envelopes, A or B. In the context version, they were asked to imagine that they had to select one of three envelopes, A, B or C. The amounts inside the envelopes (as well as the advisors' payoffs) were the same as in Study 3.

Results and Discussion

The items were reverse-coded so that all five items were aligned, with higher values indicating a more negative attitude towards the advisor's selfish behavior. The judgments on all five scales were consistent (Cronbach's alpha = .75), suggesting that they tapped the same construct of "selfish behavior." A global attitude score was therefore computed for each participant by averaging the ratings for the five items.

We found that the participants' evaluations were less negative in the context than in the basic condition, 3.77 vs 4.76, $t(78) = 4.16$, $p < .001$, $d = 0.93$. Thus, the participants were overall more forgiving towards the advisor who had given them an inferior recommendation in the context than in the basic condition. In sum, the recipients' perceptions of the selfish advice exhibited the same pattern as the advisors' choices. This finding corroborates our comparison-based account of the context effect, since such ratings do not coincide at all with the recipients' self-interest, but rather evolve from their perceptions of the alternative recommendations that were available to the advisor.

Study 7

How might social observers react to selfish advice giving? Social observers are individuals who do not partake in the advisor-recipient interaction, but form impressions of the actors whose actions they happen to observe. Observer impressions are important in the social arena of advice exchange, since they affect the reputations of the advisor(s). We tested the effect of context on observers' judgments. Our study participants were asked to take the role of someone who had observed an advisor providing a recipient with an inferior advice. The observers were fully informed about the payoff matrix. We predicted, based on our hypothesized evaluation mechanism, that the observers should also display the context effect.

Method

Participants. Seventy-nine participants (58% women, mean age 24.8) were recruited on campus. They filled out a brief questionnaire in exchange for a lottery ticket (a chance of 1 in 20 to win 20 shekels).

Procedure and design. The participants were all assigned to the role of an observer. The participants were presented with the advice exchange scenario used in Study 3, in one of two versions, either the basic or the context condition. The participants were given all the information about the recipient's gains and advisor's recommendation-based payoffs (as in Study 3). The scenario further described a particular participant, advisor #114, who had recommended card B to a particular recipient. The participants were asked to rate how they felt towards the advisor's behavior, using the following adjectives (a) fair, (b) egoistic, (c) moral, and (d) considerate. They made their ratings with respect to each adjective, on four separate scales that ranged from 1 (not at all) to 6 (very much).

Results and Discussion

The ratings of the four items were reverse-coded, so that they were all aligned, with higher values indicating attitudes that were more negative towards the advisor's behavior. The judgments on all scales were consistent (Cronbach's alpha = .79). A global score was obtained for each participant by averaging the four scale ratings. We found that the participants felt significantly less negatively about the advisor's behavior in the context than in the basic condition, 2.98 vs 4.02, $t(77) = 4.13$, $p < .001$, $d = 0.93$. We conclude that the social observers in this study displayed the context effect, just like the recipients in the previous study.

General Discussion

Our research goal was to investigate advisors' selfish behavior in conflict-of-interest settings. In the basic condition of Study 1, conflicted advisors provided mostly prosocial

recommendations (70%), suggesting they suspended their self-interest in order to help the recipients. Their tendency to behave selfishly increased dramatically however, when a third, inferior alternative was added to the consideration set (context condition). While this third option was barely chosen, it boosted the advisors' tendency to choose the selfish option.

Why did the third alternative boost selfish advice-giving? In theory, the addition of a dominated alternative should not matter since it is an inferior one in the set. Advisors however used the third alternative as a reference, against which the (dominating) selfish alternative was evaluated more favorably. This theoretical account is in accord with general principles of perception (Roberts et al., 2005), whereby the evaluation of a target stimulus depends on other stimuli that co-occur with it in space and time (Cogan, Parker, & Zellner, 2013). Study 2 tested the characteristics of the third alternative that deemed necessary for the context effect to occur. Importantly, we found that the context effect (i.e., an increase in selfish advice giving) was obtained only when the third alternative was strictly inferior to the self-serving one.

Studies 3-4 tested accountability as a moderator of the context effect (Lerner & Tetlock, 1999), in attempt to elucidate the mechanism underlying the context effect in advice-giving. Based on previous findings (Simonson, 1989), we hypothesized that forewarning advisors that they would be asked to explain their choices should enhance the context effect. We expected forewarning to increase the availability of reasons for choosing the selfish option and thereby, foster its choice (cf. reason-based choice, Simonson & Tversky, 1992). The studies yielded no evidence for this moderation effect. One possible explanation is that forewarning produced conflicting influences that canceled each other out. It is possible that it highlighted the reasons for selfish behavior, as we had expected (Simonson, 1989), and at the same time, also elevated the advisors' moral sense and inhibited their selfish tendencies.

Importantly, in their written explanations the participants raised ethical issues and fairness considerations, weighing their prosocial concern against their self-interest. This

implies that the advising task was value laden and perhaps even emotionally charged, to the extent that the advisors felt morally accountable, regardless of the forewarning manipulation.

Study 5 tested the advisors' awareness of the context effect. The participants were asked to judge which of two advisors should feel more at ease giving selfish advice, one facing a two-card game (basic condition) or one facing a three-card game (context condition). The participants indicated the latter for the most part, suggesting that they had insight into the operation of the context effect. Given this insight, advisors could exploit the context effect to their strategic advantage. We suggest that self-interested advisors could intentionally introduce an inferior alternative to their consideration set in order to increase the favorability of a selfish recommendation. Indeed a recent health policy debate on the promotion of electronic cigarettes made this point apparent. In this debate (Kessler, 2019, July 31), physicians and scientists, who served as industry consultants, promoted e-cigarettes suggesting they were less harmful than the ordinary cigarettes. These paid consultants stated their professional preference for a potentially harmful product (but lucrative to the industry), by suggesting that it compared well with an inferior alternative.

Studies 6-7 suggest that introducing an inferior alternative affects not only the perceptions of advisors, but also the perceptions of others, such as the recipients of advice (Study 6) and outside observers (Study 7). Consistent with the reference-dependent evaluation process, we see that advice recipients (who could suffer losses from using selfish advice) and social observers (who have no stakes in the advice) are just as biased as the advisors (who actually gain from selfish advice-giving). That the hypothesized mechanism accounts for the perceptions of advisors and other social actors lends strong credence, in our view, to this parsimonious theory.

Relations to Other Work

Ethical decision-making. Selfish advice giving under conflict of interest is often considered dishonest and even illegal in some professional settings. There is plenty of evidence that people care about their public as well as private self-image and as result of that, tend to refrain from unethical behavior (Hilbig & Hessler, 2013; Mazar, Amir, & Ariely, 2008; Sachdeva et al., 2009). Individuals allow themselves nonetheless to behave less ethically when they believe there exists a "moral wiggle room" that allows them to maintain their positive self-image nonetheless (Aquino & Reed, 2002; Dana, Weber, & Kuang, 2007; Shalvi, Dana, Handgraaf & De Dreu, 2011). Our work implies that the third alternative creates such a "moral wiggle room" that would justify giving selfish advice.

Is it their public or private image that our advisors cared to maintain? Individuals typically care about their public image (Dunning, 2007). The context of the third alternative could, in principle, provide the excuse advisors needed (vis-à-vis the recipients or outside observers) for recommending the advice that was suboptimal for the recipients. We suggest though that the advisors in our studies (who avoided giving the suboptimal advice), mostly cared to maintain their private image, since in their view, the recipients were not supposed to be exposed to the counterfactual alternatives. To the extent that the advisors avoided giving suboptimal advice (i.e., suboptimal for the recipients), it was because it was unacceptable to them internally.

Asymmetric dominance effect. Since its early demonstration (Huber et al., 1982), the asymmetric dominance effect (akin to the context effect studied here) has been documented in many studies of individual decision-making (Simonson, 2014). Recent research implies that the effect is less robust than it has been widely believed to be. A major criticism has been that the asymmetric dominance effect holds empirically when choice options are described numerically, but not when they are depicted visually or verbally (Frederick, Lee & Baskin, 2014; Yang & Lynn, 2014). In answering this criticism, Simonson (2014) and Huber et al.

(2014) point out that it is necessary that decision makers identify the dominance configuration in the choice set, quickly and unambiguously, in order for the asymmetric dominance effect to obtain. In other words, participants who fail to perceive an asymmetric dominance relationship among the options are less likely to exhibit the effect.

In our studies, the choice attributes – the recipient's payoff and the advisor's incentives – were numerical. We designed our procedure to make the dominance configuration clear. Participants were gradually introduced to the recipient's gains and the advisor's incentives, so that the conflict of interest and the asymmetric dominance relationship was transparent (cf. Doyle, O'Connor, Reynolds & Bottomley, 1999). Indeed, few if any of our participants chose the dominated alternative, supporting our conclusion that the dominance configuration was clear (cf. Simonson, 2014).

Final Comments

Our studies portray what we believe to be a potentially important cause of selfish advising behavior under conflict of interest. The introduction of a dominated reference alternative into the advisor's consideration set tends to bias the advisor's perception of the selfish option so it is viewed more favorably. We have suggested that a reference-dependent evaluation mechanism can account for this bias. Our studies show that not only advisors, but outside social observer as well succumb to this bias. Further, recipients of advice appear more forgiving towards selfish advisors' behavior in the very same settings where they are more likely to suffer from such behavior. These findings highlight certain elusive aspects of selfish behavior, which may sustain, and perhaps even reinforce, the occurrence of self-serving advising under conflict of interest.

References

Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423–1440.

Cain, D. M., Loewenstein, G., & Moore, D. A. (2005). The dirt on coming clean: Perverse effects of disclosing conflicts of interest. *Journal of Legal Studies, 34*(1), 1–25.

Cogan, E., Parker, S., & Zellner, D. A. (2013). Beauty beyond compare: Effects of context extremity and categorization on hedonic contrast. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 16.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd edition). New York, NY: Russell Sage Foundation.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67-80.

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

Dunning, D. (2007). Self-image motives and consumer behavior: How sacrosanct self-beliefs sway preferences in the marketplace. *Journal of Consumer Psychology, 17*(4), 237–249.

Frederick, S., Lee, L., & Baskin, E. (2014). The limits of attraction. *Journal of Marketing Research, 51*(4), 487–507.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review, 95*(1), 384-394.

Grice, H.P. (1975). Logic and conversation. In P. Cole & J.L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.

Highhouse, S. (1996). Context-dependent selection: The effects of decoy and phantom job candidates. *Organizational Behavior and Human Decision Processes, 65*(1), 68–76.

Hilbig, B. E., & Hessler, C. M. (2013). What lies beneath: How the distance between truth and lie drives dishonesty. *Journal of Experimental Social Psychology, 49*(2), 263–266.

Huber, J., Payne, J. W., & Puto, C. (1982). Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research, 9*(1), 90-98.

Huber, J., Payne, J. W., & Puto, C. P. (2014). Let's be honest about the attraction effect. *Journal of Marketing Research, 51*(4), 520–525.

List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy, 115*(3), 482–493.

Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research, 45*(6), 633–644.

Moore, D. A., & Loewenstein, G. (2004). Self-interest, automaticity, and the psychology of conflict of interest. *Social Justice Research, 17*(2), 189–202.

Moore, D. A., Tanlu, L., & Bazerman, M. H. (2010). Conflict of interest and the intrusion of bias. *Judgment and Decision Making, 5*(1), 37–53.

Parker, S., Bascom, J., Rabinovitz, B., & Zellner, D. (2008). Positive and negative hedonic contrast with musical stimuli. *Psychology of Aesthetics, Creativity, and the Arts, 2*(3), 171-174.

Kessler, D. A. (2019, July 31). Juul Says It Doesn't Target Kids. But Its E-Cigarettes Pull Them In. The New York Times. Retrieved from https://www.nytimes.com/2019/07/31/opinion/juul-kids.html

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125*(2), 255-275.

Roberts, B., Harris, M. G., & Yates, T. A. (2005). The roles of inducer size and distance in the Ebbinghaus illusion (Titchener circles). *Perception, 34*(7), 847-856.

Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science, 20*(4), 523-528.

Sah, S., & Loewenstein, G. (2012). More affected = more neglected: Amplification of bias in advice to the unidentified and many. *Social Psychological and Personality Science, 3*(3), 365–372.

Shalvi, S., Dana, J., Handgraaf, M. J., & De Dreu, C. K. (2011). Justified ethicality: Observing desired counterfactuals modifies ethical perceptions and behavior. *Organizational Behavior and Human Decision Processes, 115*(2), 181–190.

Simonson, I. (1989). Choice based on reasons: The case of attraction and compromise effects. *Journal of Consumer Research, 16*(2), 158–174.

Simonson, I. (2014). Vices and virtues of misguided replications: The case of asymmetric dominance. *Journal of Marketing Research, 51*(4), 514–519.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*(3), 281–295.

Tetlock, P. E. (1983). Accountability and complexity of thought. *Journal of Personality and Social Psychology, 45*(1), 74–83. doi.org/10.1037/0022-3514.45.1.74

Yang, S., & Lynn, M. (2014). More evidence challenging the robustness and usefulness of the attraction effect. *Journal of Marketing Research, 51*(4), 508–513.

Yaniv, I. (2004). The benefit of additional opinions. *Current Direction in Psychological Science, 13*(2), 75-78.

Yaniv, I., & Choshen-Hillel, S. (2012). When guessing what another person would say is better than giving your own opinion: Using perspective-taking to improve advice-taking. *Journal of Experimental Social Psychology, 48*(5), 1022-1028.

H E B R E W U N I V E R S I T Y O F J E R U S A L E M

מרכז פדרמן לחקר הרציונליות

THE FEDERMANN CENTER FOR THE STUDY OF RATIONALITY

List of Recent Discussion Papers

These and earlier papers can be found on our website: ratio.huji.ac.il

715. **Bezalel Peleg and Ron Holzman,** Representations of Political Power Structures by Strategically Stable Game Forms: A Survey (February 2018)

716. **Kristoffer Arnsfelt Hansen, Rasmus Ibsen-Jensen, and Abraham Neyman,** The Big Match with a Clock and a Bit of Memory (February 2018)

717. **Bezalel Peleg and Hans Peters,** Self-implementation of social choice correspondences in strong equilibrium (April 2018)

718. **Abraham Neyman,** Additive valuations of streams of payoffs that obey the time-value of money principle: characterization and robust optimization (April 2018)

719. **Bezalel Peleg and Shmuel Zamir,** Judgements aggregation by a sequential majority procedure (June 2018)

720. **Lotem Elber-Dorozko and Yonatan Loewenstein,** Striatial Action-Value Neurons Reconsidered (July 2018)

721. **Orli Bobek, Adiv Gal, David Saltz, and Uzi Motro,** Effect of Nest-Site Microclimatic Conditions on Nesting Success in the Lesser Kestrel (Falco Naumanni) (July 2018)

722. **Barry O'Neill,** Two-Party Agreements as Circular Sets (September 2018)

723. **Adiv Gal, David Saltz, and Uzi Motro,** Effect of Supplemental Feeding on Nesting Success in the Lesser Kestrel (Falco Naumanni) (September 2018)

724. **Maya Bar-Hillel,** The unbearable lightness of self-induced mind corruption (November 2018)

725. **Tomer Siedner,** Optimal pricing by a risk-averse seller (May 2019)

726. **Maya Bar-Hillel and Cass R. Sunstein,** Baffling bathrooms: On navigability and choice architecture (June 2019)

727. **Maya Bar-Hillel and Jacob Lavee,** Lay attitudes toward involuntary organ procurement from death-row prisoners: no, but (June 2019)

728. **Maya Bar-Hillel,** Why didn't I see it earlier? (July 2019)

729. **Maya Bar-Hillel, Tom Noah, and Shane Frederick,** Solving stumpers, CRT and CRAT: Are the abilities related? (October 2019)

730. **Sergiu Hart and Yosef Rinott,** Posterior probabilities: dominance and optimism (November 2019)

731. **Sergiu Hart and Dean P. Forester,** Forecast-hedging and calibration (November 2019)

732. **Maya Bar-Hillel,** The base-rate fallacy in probability judgments (December 2019)

733. **Yigal Attali and Maya Bar-Hillel,** The false allure of fast lures (February 2020)

734. **Uri Zak,** Female Chess Players Do Underperform When Playing Against Men: Commentary on Stafford (2018) (March 2020)

735. **Daniel Kahneman and Maya Bar-Hillel,** Laplace and Cognitive Illusions (June 2020)

736. **Sergiu Hart and Yosef Rinott,** Posterior Probabilities: Nonmonotonicity, Log-Concavity, and Turán's Inequality (July 2020)

737. **Maya Bar-Hillel,** An annotated compendium of stumpers (July 2020).

738. **Alex Gershkov and Eyal Winter,** Exploitative Priority Service (August 2020).

739. **Alexander Kravtsov and Eyal Winter,** An Axiomatic Approach to Sensors' Trust Measurements (August 2020).

740. **Robert J. Aumann,** Why Consciousness? (May 2021)

741. **Elon Kohlberg and Abraham Neyman,** Demystifying the Math of the Coronavirus (May 2021)

742. **Constantine Sorokin and Eyal Winter,** Pure Information Design in Classic Auctions (July 2021)

743. **Dean P. Foster and Sergiu Hart,** "Calibeating: Beating Forecasters at Their Own Game (October 2021)

744. **Sergiu Hart,** Calibrated forecasts: The Minimax Proof (November 2021)

745. **Todd R. Kaplan and Shmuel Zamir,** On the Strategic Use of Seller Information in Private-Value First-Place Auctions (December 2021)

746. **Ran Ben Moshe,** Comparing Mechanisms for Selling Correlated Goods (July 2022)

747. **Ran Ben Moshe, Sergiu Hart, and Noam Nisan,** Monotonic Mechanisms for Selling Multiple Goods (November 2022)

748. **Maya Bar-Hillel,** על פול גרייס ואריה דרעי (March 2023)

749. **Mair Barneron, Shoham Choshen-Hillel, and Ilan Yaniv,** Reaping a Benefit at the Expense of Multiple Others: How are the Losses of Others Counted? (April 2023)

750. **Mair Barneron and Ilan Yaniv,** Advice-Giving under Conflict of Interest: Context Enhances Self-Serving Behavior (April 2023)