Article

Dual-Process Reasoning in Charitable Giving: Learning from Non-Results

Zachary Grossman [1,*,†] [iD] **and Joël J. Van der Weele** [2,†]

1   Florida State University, Department of Economics, Bellamy Building, Tallahassee, FL 32306-2180, USA
2   University of Amsterdam, Center for Research in Experimental Economics and Political Decision Making (CREED), Tinbergen Institute; PO Box 1551, Amsterdam 1001 NB, The Netherlands; vdweele@uva.nl
*   Correspondence: zgrossman@fsu.edu
†   These authors contributed equally to this work.

Abstract: To identify dual-process reasoning in giving, we exposed experimental participants making a charitable donation to vivid images of the charity's beneficiaries in order to stimulate affect. We hypothesized that the effect of an affective manipulation on giving would be larger when we simultaneously put the subjects under cognitive load using a numerical recall task. Independent treatment checks reveal opposite responses in men and women and cast some doubt on the reliability of our mainstream treatment manipulations and assessment tools. We find no evidence for dual-process decision-making, even among women, whose responses to the manipulations conformed most to our expectations. These results highlight the need for caution in the use of these common manipulations, the importance of independent manipulation checks, and the limitations of dual-process models for understanding altruistic behavior.

1. Introduction

Dual-process theories assert that decision-making is comprised of two systems: a slow, deliberative cognitive system and a fast, intuitive/affective system. These theories have been applied fruitfully to economic domains such as decision-making under risk or intertemporal choice [1,2]. More recent research has examined whether dual-process reasoning also underlies altruistic or cooperative behavior and whether people are instinctually selfish or altruistic.

Identifying dual-process reasoning requires an indirect approach as the processes under consideration are hard to observe directly.[1] One method is to place the decision maker under cognitive load: if the cognitive system is occupied or taxed then behavior more strongly reflects the influence of the affective system.[2] A problem with this approach is that if both systems approximately "agree" on the right amount of giving, manipulations of the cognitive system will not result in different decisions. This may well be the case in the giving situations that are the focus of the literature, like the dictator game, and can explain why a growing number of studies implementing cognitive load yield inconclusive results. For example, Schulz et al. [11] find that cognitive load makes dictators somewhat

---

1   Neuroscientific evidence suggests that brain areas associated with cognitive and affective processing are active during the process of charitable giving [3–5].
2   Another approach looks at response times. Piovesan and Wengström [6] finds that subjects who decide faster are more selfish in a modified dictator game. In contrast, Rand et al. [7] and Rand et al. [8] find that faster subjects are more cooperative across a range of games. The usefulness of response time as a window into cognitive processes is not universally accepted [9,10].

more generous and Hauge et al. [12] do not find any difference. In a charitable giving experiment, Kessler and Meier [13] find a small increase in giving that is not robust to small manipulations. Furthermore, Benjamin et al. [14] find a small negative effect and Cornelissen et al. [15] find that the effect of cognitive load depends on social value orientation. Thus, the current literature provides little consensus on the brain processes that determine giving.

To overcome this problem, we conducted a charitable-giving experiment with manipulations of both the affective and the cognitive system in a 2 × 2 design. Based on the model in Loewenstein et al. [1], we hypothesized that the affective manipulation would lead to stronger increases in dictator giving when the cognitive system is under load. Thus, the change in giving across cognitive load levels is not our main outcome of interest. Rather, we look at how varying cognitive load affects the between-subjects *change* in giving driven by the affect manipulation—the difference in differences. This design avoids the problem mentioned above.[3]

We chose manipulations that have been used with apparent success in many previous studies. To manipulate the cognitive system, we asked participants to remember either a long or short string of numbers. Variations of these numeric recall tasks have been used in numerous studies published in both psychology and economics journals, including the studies on dictator giving cited above and some classic studies that are heavily cited.[4] The affective system manipulation consisted of vivid images of victims helped by the charity under consideration, a standard method used to induce affect (e.g., Olofsson et al. [19]), and featured individual refugees, playing on the identifiable-victim effect [16,20].[5] Some of the images were taken from the charity's own website, reflecting real-world communication strategies.

We performed independent behavioral checks of both manipulations to study whether they succeeded in manipulating cognitive and affective processes.[6] The results show that the manipulations, despite being standard, did not work entirely as we expected. We do not find statistically-significant effects of our treatments on aggregate giving levels. In light of previously documented gender differences in response to intuitive and deliberative manipulations [21], we examine men and women separately and find that they have opposite responses to both manipulations. Re-evaluating our results for women, whose response to the treatment hewed most closely to our expectations, we still find no support for dual-process decision-making. These results raise important concerns for the study of dual-process theories and altruism, which we discuss in the final section.

2. Design

We conducted the experiments at the Frankfurt Laboratory for EXperimental economics (FLEX) at Goethe University Frankfurt with subjects randomly recruited from the FLEX subject pool using the online system ORSEE [23]. The treatments were programmed with z-Tree [24] and participants received a show-up fee of €2. The Supplementary Materials provide screenshots containing the instructions. We conducted 12 sessions, each lasting about 25 min and with 12 to 23 participants each, for a total of 224 participants.[7]

---

[3] Small et al. [16] also investigate both affective and cognitive processes using a 2 × 2 design. However, that study is not designed to separate the two systems, and some of the treatment manipulations may affect both the cognitive and the affective system simultaneously. Skitka et al. [17] show that the effect of cognitive load on social attitudes depends upon a subject's political leanings. More liberal subjects expressed more sympathy than conservatives towards a personally-responsible aid claimant under low cognitive load, but high load significantly reduced this correlation. This study was hypothetical and not incentivized. Furthermore, the affective manipulation was not induced in a randomized way but based on differences between subgroups.

[4] For example, Shiv and Fedorikhin [18] has over 1500 citations indexed by Google Scholar.

[5] Small and Loewenstein [20] and Small et al. [16] each have over 500 citations indexed by Google Scholar.

[6] Skitka et al. [17] and Hauge et al. [12] assess the success of their respective manipulations, but use introspective self-reports, whereas ours are based on behavioral and out-of-sample measures.

[7] All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University of California, Santa Barbara (2-13-0749).

Participants had the opportunity to donate to the German Red Cross (GRC) programs that provide aid to victims of conflict in Syria. After reading basic instructions on an introductory screen, participants advanced to a screen displaying information about the GRC programs in Syria. The text was adapted from the GRC website and described the victims of the conflict, the activities of the GRC, and how the activities help the victims. When reading this information, the participants did not yet know that they would be asked to make a donation to the GRC.

A third screen introduced the cognitive manipulation, which consisted of a numeric recall task. Participants were given a number and told that on two subsequent screens they would be asked to add a single digit to the number and keep a running total in their head. This total was to be reported later in the session. Participants were not allowed to write or use electronic devices, so this running total had to be calculated and stored in participants' memories.

The starting number was 13 in the *Low Load* treatment and 13,987 in the *High Load* treatment. The two numbers to be added were 7 and then 8, yielding interim and final totals of 20 and 28 in the *Low Load* treatment and 13,994 and 14,002 in the *High Load* treatment. This constitutes a small variation of the standard numeric recall task, in which a high load typically requires memorizing a single six or seven digit number. The purpose of requiring participants to perform recurrent addition tasks was to keep them engaged with the memory task throughout both the donation choice and the subsequent stage. Because this increased the level of difficulty of the task, we used a slightly shorter number to avoid discouraging participants. Like Kessler and Meier [13], we did not provide financial incentives for this recall task to avoid creating a disparity in income effects across experimental treatments.

We then gave participants an endowment of €10 and asked them to choose how much of it to donate to the GRC. This donation choice was our main outcome of interest and was made by selecting an amount from a menu of choices, enumerated in whole euro amounts, from €0 to €10. Above the donation-choice menu, participants were asked to add 7 to the number they were shown on the previous screen and remember the running total, thus extending the cognitive load manipulation.

Next, participants took a version of the cognitive reflection test (CRT) [25] and were also asked to add 8 to the running total in their head. This was our first manipulation check, as the CRT is designed to assess participants' ability to overcome an intuitive incorrect answer in favor of a correct answer that requires more reflection. Our version of the CRT consisted of four questions.[8] Lower CRT performance in the *High Load* treatment would be consistent with an effective cognitive manipulation [26]. After the subjects completed the CRT, they advanced to a screen where they were asked to input the memorized number. Participants then completed a brief questionnaire, the purpose of which was to collect basic demographic information and to return them to a more neutral cognitive state.

As a further manipulation check, after the load had been lifted, we gave subjects the unexpected opportunity to revise their donation decisions. Under low cognitive load, we would expect little change in their behavior, especially because phenomena such as anchoring and a desire for consistency would reinforce the previous donation choice. However, dual-process theories predict that when there is conflict between systems, different behavior may manifest depending upon whether or not the person is under cognitive load. In the *High Load* condition, the initial donation decision is made without much influence from the cognitive system. After completing the recall task and the questionnaire, these subjects may have returned to a more neutral, unloaded cognitive state, and thus their decision may be subject to more influence from the cognitive system. Although in the introduction we argued that cognitive load might not lead to different decisions, larger and more frequent revisions in the *High Load* treatment would be *consistent* with an effective cognitive manipulation. We reiterate that the donation revision was *not* included so as to test the hypotheses of dual-process models. It served only to assess the effectiveness of our manipulation of the cognitive system.

8　　The CRT questions we used are available in the Supplementary Materials.

The affective manipulation consisted of displaying images chosen to increase empathy for the conflict victims and willingness to donate to a charity that helps them. In the *High Affect* treatment, three screens (the GRC-information screen, the donation screen, and the revised donation screen) featured evocative images of victims and aid workers in areas affected by conflict in Syria. The images were chosen to elicit affective responses without providing additional information about the conflict victims or the charity. Three of those images had been taken from the GRC website, so this manipulation reflects the actual communication strategy of the GRC.[9] The same four images were shown on each screen. In the *Low Affect* treatment, no images were displayed on any screens.

3. Hypotheses

We base our hypotheses on the model in Loewenstein et al. [1], which gives a rather literal interpretation of the dual-system approach. The authors argue that optimal level of giving may differ between the affective and cognitive systems, in which case the resulting donation level will lie somewhere in between. The degree to which the cognitive system will be able to influence the decision away from the 'affective optimum' towards the 'cognitive optimum' depends on contextual factors, such as the degree of depletion of the cognitive system and the presence of other tasks that require cognitive attention.

This experiment departs from previous studies by featuring not only a cognitive manipulation, but also a manipulation designed to increase the level of giving desired by the affective system. The use of pictures to evoke affect is standard in social psychology [19] and is designed to increase pity and compassion. Moreover, focusing on individual victims can increase giving through 'the identifiable victim effect', which is typically associated with the affective system (see [16,27]). We assume that being exposed to images of victims will increase the amount that the affect system regards as appropriate to give. We assume that the cognitive system's appropriate amount is less than that of the stimulated affect system, so that the unloaded cognitive system constrains the giving off the image-exposed affective system, but the loaded cognitive system does not have as strong an effect.[10]

We thus hypothesize that *High Affect* will increase the amount of giving favored by the affective system, leading to higher donations. Moreover, under *High Load*, the cognitive system will have less capacity available to influence the decision away from the 'affective optimum' than under *Low Load*. Thus, we hypothesize that the increase in giving will be larger under *High Load* than under *Low Load*. In other words, under *High Load*, we expect the 'whims' of the affective systems to have a more powerful influence on behavior than under *Low Load*. If the cognitive and affective manipulations work as intended, then the failure to find a larger increase in giving under *High Load* would be inconsistent with dual-process models.

While we believe Loewenstein et al. [1] is a plausible model, there are many different dual-process theories, and not all of them would yield these results. For instance, based on findings underlining the primacy of the affective system [28,29], one could conjecture that the impact of the affective manipulation should be independent of the cognitive-load manipulation. Disentangling different dual-process models is an important task for empirical research.

4. Results

We first examine the effectiveness of the cognitive and affective manipulation and then turn to an analysis of the charitable giving choices.

---

[9] The fourth image was slightly more graphic and featured an urban area after a bombardment, with residents gathering around destroyed properties. It was added to increase the affective shock from the images.

[10] We acknowledge the possibility that the cognitive system's appropriate amount is higher than the amount desired by the affective system in either treatment. This could potentially explain our negative results, but it would not be consistent with the revision behavior that we observe.

4.1. Effectiveness of the Cognitive Manipulation

We assess the effectiveness of the cognitive manipulations by looking at four different measures: how correctly subjects reported the number they were asked to compute and recall in the cognitive load task, how quickly they made their decisions in each stage, their performance on the CRT, and the frequency and amount of their revisions of their donations.

4.1.1. Number Reporting

In the *Low Load* treatment, 98 out of 112 participants (88%) correctly reported the number they were asked to compute and 74 out of 112 (66%) did in the *High Load* treatment. Thus, it seems that the *High Load* task was indeed harder than the *Low Load* task. Relative performance notwithstanding, the fact that the majority of *High Load* participants reported the number correctly indicates that they were taking the task seriously.[11]

4.1.2. Decision Time

Our analysis of the amount of time participants took to complete each screen reveals that participants took significantly longer to make their decisions under *High Load*, with the notable exception of the CRT screen. Table 1 reports the mean decision time in seconds, by cognitive-load treatment, for each screen encountered by participants, as well as *p*-values for an MWU test of equal distributions across the two conditions. It shows no significant difference in decision times for the Introductory and Charity Information screens, before the cognitive-load manipulation had been introduced, as well as no significant difference for the Questionnaire and Donation Revision screens, which came only after the cognitive-load task was complete. However, with the exception of the CRT screen, the other three screens that participants encountered during the cognitive-load task—the screen initiating the Cognitive Loading, the Donation Decision screen, and the Number Recall Screen—all featured decision times from participants under *High Load* that were markedly and statistically-significantly higher. It is *only* for the CRT screen that this is not the case.

Table 1. Comparison of mean decision times for each screen across cognitive-load treatments with *p*-values for an MWU test of equal distributions. Also included for comparison are *p*-values for the comparison of the low and high affect treatment.

Screen	Mean Decision Time (s)		MWU Test *p*-Value	
	Low Load	*High Load*	Cognitive	Affective
Intro	22.6	23.2	0.66	0.39
Charity Info	50.8	55.3	0.22	0.12
Cognitive Loading	33.6	50.1	0.00	0.12
Donation Decision	37.3	45.3	0.00	0.00
CRT	94.4	94.8	0.58	0.58
Number Recall	13.4	27.1	0.00	0.63
Questionnaire	32.2	32.4	0.99	0.49
Revise Donation	35.8	38.8	0.87	0.17

Furthermore, as a placebo test, we also examined the difference in response times across affective treatments and find that, with the exception of the Donation Decision screen, participants' response times are not different across affective treatments. It is also worth noting that women's decision time
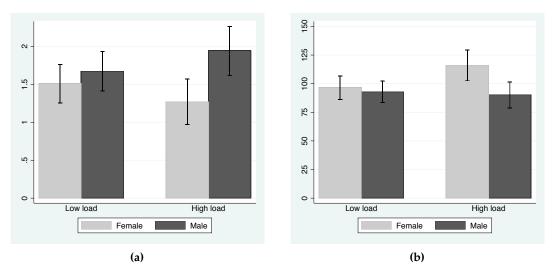
---

[11] Eighteen *High Load* participants reported an incorrect number, but were close to either the final number or the intermediate number, or omitted one of the digits of the correct number. These participants tried to complete the task and were likely still under cognitive load while choosing the donation amount. Pooling these with those who reported correctly suggests that at least 81% of the participants spent cognitive resources on the task in the *High Load* condition.

is more impacted by *High Load*. For example, while decision time on the Donation Decision screen jumps from 38.4 to 42.2 with the introduction of *High Load* for men, for women the same manipulation causes a jump from 36.0 to 48.7, an increase of roughly more than 33 percent. Both of these jumps are statistically significant, but for women it is much larger.

4.1.3. CRT Performance

The average number of CRT questions answered correctly was 1.60 out of 4 in *both* load treatments, so on the basis of performance we cannot reject the possibility that the cognitive manipulation was ineffective. Note that this goes against the evidence in Johnson et al. [26], who find that load, imposed using a similar memory task, does decrease CRT performance. *Low Load* participants took on average 94 s to answer the questions, while they took on average 103 s in the *High Load* treatment. This difference of 9 seconds (10 percent) is not significant at the 10-percent level ($p = 0.115$, *t*-test).[12] The absence of any significant differences in performance and completion time casts doubt on the effectiveness of the manipulation.

There is evidence that women perform differently on the CRT test [25], so we disaggregate CRT performance and speed by gender. This reveals a divergent response to cognitive load. Figure 1 shows the mean CRT score and response time across genders and load treatments. Women in the *High Load* treatment answer on average 1.27 out of four CRT questions correctly, which is lower than the 1.51 average score for women in the *Low Load* treatment, although not significantly so ($p = 0.170$). Similarly, women's average response time increases from 96.5 s to 115.9 s from the *Low Load* to *High Load* treatments, which is a significant difference ($p = 0.020$, *t*-test). In contrast, men's average performance *increases* from 1.67 to 1.94 from *Low* to *High Load* treatments, although the change is not statistically significant ($p = 0.203$), and their average response time drops trivially from 92.8 to 90.1 ($p = 0.364$, *t*-test). As Figure 1 indicates, these opposite tendencies by men and women result in a large gender gap under high load, both with respect to performance ($p = 0.002$) and decision time ($p = 0.004$, *t*-test).[13]
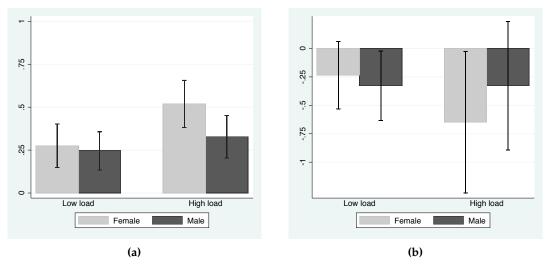

Figure 1. Mean CRT results by cognitive load and gender with 95% confidence intervals. (**a**) CRT score; (**b**) CRT answer time in seconds.

---

12   Unless noted otherwise, reported *p*-values are for Mann–Whitney U (MWU) tests, for which the null hypothesis is equal distributions. One can make interpretations about differences in means if one assumes that the other parameters of the distribution are the same. Where appropriate, we also performed *t*-tests of differences of means, which did not change the conclusions.

13   We find a gender difference in CRT performance overall. The average number of correct answers is significantly lower ($p = 0.01$, *t*-test) for women (1.41) than for men (1.77), echoing the results in Frederick [25].

4.2. Donation Revisions

In addition, we judge the cognitive manipulation by evaluating how subjects revised their decisions after cognitive load was lifted.[14] The average participant revised her donation downward by €0.38 and 34 percent of participants made a non-zero revision. Downward revisions were larger on average in the *High Load* treatment (€0.48) than in the *Low Load* treatment (€0.29), but not significantly so ($p = 0.404$, *t*-test).[15] While 26 percent in the *Low Load* treatment chose to revise the original donation amount, this increased significantly to 42 percent in the *High Load* treatment ($p = 0.016$, Fisher Exact Test (FET)). These latter results constitute the only statistically-significant evidence that is consistent with the expected effect of the cognitive-load manipulation.

Figure 2a shows the fraction of subjects who revised their decision by gender and by cognitive load. Revision rates are higher in the *High Load* treatment for both sexes, consistent with the idea that reflection is impaired in this treatment. However, while this increase is significant for women ($p = 0.016$, FET), it is not for men ($p = 0.417$, FET). Figure 2b shows the average size of revisions by gender and cognitive load. Men do not respond to higher load with a higher average revision size. Women do but the effect is not statistically significant.

Figure 2. Revised donations by cognitive load and gender with 95% confidence intervals. (**a**) Revision rate; (**b**) Mean revision size in euros.

Result 1. *We find evidence for the effectiveness of the cognitive load treatment on some measures but not others. Men and women displayed different effects on decision time, CRT performance and revision rates, with women's responses much more consistent with the expected effects.*

4.3. Effectiveness of the Affect Manipulation

We now investigate the effect of the affective manipulation. A first noteworthy observation comes from Table 1. In the final column of the Table, we see that there is a highly significant difference in response times on the donation page, which is caused by an average increase in response time of 4 s in

---

14   In the *High Affect* treatment, the images shown before and with initial donation choice were also provided in the revision screen, so that the affect treatment was maintained.

15   Our hypothesis test could reject the null hypothesis of no difference between mean revision amounts across cognitive-load treatments at the five-percent level with an effect size of 0.46 (11 percent of the average initial donation amount), meaning that if each person in the *High Load* condition had revised their donation downward by 0.26 more Euros than they actually did (6 percent of the average initial donation amount), then we could reject at the 5 percent level.

the high affect treatment compared to the low affect treatment ($p = 0.027$, *t*-test). This indicates that subjects did pay attention to the pictures when making their donation decision.

To understand whether the affective manipulation actually increased the level of giving preferred by the affective system, we elicited affect ratings from participants in other, independent sessions.[16] These participants only reported their feelings and were not asked to make any donation. They were presented with the same information about the GRC as in the main experiment and shown the same vivid images of the conflict in Syria in the *High Affect* treatment, but not in the *Low Affect* treatment. Next, they completed a six-item Likert scale pertaining to their feelings about victims of the conflict in Syria and the GRC programs assisting those victims. The items included declarations of feelings of pity, sympathy, compassion, and an obligation to help the victims of conflict in Syria, as well as the desire to do more to help the victims, and a statement of approval of the GRC's programs in Syria.[17] The scale featured six levels of agreement or disagreement and did not include a neutral option. Finally, participants completed the same questionnaire as in the donation experiment. Ninety participants in ten sessions completed the affect ratings, of whom 44 percent identified as male.

For each item, we mapped responses to numerical scores from zero to five, with zero corresponding to strong disagreement and five corresponding to strong agreement. For each individual, we summed these scores over the six items, yielding a total affect score with a range of 0 to 30. The overall average affect score was 14.4 with a standard deviation of 5.6. Table 2 breaks down the affect score by gender and affect treatment. We cannot reject the hypothesis that the affect score is distributed the same across both affect treatments ($p = 0.643$), nor can we reject the hypothesis that it is the same across gender ($p = 0.705$).

Table 2. Affect score by treatment and gender.

	Low Affect	*High Affect*	**Total**
Female	13.7	15.4	14.5
N	27	23	50
Male	14.5	13.8	14.2
N	23	17	40
Total	14.1	14.8	14.4
N	50	40	90

Looking at treatment effects by gender, the average score for women increases from 13.7 to 15.3 when shown the images of conflict victims ($p = 0.402$). Men respond to the images with a drop in average score from 14.5 to 13.8 ($p = 0.752$). However, since the aggregate score combines different answer distributions, it has high variance and the test has relatively low power. We therefore look at individual Likert items, which suggests different affective responses across gender. In each affect dimension, average women's ratings go up with the display of pictures, except for the "help" measure, for which there is no change. These changes are at or close to marginal significance levels for feelings of pity and compassion ($p = 0.103$ and $p = 0.068$, respectively). This contrasts with the men, whose average rating *decreases* under *High Affect* in four out of six dimensions, with statistical significance when it comes to approval of the GRC's activities ($p = 0.077$). These opposite reactions lead women to have slightly higher scores for all items under *High Affect*, with virtual ties for "help" and "sympathy" and marginally significant differences for compassion ($p = 0.069$) and approval ($p = 0.056$).[18]

16　These subjects had participated in another, unrelated experiment and were asked before they were paid to complete one final, unrelated and unpaid task.

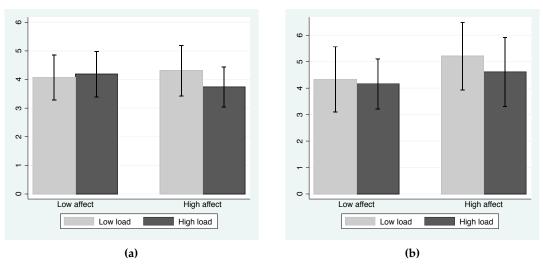17　The full text of the Likert items is available in the Supplementary Materials.

18　While our data do not provide any evidence as to why men and women exhibit opposite reactions to the images, one possibility is that images of foreign victims may trigger group identity. Men are known to act less positively towards out-group members than women [30].

As further test, we examine the three items—pity, sympathy, and compassion—that most directly measure affect. The Cronbach's alpha for these three items is 0.76, which indicates an acceptable level of reliability. Aggregating the scores of these three items and testing the hypothesis of equal means across affect levels yields a *p*-value of 0.13 (*t*-test, one-sided), falling short of significance in a manner similar to that of the broader analysis. However, the difference among the women-only sample is borderline significant ($p = 0.10$, *t*-test, one-sided).

Result 2. *A display of vivid pictures has a weak positive effect on women's affect across the board, with more pronounced increases in pity and compassion. For men, the opposite is true, as affect ratings go down in most dimensions.*

4.4. Analysis of Donation Choices

Figure 3a shows the mean initial donation levels across treatments. The overall mean donation was €4.07, so the average subject took home €7.93, including the €2 show-up fee. The 113 *High Affect* participants gave 4.02 on average, which is not higher than the 4.13 average for the 111 *Low Affect* participants. Under *Low Load*, the average donation did increase slightly from 4.07 ($N = 56$) in the *Low Affect* treatment to 4.30 ($N = 56$) in the *High Affect* treatment, but it actually decreased from 4.18 ($N = 55$) to 3.74 ($N = 57$) in the *High Load* treatment. This is inconsistent with our hypothesis that there would be a larger jump in giving under *High Load*. A series of two-sided Mann–Whitney U (MWU) tests do not reject the null hypothesis that the donation distributions are the same when comparing any two treatments.[19]

(a)　　　　　　　　　　　　　　　　　　　　　　(b)

Figure 3. Mean donations by treatment with 95% confidence intervals. (**a**) Mean donations by men and women; (**b**) Mean donations by women only.

Result 3. *There are no significant differences in aggregate donation levels across treatments.*

Having found above that our manipulations affect women in a manner closer to that which was intended than they do for men, we re-examine women's donation decisions. Overall, women contribute on average 88 cents more than men (4.54 vs. 3.66) and the distribution of their donations is significantly

---

[19]　A multivariate OLS analysis also fails to find statistically significant effects of our treatments or their interaction. We also do not find a direct relation between CRT scores and giving, as some other authors have reported [31]. The only significant result, which comes at the 10% level, occurs for the interaction between gender and affect, with males responding to *High Affect* by reducing their donation by 1.34.

different ($p = 0.012$) from that of men's donations.[20] Figure 3b shows mean giving levels for women, by treatment. Under *Low Load*, the mean initial donation increased from 4.33 to 5.21 in response to the affect manipulation, a difference of 0.88 ($p = 0.289$). Under *High Load*, the mean donation increased from 4.16 to 4.61, a difference of only 0.45 ($p = 0.734$). Thus, if anything, higher cognitive load causes a *smaller* increase in giving and our test provides no support for the dual-process model.

Result 4. *Women increase giving in response to our affect manipulation, but cognitive load does not lead to a larger increase.*

4.5. Power

For our design, it was important to generate a sizeable effect of the affective manipulation, so we could study how it varied in different cognitive load treatments. We computed the affective treatment effect that we could have detected at the 5% level with 80% power on a two-sample test of means. For the overall sample with a size of 112 in each of the affect treatments and using the fact that the mean donation in the *Low Affect* treatment is 4.13, we could have rejected the null hypothesis of no treatment effect for an effect size of 0.35, implying a mean donation in the *High Affect* treatment of 4.76. Below, we separately examine the subsample of women. Assuming a sample size of 52 in each treatment and using a mean donation of 4.24, we could have rejected the null hypothesis of no treatment effect in this subsample for an effect size of 0.55, implying a mean donation in the *High Affect* treatment of 4.79.

5. Discussion and Conclusions

To identify dual-process reasoning in altruistic decisions, we attempted to manipulate the cognitive and affective system in a giving task, using commonly-used procedures. Our treatment checks show that these manipulations did not work entirely as we expected and we find no effect aggregate giving levels. We find that the cognitive load manipulation affected outcomes for females, but not males, and that a display of vivid pictures raises women's affective ratings but lowers men's. Behavior among women, the group in which the manipulations had the expected effects, was not consistent with the dual-process model.

These observations contain several lessons for the nascent field of dual-process research in economics. First, the effects of even standard treatment manipulations may depend on unanticipated details. In the case of cognitive load, we are strengthened in this belief by a number of recent studies that find conflicting or null results from this manipulation (see the introduction). For example, Kessler and Meier [13] find that the effect of cognitive load manipulations depends on whether the manipulation was implemented early or late in the experiment.[21]

Second, manipulations may have different effects on different subgroups. This finding is consistent with that of Rand et al. [21], who find that promoting intuition relative to deliberation increases giving among women, but not men, and argue that whether one is intuitively altruistic or selfish may depend on who you are. The response to the treatments might depend on other characteristics such as education level or age, of which there was little variation in our study. Together, these two points underline the importance of a critical evaluation of these manipulations and the use of independent treatment checks.

Third, our results among the female sample are not consistent with the idea that cognitive capacity constrains the "whims" of the affective system. If anything, we observed a lower impact of affective

---

[20] Appendix A provides more details on giving differences between men and women.
[21] Using a similar manipulation, Balafoutas et al. [22] find that ego depletion moderates social preferences, which may explain why suppressing the cognitive system produces such inconsistent effects on giving behavior.

manipulations under cognitive load. While more research is obviously needed, this finding does not agree with most standard interpretations of the dual-system approach.

Thus, our results show the challenges involved in testing the dual-process model. As psychologists and neuroscientists have long realized, this model is a metaphor for a much more complex interaction of different distributed modules in the brain. Despite the popularity of this metaphor, there is an ongoing debate about its cogency, its usefulness for scientific practice, and the robustness of the supporting evidence (e.g., [32–34]). There is increasing evidence that the relation between behavior and affective and cognitive manipulations is very complex.[22]

Our results attest to the fact that the reasoning underlying giving decisions is probably much more complicated than can be usefully described by a simple dual-process model. For example, much more research is needed for us to understand the mechanisms that lead to the gender differences we observe and other concepts outside of the dual-process model, such as routines or social-impact theory, that might be useful for understanding giving behavior. However, any alternative theories would face the burden of having to make sense of the findings that inspired dual-process theories in the first place. Our paper provides a cautionary note for those seeking to do research in this area, by showing that treatments based on highly simplified models of the brain may fail to generate consistent results.

Author Contributions: "Zachary Grossman and Joel J. van der Weele conceived and designed the experiments; Joel J. van der Weele performed the experiments with research assistance from Ana Andrijevik and Paul Gortner; Zachary Grossman and Joel J. van der Weele analyzed the data; Zachary Grossman and Joel J. van der Weele wrote the paper".

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A. Detailed Donation Results

Tables A1 and A2 provide mean donation levels by treatment for women and men, respectively. On closer inspection, it appears that the gender difference is due to opposite reactions to the pictures of victims displayed in the *High Affect* treatment. While women increase their donations, men decrease them. This is in line with their reaction to our treatment manipulations in our independent check, inspiring confidence that the treatment did indeed have the hypothesized effect. Pooling both cognitive load treatments, the mean giving numbers when affect is high are 4.91 for women and 3.38 for men. A MWU test rejects the null hypothesis of equal distributions between men and women under high affect at the 1% level ($p = 0.0048$).

---

[22] For example, Neys and Schaeken [35] find that people are more logical under cognitive load, Blanchette et al. [36] show that increasing affect may improve reasoning when the emotions are relevant to the decision in question, and Lench and Bench [37] provide evidence that affective reactions can reduce judgment biases.

Table A1. Mean donations made by women.

	Low Affect	*High Affect*	**Total**
Low Load	4.33	5.21	4.75
N	27	24	51
High Load	4.16	4.61	4.35
N	31	23	54
Total	4.24	4.91	4.54
N	58	47	105

Table A2. Mean donations made by men.

	Low Affect	*High Affect*	**Total**
Low Load	3.83	3.63	3.72
N	29	32	61
High Load	4.21	3.15	3.59
N	24	34	58
Total	4.00	3.38	3.66
N	53	66	119

References

1. Loewenstein, G.; O'Donoghue, T.; Bhatia, S. Modeling the interplay between affect and deliberation. *Decision* **2015**, *2*, 55–81.
2. Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.
3. Moll, J.; Krueger, F.; Zahn, R.; Pardini, M.; de Oliveira-Souza, R.; Grafman, J. Human fronto-mesolimbic networks guide decisions about charitable donation. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15623–15628.
4. Sanfey, A.; Rilling, J.; Aronson, J. The neural basis of economic decision-making in the ultimatum game. *Science* **2003**, *300*, 1755–1758.
5. Knoch, D.; Pascual-Leone, A.; Meyer, K.; Treyer, V.; Fehr, E. Diminishing reciprocal fairness by disrupting the right prefrontal cortex. *Science* **2006**, *314*, 829–832.
6. Piovesan, M.; Wengström, E. Fast or fair? A study of response times. *Econ. Lett.* **2009**, *105*, 193–196.
7. Rand, D.G.; Greene, J.D.; Nowak, M.A. Spontaneous giving and calculated greed. *Nature* **2012**, *489*, 427–430.
8. Rand, D.G; Peysakhovich, A.; Kraft-Todd, G.T.; Newman, G.E.; Wurzbacher, O.; Nowak, M.A.; Greene, J.D. Social heuristics shape intuitive cooperation. *Nat. Commun.* **2014**, *5*, 3677.
9. Tinghög, G.; Andersson, D.; Bonn, C.; Böttiger, H.; Josephson, C.; Lundgren, G.; Västfjäll, D.; Kirchler, M.; Johannesson, M. Intuition and cooperation reconsidered. *Nature* **2013**, *498*, E1–E2.
10. Recalde, M.; Riedl, A.; Vesterlund, L. *Error Prone Inference from Response Time: The Case of Intuitive Generosity*; CESIfo Working Paper 4987; CESIfo: Munich, Germany, 2014.
11. Schulz, J.J.F.; Fischbacher, U.; Thöni, C.; Utikal, V. Affect and fairness: Dictator games under cognitive load. *J. Econ. Psychol.* **2014**, *41*, 77–87.
12. Hauge, K.E.; Brekke, K.A.; Johansson, L.O.; Johansson-Stenman, O.; Svedsäter, H. Keeping others in our mind or in our heart? Distribution games under cognitive load. *Exp. Econ.* **2016**, *19*, 562–576.
13. Kessler, J.B.; Meier, S. Learning from (Failed) Replications: Cognitive Load Manipulations and Charitable Giving. *J. Econ. Behav. Organ.* **2014**, *102*, 10–13.
14. Benjamin, D.J.; Brown, S.A.; Shapiro, J.M. Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences. *J. Eur. Econ. Assoc.* **2013**, *11*, 1231–1255.
15. Cornelissen, G.; Dewitte, S.; Warlop, L. Are social value orientations expressed automatically? Decision making in the dictator game. *Pers. Soc. Psychol. Bull.* **2011**, *37*, 1080–1090.
16. Small, D.A.; Loewenstein, G.; Slovic, P. Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organ. Behav. Human Decis. Process.* **2007**, *102*, 143–153.
17. Skitka, L.J.; Mullen, E.; Griffin, T.; Hutchinson, S.; Chamberlin, B. Dispositions, scripts, or motivated correction? Understanding ideological differences in explanations for social problems. *J. Pers. Soc. Psychol.* **2002**, *83*, 470–487.

18. Shiv, B.; Fedorikhin, A. Heart and mindin Conflict: The interplay of Affect Decision and Cognition in Consumer Decision Making. *J. Consum. Res.* **1999**, *26*, 278–292.

19. Olofsson, J.K.; Nordin, S.; Sequeira, H.; Polich, J. Affective picture processing: an integrative review of ERP findings. *Biol. Psychol.* **2008**, *77*, 247–265.

20. Small, D.A.; Loewenstein, G. Helping a victim or helping the victim: Altrusim and identifiability. *J. Risk Uncertain.* **2003**, *26*, 5–16, arXiv:astro-ph/0005074v1.

21. Rand, D.G.; Brescoll, V.L.; Everett, J.A.C.; Capraro, V.; Barcelo, H. Social heuristics and social roles: Intuition favors altruism for women but not for men. *J. Exp. Psychol. Gen.* **2016**, *145*, 389–396.

22. Balafoutas, L.; Kerschbamer, R.; Oexl, R. *Distributional Preference Types and the Deliberative System*; Mimeo; Innsbrück University: Innsbrück, Austria, 2015.

23. Greiner, B. An Online Recruitment System for Economic Experiments. *Forsch. Wiss. Rechn.* **2003**, *63*, 79–93.

24. Fischbacher, U. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Exp. Econ.* **2007**, *10*, 171–178.

25. Frederick, S. Cognitive Reflection and Decision Making. *J. Econ. Perspect.* **2005**, *19*, 25–42.

26. Johnson, E.; Tubau, E.; Neys, W.D. The unbearable burden of executive load on cognitive reflection: A validation of dual process theory. *Proc. Annu. Conf. Cogn. Sci. Soc.* **2014**, *36*, 2441–2446.

27. Kogut, T.; Ritov, I. The "identified victim" effect: An identified group, or just a single individual? *J. Behav. Decis. Mak.* **2005**, *18*, 157–167.

28. Chang, H.H.; Tuan Pham, M. Affect as a Decision-Making System of the Present. *J. Consum. Res.* **2013**, *40*, 42–63, arXiv:1011.1669v3.

29. Shiv, B.; Loewenstein, G.; Bechara, A.; Damasio, H.; Damasio, A.R. Investment behavior and the negative side of emotion. *Psychol. Sci.* **2005**, *16*, 1–6.

30. Winterich, K.P.; Mittal, V.; Ross, W.T. Donation Behavior toward In-Groups and Out-Groups: The Role of Gender and Moral Identity. *J. Consum. Res.* **2009**, *36*, 199–214.

31. Ben-Ner, A.; Kong, F.; Putterman, L. Share and share alike? Gender-pairing, personality, and cognitive ability as determinants of giving. *J. Econ. Psychol.* **2004**, *25*, 581–589.

32. Keren, G.; Schul, Y. Two Is Not Always Better Than One: A Critical Evaluation of Two-System Theories. *Perspect. Psychol. Sci.* **2009**, *4*, 533–550.

33. Evans, J.S.B.T.; Stanovich, K.E. Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspect. Psychol. Sci.* **2013**, *8*, 223–241.

34. Carter, E.C.; McCullough, M.E. Publication bias and the limited strength model of self-control: Has the evidence for ego depletion been overestimated? *Front. Psychol.* **2014**, *5*, 1–11.

35. Neys, W.D.; Schaeken, W. When people are more logical under cognitive load. *Exp. Psychol.* **2007**, *54*, 128–133.

36. Blanchette, I.; Gavigan, S.; Johnston, K. Does emotion help or hinder reasoning? The moderating role of relevance. *J. Exp. Psychol. Gen.* **2014**, *143*, 1049–1064.

37. Lench, H.C.; Bench, S.W. Strength of affective reaction as a signal to think carefully. *Cogn. Emot.* **2014**, 37–41.