Joint Discussion Paper
Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 06-2016

Matthias Greiff, Kurt A. Ackermann and Ryan O. Murphy

The influences of social context on the measurement of distributional preferences

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

The influences of social context on the measurement of distributional preferences*

Matthias Greiff[†], Kurt A. Ackermann[‡], Ryan O. Murphy[§]

Tuesday 12th January, 2016

A variety of different social contexts have been used when measuring distributional preferences. This could be problematic as different degrees of social interdependence may affect people's distributional preferences, and this contextual variance may inadvertently muddle the measurement process and complicate comparisons between different strands of research. In the present study, we systematically measure distributional preferences embedded within different social contexts, as well as cooperative choices in a strategic setting. More specifically, we use a within-subjects design and compare the choices people make in resource allocation tasks with role certainty, role uncertainty, decomposed games, and matrix games. Results show that, at the aggregate level, role uncertainty and decomposed games both lead to higher degrees of prosociality when compared to role certainty. At the individual level, we observe considerable differences in behavior across the social contexts, indicating that the majority of people are sensitive to these different social settings but respond in different ways. We conclude with some recommendations for measuring distributional preferences as an individual difference and reiterate that social context is an inherent part of measurement methodology when considering social motivations.

Keywords: Distributional Preferences, Social Preferences, Other Regarding Preferences, Social Value Orientation (SVO), Measurement Methods, Individual Differences
JEL: C91, D03, D64

*We thank Max Albert, Andreas Hildenbrand, Hannes Rusch, and Rudolf Kerschbamer for helpful comments. We also thank Felix Krieg and Simone Schweizer for help with conducting the experiment. Any errors herein are our own.

[†]Department of Economics, VWL VI, Justus-Liebig-University Giessen, Licher Str. 66, 35394 Giessen, Germany, matthias.greiff@wirtschaft.uni-giessen.de

[‡]Fehr Advice & Partners, Klausstrasse 20, 8008 Zürich, Switzerland, kurt.alexander.ackermann@gmail.com

[§]University of Zürich, Department of Economics, Blümlisalpstrasse 10, 8006 Zürich, Switzerland, ryan.murphy@econ.uzh.ch

1 Introduction: Measuring distributional preferences

The measurement of preferences over payoff allocations, or distributional preferences, is an important issue in economics and psychology (see e.g., Andreoni and Miller, 2002; Camerer and Fehr, 2004; Balliet et al., 2009; Murphy and Ackermann, 2014; Kerschbamer, 2015a). A widely used technique for evaluating distributional preferences is a resource allocation task, in which a decision maker (DM) determines a payoff for herself and some unknown other person. Typically these choices are one-shot, mutually anonymous, and incentive compatible.

A method for measuring distributional preferences has two distinct components. First, there are the *choice sets*: well-defined sets of options that correspond to different allocations of resources. These are the feasible sets from among which the DM selects her most preferred allocations. There are several established sets of resource allocation tasks, which we discuss in section 2.1.

The second component is the *social context* in which a DM makes her choices. Typically, social features of the decision task are minimized, but there remains at least some minimal social context in which the choices are embedded, as there must be some "other" who is affected by the choices of the DM. Social context differs by role assignment and can be accented by the display of information, as we will discuss in section 2.2.

Differences in social context are potentially problematic, as they may interact with other social preferences apart from distributional preferences, for instance, preferences for reciprocity (Kerschbamer, 2015b). We believe that a systematic empirical evaluation of the role of social context in measuring distributional preferences, as we undertake in the present paper, would be a valuable addition to the literature. Moreover, a systematic comparison of social contexts tells us something about how humans perceive situations in the laboratory, which should be of interest to experimental economists.

We use a within-subjects design to compare the average degree of prosociality and the consistency of subjects' revealed preferences across different social contexts. Specifically, we use the same measurement stimuli (the choice sets that define the SVO Slider Measure) but systematically vary the social context. Our experimental design allows us to test several hypotheses, and to attempt the replication of some previous results.

We proceed as follows: In section 2, we give a systematic review of the different procedures that have been used to measure distributional preferences, focusing on *choice sets* and *social contexts*. In section 3, we briefly sketch the experimental design and use this sketch as a basis for outlining the main research questions and deriving our hypotheses. We then provide details of our experimental procedure (section 4). In section 5, we present and discuss our results and offer some suggestions for measuring distributional preferences before we conclude in section 6.

2 Procedural differences

2.1 Different sets of stimuli and the slider measure

The first component for the measurement of distributional preferences is the resource allocation task itself. Arguably the simplest task is a dictator game. Other experimental measurements of distributional preferences are based on multiple resource allocation tasks that have a more nuanced structure (i.e., various marginal rates of substitution). In all these allocation tasks the DM chooses her most preferred options, and, based on her choices, the DM's distributional preferences are revealed.

We can distinguish between two general approaches to measuring distributional preferences. The first approach categorizes DMs according to different types (selfish, maximin, welfare maximizing, inequality averse, and many more). The second approach measures the intensity of distributional preferences. Examples for the former are modified dictator games (Andreoni and Miller, 2002), the Ring Measure (Liebrand, 1984), the Triple-Dominance Measure (van Lange et al., 1997), or other resource allocation tasks designed to discriminate between different types (e.g., Engelmann and Strobel, 2004; Fehr et al., 2008; Iriberri and Rey-Biel, 2011; Blanco et al., 2011). Examples for the latter are the SVO Slider Measure (Murphy et al., 2011) and the Equality Equivalence test (Kerschbamer, 2015a). Kerschbamer's test can be used to differentiate nine different types and yields a two-dimensional index of preference intensity. The primary items from the

Slider Measure produce a one-dimensional measure called social value orientation (SVO).[1]

In this study, we use the SVO Slider Measure because the discretization of the data undermines statistical power and harms the reliability of results. The SVO Slider Measure has demonstrable psychometric benefits over alternative measures, including its production of a continuous score and a built-in transitivity check. SVO, as measured by the Slider Measure, has been shown to be highly stable as an individual difference as evidenced by a test-retest reliability of $r = 0.915$ (Murphy et al., 2011). The Slider Measure can also be readily extended into a matrix game. It is for these reasons we use it here as the stimuli for measuring distributional preferences.

The version of the Slider Measure used in this study works as follows: There are six resource allocation tasks (see Figure 1). For each task, five different allocations are presented to the DM, who is asked to choose her most preferred allocation. Based on her choices, the DM's SVO is computed as

$$SVO^{\circ} = \text{arctan}\left(\frac{\bar{\pi}_o - 50}{\bar{\pi}_s - 50}\right), \quad (1)$$

where $\bar{\pi}_s$ and $\bar{\pi}_o$ denote the DM's and the other's mean payoff. The SVO angle can range between -16.26° and 61.39°, while negative angular degrees indicate negative concerns for others, angular degrees close to zero indicate narrow self-interest, and positive angular degrees indicate positive concerns for others.[2]

2.2 Different social contexts

The simplest way to conceptualize a resource allocation tasks is to consider a decision situation with two distinct roles, an active *Decider* and a passive *Receiver*. For each resource allocation task,

[1]For a review of the different methods for measuring social preferences see Murphy and Ackermann (2014) and Kerschbamer (2013).

[2]The trigonometric scaling of the measure harkens back to early work from psychology on SVO (Griesinger and Livingston, 1973) that used a geometric model of preferences and *degrees* are commonly used when reporting SVO scores that reflect individual differences in social preferences. A person's SVO angle can be translated (by taking its tangent) into an approximation of that person's parameter α in a joint utility function $u(\pi_s, \pi_o) = \pi_s + \alpha \cdot \pi_o$, where π_s is the DM's payoff and π_o is the other individual's payoff (for further details, see Murphy et al., 2011; Murphy and Ackermann, 2014, 2015).

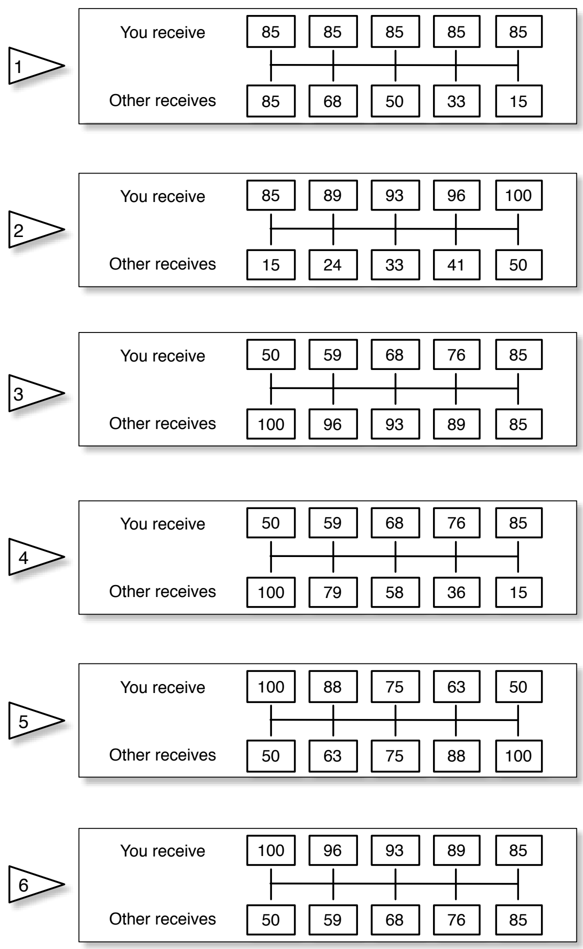

Figure 1: The six resource allocation tasks from the SVO Slider Measure. In treatments RCT, RUT and DGT, these six tasks were presented and subjects were asked to choose their most preferred allocations. In treatment MGT the same payoffs were used to create proper matrix games.

the *Decider* chooses an allocation which determines her own payoff and the *Receiver's* payoff. This is not a proper game as only the *Decider* makes a choice. However, it is a social decision as the *Decider*'s choice has an effect on some other subject. There has been substantial contextual heterogeneity in how resource allocation tasks have been implemented. This heterogeneity stems from differences in social context, more precisely, from differences in role assignment. Next we elaborate on the different role assignment procedures and provide examples from the literature that use these particular procedures.

2.2.1 Role certainty

In experimental designs with *fixed role assignment* or *role certainty*, subject A is in the role of the *Decider* and subject B is in the role of the *Receiver*. B is passive and has no choice to make; A knows this and all of these features are common knowledge. This is arguably the cleanest method when measuring distributional preferences (Kerschbamer, 2015b). In experimental practice, this means that the sample is split in half and the role of the *Decider* is assigned to subjects in one half of the sample, while the role of the *Receiver* is assigned to subjects in the other half. Role certainty has been employed in several experiments (e.g., Forsythe et al., 1994; Dana et al., 2007; Bardsley, 2007).

A downside of this procedure is inefficiency in that the distributional preferences of only half of the subjects are assessed. One other consequence is that this context creates a substantial imbalance in power that is obvious to *Decider* and *Receiver*. It is conceivable that this imbalance may evoke particular feelings of responsibility or entitlement on behalf of some *Deciders*, whereas others may not be affected by this imbalance. This imbalance in power is not a problem *per se* but it is worth noting and recognizing that even in the most austere and barren of experimental contexts, there exist social elements.

2.2.2 Role uncertainty

A procedure that avoids the inefficiency of only measuring half the sample is *role uncertainty*. All subjects in the sample perform the resource allocation task but do not know *a priori* whether they are going to be assigned the role of the *Decider* or the role of the *Receiver*. Only after the decisions are made by all subjects, they are matched randomly in pairs and another random process determines which role is assigned to which subject in each pair. Role uncertainty has been employed in several experiments (e.g., Charness and Grosskopf, 2001; Engelmann and Strobel, 2004).

Role uncertainty also mitigates the issue of power imbalance and preserves symmetry. With role uncertainty, all subjects are equally entitled because all subjects perform the same task (Engelmann and Strobel, 2004, 859). The social context corresponds to a choice of an allocation which might affect another subject's payoff; while the other subject is an anonymous subject who faces the same decision situation, and the other's choices might affect the DM's payoff. However mutual interdependence is not possible in this treatment as only one subject can be the *Decider* and the other must be the *Receiver*.

2.2.3 Double role assignment

The third procedure is called *double role assignment*. All subjects simultaneously perform the resource allocation tasks. A subject's total payoff comes from two sources, the payoff from her role as *Decider*, and her payoff from the role as *Receiver*. The mapping from choices to payoffs is common knowledge, and, because the two subjects are mutually interdependent, this is by definition a strategic setting (i.e., a proper game).

We can distinguish between double role assignment with fixed and random pairs. With fixed pairs, A is matched with B, and A's payoff depends on her own and B's choices. With random pairs, A, in the role of the *Decider*, chooses an allocation which affects her own and B's payoff, while in the role of the *Receiver*, A's payoff depends on the allocation chosen not by B but by another subject. Double role assignment with random pairs is used in many economic experiments (e.g., Andreoni and Miller, 2002; Fisman et al., 2007; Balafoutas et al., 2012, 2014). Double role assignment with fixed pairs has been employed in several experiments, mostly in psychology where it is most frequently used in the measurement of SVO (e.g., Pruitt, 1967; Messick and McClintock, 1968; Liebrand, 1984; Offerman et al., 1996; van Lange et al., 2007).

With random pairs, direct reciprocity is not an issue, but decisions might be affected by indirect reciprocity. The crucial point is that with both fixed and random pairs, each subject simultaneously gives and receives, so that a subject's decision to give might depend on her expectations about what she receives. Strategic choice behavior is determined by preferences and expectations, which implies that measuring distributional preferences with double role assignment may yield a confound as home grown expectations may inform choices as well as preferences. Along these lines Rigdon and Levine (2009) find significant differences in choice behavior contingent upon subjects' expectations. Since our experiment does not rely on double role assignment with random pairs, we refrain from further discussion.

With double role assignment with fixed pairs, there are two ways of displaying information about choices and payoffs. If information is displayed in a *payoff matrix*, strategic interdependence is salient because a subject can easily see that there is another subject whose decision affects the own payoff. Note that matrix games have not been used to measure social preferences because of obvious strategic interdependence. One attempt to control for the strategic interdependence is the use of the so-called *decomposed games method* (Pruitt, 1967; Messick and McClintock, 1968). Decomposed games attempt to accomplish this by manipulating how information about the payoffs is displayed to subjects. Assume that A and B simultaneously choose between allocations $X = (100, 50)$ and $Y = (85, 85)$. For each allocation, the first number denotes the payoff to the subject who chooses the allocation and the second number denotes the other subject's payoff. With the *decomposed games method* each subject sees both allocations (her possible choices) and is informed that the other subject chooses between the exact same allocations. The social context corresponds to a choice of an allocation, which affects another subject's payoff; and the other subject is an anonymous subject who faces the same decision situation, and the other's choice affects the DM's payoff for sure. Note that this is still a strategic interaction and a proper game but the interdependence has been deemphasized by the formatting of the information. The normal form representation and the corresponding decomposed game can be seen in Figure 2.[3]

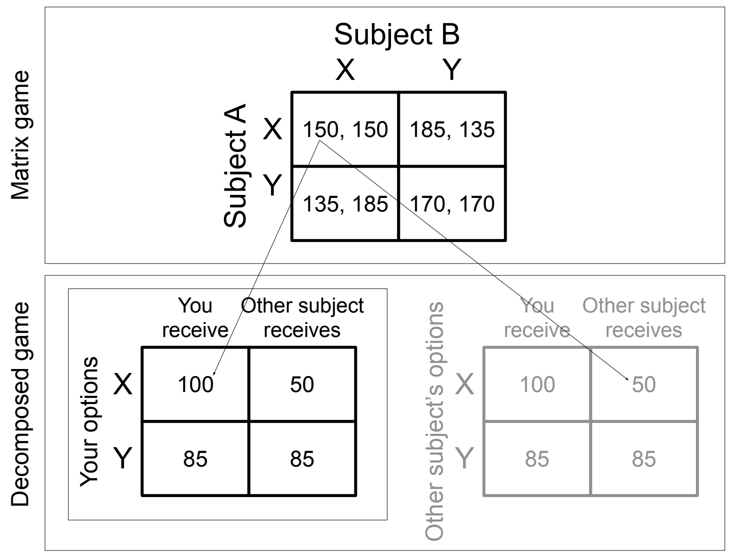

Figure 2: Decomposed game and matrix game. In the matrix game subject A chooses a row and subject B simultaneously chooses a column. In the decomposed game, both subjects simultaneously choose an allocation and the final payoffs are given by the sum of the allocations.

The use of decomposed games may be a misguided experimental manipulation as it does nothing to mitigate strategic interdependence. Decomposed games are still strategic games, but merely designed to be obtuse and thus harder for subjects to understand. Obfuscation is not the same as experimental control, and experimenters can not blur away strategic interdependence. Other psychologists have recognized this and have used fundamentally different experimental settings (without strategic interdependence) in order to establish an unconfounded measure of distributional preferences (Kahneman et al., 1986). However, due to tradition and inertia, decomposed games are still the predominant method employed when measuring distributional preferences under the rubric of SVO.[4]

[3]Note that there are many different possible decompositions of the same matrix game, but only one matrix game that can be recomposed from any decomposed game. Matrix games that can be decomposed are also called separable games (Hamburger, 1969).

[4]A search on Google scholar (for January 2013 to June 2015) returned 23 experiments in which the measurement of SVO was incentivized. Out of those 23 experiments, 16 used the decomposed game method while only 3 used the role

A charitable interpretation of the decomposed games approach is that the technique acts like a kind of choice framing and directs a subject's attention to the consequences of her own action, downplaying the potential consequences from the other's choice. Pruitt (1967) makes exactly this argument by suggesting that the decomposed representation better captures how people perceive strategic interactions in the wild, not as the more austere matrix representation that lays bare the strategic interdependence.

3 Hypotheses

Before we derive the hypotheses we briefly discuss the experimental design, because understanding the design helps make the hypotheses clearer. The design is summarized in Figure 3. The main motivation is to use the same stimuli, but systematically vary the social context. There are three different procedures for role assignment and one of the role assignment procedures has two ways of displaying information about choices and payoffs. This yields four treatments: the Role Certainty Treatment (RCT), the Role Uncertainty Treatment (RUT), the Decomposed Game Treatment (DGT), and the Matrix Game Treatment (MGT). Although matrix games have not been used to measure distributional preferences, we include the MGT for several reasons. It allows us to compare results from DGT and MGT and thus analyze how the description (i.e., decomposed framing) affects behavior. The inclusion of a standard matrix game also can serve as a benchmark for behavior and contrasted against the other treatments. We use a fully within-subject design with random orderings of treatments to mitigate potential carryover effects.

3.1 H1: Incentives reduce distributional preferences

In the RCT, half of the subjects filled out the Slider Measure with incentives knowing they were the *Deciders*. The other half filled out the Slider Measure knowing they were the *Receivers* and that

uncertainty method. The remaining 4 could not be classified because the papers did not contain enough information to make a determination.

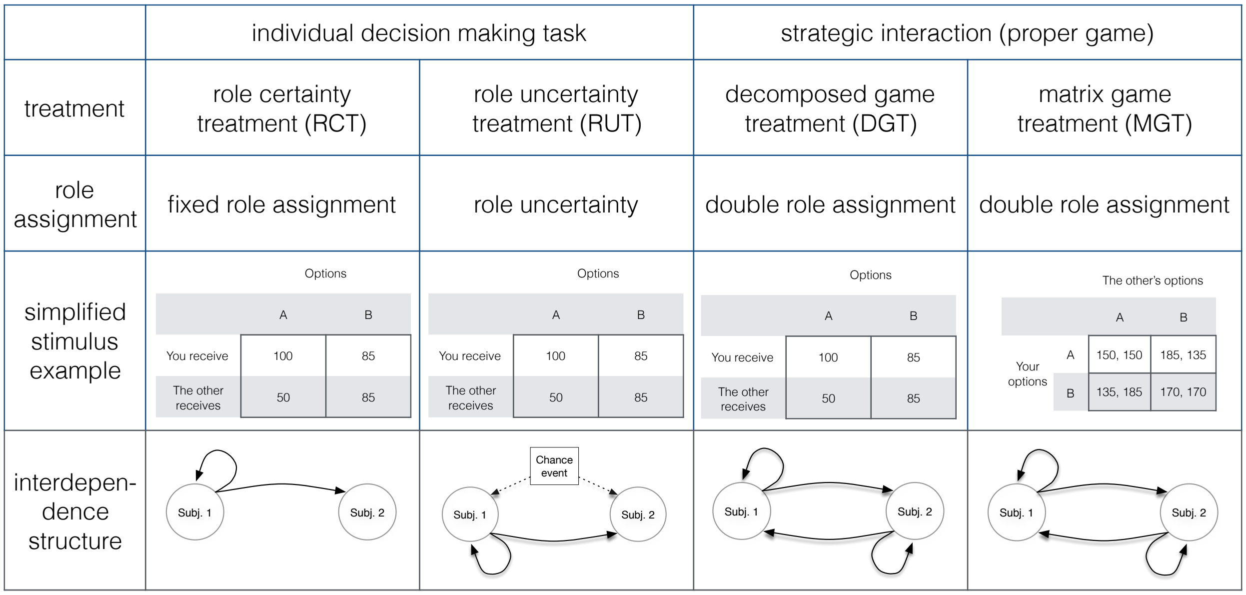

Figure 3: This provides an overview of the experimental design. The figures in the last row show how influence propagates from a subject to affect her own payoff as well as the payoff of another subject.

their responses would not have any effect on themselves nor any other person. This implementation naturally yields a test of the effect of incentives on distributional preferences.

This kind of study has been conducted before. Mentzakis and Mestelman (2013) and Engel (2011) found no significant effects of incentives on generosity, but in Forsythe et al. (1994), subjects are more generous if decisions are hypothetical as opposed to incentivized (see also Camerer and Hogarth, 1999).

To see how financial incentives affect behavior, we compare mean SVO scores within the RCT, comparing subjects who made incentivized choices with subjects who made hypothetical choices. Although previous findings are mixed, we expect that incentivized choices will decrease distributional preferences.

3.2 H2: Role certainty reduces distributional preferences

Iriberri and Rey-Biel (2011) compare behavior in modified dictator games with role certainty and role uncertainty and find that subjects are more generous with role uncertainty. This could be driven by fairness considerations or by the cost of expressing generosity.

First, fairness considerations and positive expectations could lead to more prosocial choices in the RUT. Assume that the choice of an allocation that would result in a high payoff for the other subject is perceived as "nice". If A has a preference for fairness and expects B to choose the "nice" allocation, A might reciprocate by also choosing the "nice" allocation (Rabin, 1993). Here, A's choice depends on her expectation about B's intention.

Second, higher generosity with role uncertainty could be due to the low costs of expressing generosity (see also the literature on low cost expressive voting, e.g., Kirchgässner, 1992; Hillman, 2010). Assume that being generous increases utility because it produces a *warm glow* (Andreoni, 1990) or because subjects derive utility from expressing their generosity (Hillman, 2010). With role certainty, the additional utility from being generous comes with a reduction in the own payoff. This is different with role uncertainty. With role uncertainty, choosing a more prosocial allocation in the role of *Decider* reduces the expected payoff. The actual payoff is reduced only if the subject is assigned to the role of *Decider*. Hence, the expected price of expressing generosity is lower in the RUT.[5]

Due to fairness considerations and the low costs of expressing generosity, we expect subjects to be more generous with role uncertainty compared to role certainty. To address this hypothesis we compare subjects' SVO scores in the RCT with their SVO scores in the RUT. We expect SVO to be higher with role uncertainty.

3.3 H3: Strategic interdependence increases distributional preferences

Shafir and Tversky (1992) report the results of an experiment where 25% of choices from subjects playing a prisoner's dilemma had the following pattern: Subjects chose to defect when they knew the other subject had chosen to defect, they chose defect when they knew the other subject had cho-

[5] Say as an extreme example that subject A would be the *Receiver* with $p = 0.999$; then it costs her almost nothing in expectations to act nicely; she can "buy" all the warm glow at almost none of the cost, or express her generosity without reducing her material payoff. If the tables were turned and she were to be the *Decider* with $p = 0.999$, then she may be more careful in how she traded off her resources for subject B.

sen to cooperate, yet they chose to cooperate when they did not know the other subject's choice. This kind of non-consequentialist reasoning indicates that a sizable proportion of subjects are sensitive to strategic interdependence and engage in a kind of "magical thinking" about the potential effects of their choices. A possible interpretation is that in simultaneous games, subjects change their perspective towards "collective rationality" and choose to cooperate with the goal of maximizing the sum of payoffs (Shafir and Tversky, 1992, 457). In the sequential prisoner's dilemma, the second mover has control over both outcomes once the first mover has made her choice. This undermines team reasoning and subjects tend to choose the alternative that maximizes their own payoffs (Shafir and Tversky, 1992, 457).

A similar kind of non-consequentialist reasoning could be at work in DGT and MGT, the two treatments in which subjects are exposed to strategic interdependence. If this is the case, subjects would choose prosocial allocations in treatments DGT and MGT, compared to the treatments RCT and RUT, in which the outcome depends only on one subject's choice. From this we conjecture that strategic interdependence will increases distributional preferences.

3.4 H4: Decomposed games increase distributional preferences relative to matrix games

Empirical results reveal that behavior in decomposed games and strategically equivalent matrix games differs. Pruitt (1967) compared behavior in a repeated prisoner's dilemma and found a significant effect on behavior. For most decompositions, cooperation rates where higher when information was presented in form of a decomposed game. Qualitatively similar results have been obtained in other studies (Evans and Crumbaugh, 1966; Crumbaugh and Evans, 1967; Gallo Jr. et al., 1969; Pincus and Bixenstine, 1977). Analogously, we expect that, compared to the MGT, the mean SVO will be higher in the DGT.

3.5 H5: Behavior in decomposed games is more similar to behavior in matrix games than to behavior in resource allocation tasks with role certainty

Our four treatments correspond to four different methods for the measurement of distributional preferences. The correlations among the different methods can be treated as a method similarity matrix, which itself can be the basis for further analysis using classical multi-dimensional scaling with one dimension. This kind of analysis would provide evidence into the "distance" among the different methods; methods that yield similar behaviors from DMs would be closer to each other, whereas methods that yielded different behaviors would be further away. Moreover, if the different methods were inconsequential to behavior (i.e., if they all measure the same thing), then the correlations between the treatments would be very high and approximate test-retest correlation levels.

It is worth noting that this hypothesis is not concerned with the mean levels of prosociality, as the previous four hypotheses are. Rather, it is concerned with the consistency of behavior from individual subjects across the different treatments. This analysis is useful in that it is sensitive to heterogeneity in responses. Say for example that half of the subjects became more prosocial in a particular treatment and half become less prosocial in the same treatment (compared to a baseline treatment). This result would yield a consistent mean level of prosociality, but it would not indicate that subjects were behaving consistently. It would be evidence that the different experimental treatments had effects, but the response from subjects are not systematic.

4 Experimental procedure

The experiment was conducted at the University of Giessen using zTree (Fischbacher, 2007). A total of 192 subjects participated. Upon arrival, subjects were asked to take a seat at a randomly assigned computer workstation where they found general instructions which informed subjects that the experiment consists of four independent treatments. In addition to the general instructions, each subject found four sealed and differently colored envelopes on her desk, containing the instructions for the four different treatments.

Subjects were informed that at the beginning of each part, the on-screen instructions would inform them which colored envelope to open. This procedure ensures that subjects are informed about each treatment only when the treatment begins, but that the treatments exist *a priori* and are independent of their choices.

In order to avoid that subjects act prosocially in order to elicit kindness from reciprocators in subsequent tasks or treatments, subjects received feedback about results and earnings only at the very end of the experiment.

The order in which subjects encountered the four treatments was fully counterbalanced across sessions. That is, we implemented all permutations of the orders of treatments, resulting in $4! = 24$ orders that were each implemented in a separate session. The number of subjects per session was eight, and a no-contagion matching, which was also explained to subjects, was used (Kamecke, 1997).

At the beginning of each treatment, subjects opened the envelope containing the treatment-specific instructions which were also read aloud by the experimenter, and subjects had the opportunity to ask questions privately. Before starting the experimental task subjects had to answer a control question correctly to verify their comprehension.

Each treatment consists of six different decision situations. In treatments RCT, RUT and DGT the six decision situations correspond to the six resource allocation tasks which comprise the six primary items of the SVO Slider Measure (see Figure 1). In treatment MGT the six decision situations correspond to six matrix games (see Table 1 for an example and appendix B for all six matrix games).[6]

[6] The resource allocation tasks were implemented in terms of zTree modules (Crosetto et al., 2012) that were slightly modified so that items consists of five options per task rather than nine options. We provide subjects with five options per task to ensure that the recomposition of the corresponding decomposed games into matrix games in normal form would not overwhelm subjects with too many strategies. In the most commonly used version, each item from the Slider Measure corresponds to the choice between nine alternatives. Recomposing these Slider Measure items would have resulted in a 9 x 9 payoff matrix with 81 cells. Reducing to 5 options yielded a matrix game that had only 25 cells which we considered experimentally tractable.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	150 , 150	159 , 146	168 , 143	176 , 139	185 , 135
	Option B	146 , 159	155 , 155	164 , 152	172 , 148	181 , 144
Your options	Option C	143 , 168	152 , 164	161 , 161	169 , 157	178 , 153
	Option D	139 , 176	148 , 172	157 , 169	165 , 165	174 , 161
	Option E	135 , 185	144 , 181	153 , 178	161 , 174	170 , 170

Table 1: Recomposition of task number six into a matrix game. In treatment MGT, six matrices (derived from the corresponding decomposed games) were presented to subjects and subjects were asked to choose their most preferred allocations.

Within each treatment, task order and option order of the six resource allocation tasks was fully randomized. These randomization procedures were implemented to control for order effects both within and across treatments, and moreover, to mitigate carryover effects from subjects remembering their previous choices in particular items and simply attempting to reproduce their former choice patterns rather than reflecting anew on their preferences in each task and treatment.

Subjects were informed that each subject's final payoff was given by the sum of each treatment's payoff plus a fixed show-up fee (5 euros). For each treatment, one of the six decision situations from each treatment was randomly selected, and the decisions made by both members in that decision situation determined payoffs from this treatment.

Role Certainty Treatment (RCT): Subjects are matched pairewise. One subject is assigned the role of the *Decider* while the other is assigned the role of the *Receiver*. Both subjects make the six resource allocation decisions. Only the decisions of the *Decider* are relevant for payment. The *Receiver* makes only hypothetical decisions that are neither implemented nor communicated.

Role Uncertainty Treatment (RUT): Subjects are matched pairewise. Both subjects make the six resource allocation decisions. Only after both subjects of the pair have made all six decisions, a coin flip decides whose decisions will determine payoffs.

Decomposed Game Treatment (DGT): Subjects are matched pairewise. Both subjects make the six resource allocation decisions. After all decisions have been made, the decisions made by both subjects determine payoffs.

Matrix Game Treatment (MGT): The MGT is strategically identical to the DGT, however the information about payoffs was presented as a payoff matrix. One subject was told to select a row while the other subject would simultaneously select a column. The resulting payoffs would be the intersection of these choices.

5 Results

Each of the 192 subjects participated in all four treatments and made six decisions in each treatment. In total, this makes $192 \times 4 \times 6 = 4,608$ decisions. In RCT only half of decisions were incentivized. If not explicitly stated otherwise, only incentivized decisions are used for statistical analysis. Recall that in each treatment, the six decisions correspond to the six primary items of the SVO Slider Measure (see Figure 1). Hence, we can compute a subject's SVO for each treatment. The experiment lasted 60 minutes and the average payment, including the show-up fee, was 14.55 Euros.

5.1 Aggregate data on SVO

The observed preference distributions are consistent with data from earlier experiments in which SVO was measured (Murphy et al., 2011). In all treatments the most common SVO score is 7.82, corresponding to the maximization of the own payoff.[7] Table 2 summarizes the data on subjects' SVOs. The first four columns contain the descriptive statistics for incentivized choices by treatment. In RCT half of the subjects made hypothetical decisions which are summarized in column 6. Pooled observations from RCT (incentivized and hypothetical decisions) are summarized in column 5.

The scatterplots (Figure 4) and the corresponding rank correlations reveal that there is substantial heterogeneity at the individual level. Each scatterplot compares two treatments and reports the

[7] The reason the score is not exactly 0 is because one of the SVO slider items (task 1 in Figure 1) is structured such that the DM's payoff is invariant at 85 and the DM chooses a payoff for the other ranging between 15 and 85 . The vast majority of subjects from Western countries, including otherwise completely self-interested ones, choose 85 for the other person on this item. This choice has no cost to the DM and the result is their SVO score is not exactly 0.

	RCT(inc.)	RUT	DGT	MGT	RCT(all)	RCT(hyp.)
n	96	192	192	192	192	96
Mean	16.46	17.98	17.30	18.83	17.98	19.51
Median	11.36	16.26	14.94	16.26	14.62	16.73
Std	13.06	14.40	13.97	16.92	14.74	16.18

Table 2: Descriptive analysis on subjects' SVOs. The first four columns contain the descriptive statistics for incentivized choices by treatment. In RCT half of the subjects made hypothetical decisions which are summarized in column 6. Pooled observations from RCT (incentivized and hypothetical decisions) are summarized in column 5.

corresponding rank order correlation. If all subjects' SVOs were identical in two treatments, all of the points would be on the 45-degree line and the correlation would be 1. If all subjects were more prosocial in one treatment, then the points would be either above or below the 45-degree line (depending on which treatment corresponds to which axis). Nonetheless, the correlation could still be 1 if the magnitude of change was the same for all subjects. However, we see that for all pairwise comparisons, there are data points both above and below the 45-degree line. This indicates changes in distributional preferences at the individual level. When moving from one treatment to the other, for some subjects SVO increases while for other subjects SVO decreases. In fact, there are only 16 out of 192 subjects for whom SVOs were identical in all treatments in which they made incentivized decisions.[8] For a total of 37 subjects, SVOs do not differ by more than 5 between all treatments in which they made incentivized decisions.

Hypothesis 1-4 are not concerned with inconsistency at the individual level. Rather, they analyze whether subjects are, on average, more prosocial in one treatment compared to another. These aggregate level effects are identified by looking for shifts in the central tendency of the distributions. Hypothesis 5 is different and is concerned with preference stability at the individual level, not general aggregate shifts in the distribution nor its central tendency.

[8]For 9 subjects, who made hypothetical decisions in RCT, SVOs were exactly the same in RUT, DGT and MGT; for 7 subjects, who made incentivized decisions in RCT, SVOs were exactly the same in all four treatments.

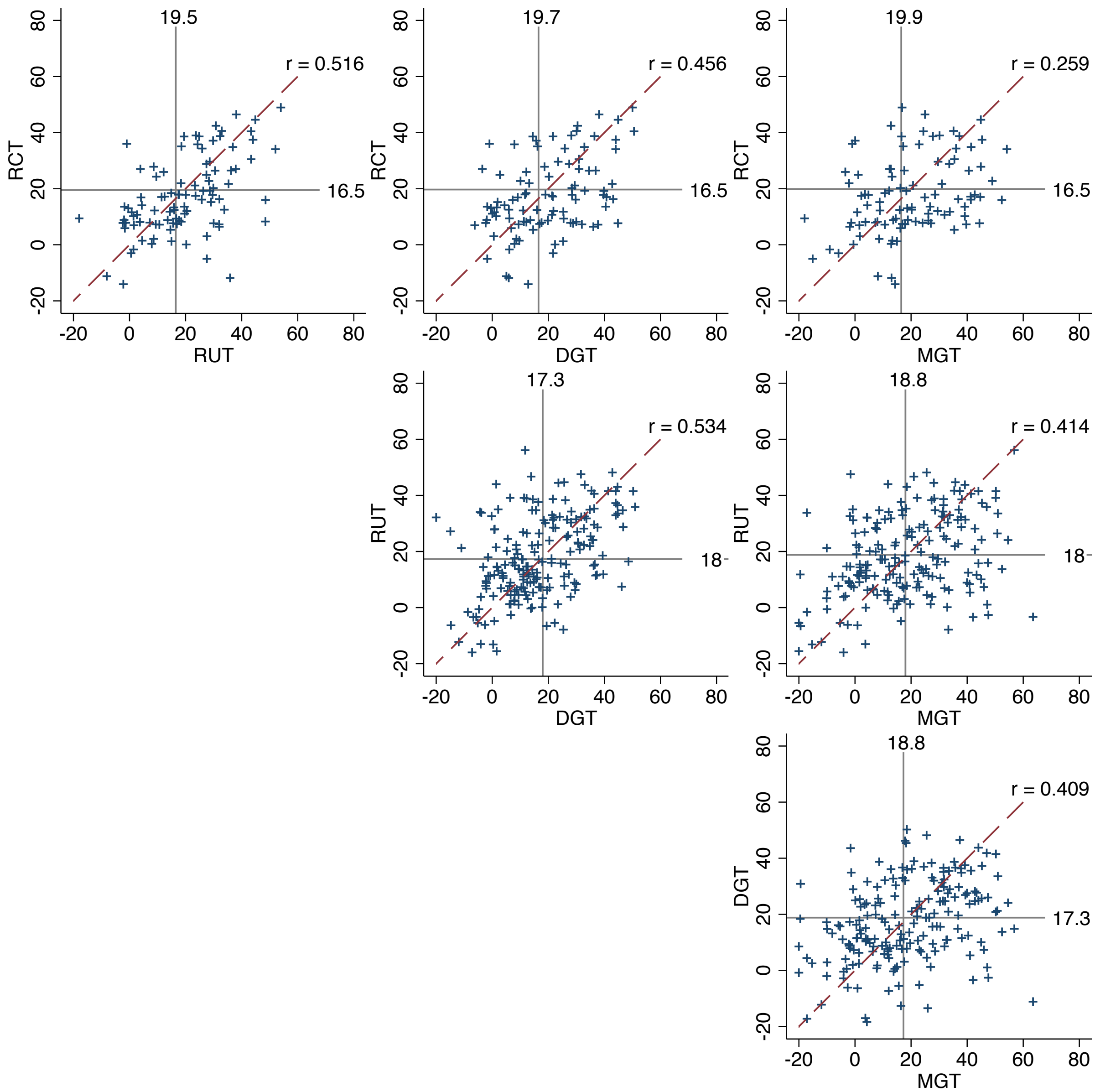

Figure 4: SVO scatterplots and Spearman's rank correlations (r). Horizontal and vertical lines represent mean values of the corresponding treatment. (Some mean values differ from the mean values in Table 2 because in this figure, mean values are computed over all subjects who made incentivized decisions in both treatments, the treatment on the x-axis and the treatment on the y-axis.)

5.2 Hypotheses tests and discussion

In this section we present the results for the hypotheses discussed above. After stating each result, we describe the statistical tests that support the result, followed by a short discussion.

Result 1: *In RCT there is no significant difference between incentivized and hypothetical decisions.*

A Mann-Whitney test reveals that there is no shift on average since the medians are not significantly different ($p = 0.140$). A two-sided t-test ($p = 0.153$) also shows that the means are not significantly different and corroborates this result.[9]

Although we do not find a difference in generosity, financial incentives reduce variance (see also Camerer and Hogarth, 1999, 31). A Levene test shows that variances are significantly smaller if decisions are incentivized ($p = 0.004$). If hypothetical decisions inflate variance because subjects are more likely to make mistakes or respond haphazardly, then it appears that incentives are an effective way to reduce noise in the measurement of distributive preferences.

Result 2: *On average, we find that subjects in the RUT are more prosocial than in the RCT.*

Using a paired t-test we find that in RUT, mean SVO is significantly larger than mean SVO in RCT ($p = 0.029$).

Our experiment cannot tell whether this result is driven by fairness considerations or by the cost of expressing generosity. However, the implication is clear: If SVO is measured with role uncertainty, a subject's SVO should be considered as an upper bound.

Result 3: *SVO does not differ significantly between treatments (RCT,RUT) and (DGT,MGT), indicating that people are prosocial to about the same degree when contrasting strategic and non-strategic settings.*

To test for this effect, we computed each subject's average SVO from treatments (RCT,RUT) and (DGT,MGT). Call these averages $SVO(nonstr.)$ and $SVO(str.)$. The average $SVO(nonstr.)$ is 17.97 and the average $SVO(str.)$ is 19.80. A Wilcoxon signed rank test shows that the median

[9] To check the robustness of results 1 to 4 we performed bootstrapped t-tests and excluded subjects who violated transitivity in RCT or RUT from the analysis. The results are qualitatively the same.

of differences is not significantly different from zero ($p = 0.080$). A paired t-test corroborates this finding ($p = 0.122$).

We find no evidence for "magical thinking" or "team reasoning" in the treatments with strategic interdependence. Nonetheless, the mean SVO is highest in the MGT. A possible explanation is that in the DGT, the framing sufficiently "hides" the other subject so that "magical thinking" is less likely.

Result 4: *On average, SVO is not significantly different between the DGT and the MGT, indicating that the formatting of information does not have a significant effect on prosociality.*

Using a Wilcoxon signed rank test ($p = 0.295$) we find the median of SVO difference scores is not significantly different from zero. A paired t-test ($p = 0.216$) shows the same.

This result does not imply that both methods lead to the same result. In fact, the correlation between both treatments is only 0.409, and variances are significantly higher in the MGT (Levene test, $p = 0.001$). How to interpret the low correlation and the difference in variances? Our results indicate that the majority of subjects are sensitive to the strategic framing, but respond in different ways. The framing effect is driven by the salience of strategic considerations, which invokes fairness preferences and home grown expectations. The reasoning is as follows: The availability of the payoff matrix increases the salience of the strategic nature of the interaction. With respect to DGT and MGT, this implies that the other subject's expected decision is less salient in the DGT, and more salient in the MGT. With strategic considerations being more salient, subsequent choice behavior could be driven by home grown expectations as well as preferences. Because of psychological considerations, like fairness or reciprocity, this decision task is about coordination, not just cooperation, and individual differences in expectations, which are amplified by the salience of interdependence, can explain the larger variance in MGT.[10]

Result 5: *Behavior in decomposed games is closer to resource allocation tasks with role certainty than to matrix games.*

[10]Theoretically, in DGT and MGT there are multiple equilibria if subjects have preferences for fairness (Rabin, 1993) or inequality aversion (Fehr and Schmidt, 1999) provided these social preferences are common knowledge. This gives rise to the problem of equilibrium selection.

This result is consistent with both a visual inspection of the difference matrix in Table 3 and multidimensional results (explained below). The RCT and MGT produce the largest difference between methods (0.741), and the DGT and RUT produced the smallest difference (0.466). However all of these differences are more that would be expected from chance alone. The test-retest reliability of the SVO Slider Measure is reported as 0.915 (Murphy et al., 2011), implying an expected difference of about 0.085 (95% CI 0.065 to 0.111) due to measurement error alone. The observed differences are substantially larger than this, indicating significant effects of social context on the measurement of distributive preferences at the individual level.

	RCT	RUT	DGT	MGT
RCT	0	-	-	-
RUT	0.484	0	-	-
DGT	0.544	0.466	0	-
MGT	0.741	0.586	0.591	0

Table 3: Difference matrix among the different treatments. The entries are 1 minus the correlation between two methods. Higher numbers indicate more inconsistency in subject's behavior between two methods.

Results from multidimensional scaling yield the emergence of the rank ordering of the methods in terms of their similarity. The resulting ordering is: RCT, RUT, DGT, and MGT. The output scores are -0.4404, -0.1296, -0.0434, 0.6135 respectively. The stress of the model is (1.65e-16) indicating adequacy of the goodness of fit of the simple unidimensional model. These multidimensional scaling scores can be normalized between 0 and 1 (see Figure 5), and the results show that behavior in DGT is closer to behavior in RCT than to MGT.[11]

Result 5 shows a substantial effect of "framing" on decision making, and how a game's presentation can have a substantial effect on subsequent choices, even though the underling game itself remains the same. Description invariance is clearly violated for the vast majority of subjects.

Note that the large differences in scores between DGT and MGT is not inconsistent with Result

[11] Another way to evaluate hypothesis 5 is to compute a new variable for the absolute distance between a subject's SVO in treatments x and y, $\Delta(x, y) = abs(SVO(x) - SVO(y))$. Using a sign test, we can show that $\Delta(DGT, MGT)$ is significantly larger than $\Delta(RCT, DGT)$, ($p = 0.0024$), corroborating the results from the scaling analysis.

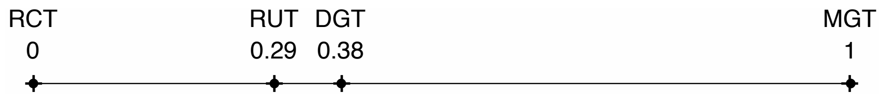

Figure 5: Normalized multidimensional scaling scores from each of the four treatments mapped onto a unidimensional solution. This analysis yields the relative similarity (behavioral consistency) among the different treatments.

4 since Result 4 is concerned with shifts in central tendency whereas Result 5 is concerned with preference stability (at the individual level) between treatments. The central tendency of the DGT and MGT distributions are similar but that does not imply that individual behavior in the treatments is similar. Rather we see that many decision makers respond to the strategic framing, but do so in different ways. Some DMs become more prosocial, some less so. There is not a general tendency or direction in the direction the scores change (as indicated by Result 4). However there is quite a bit of between treatment variance in scores (as indicated by Result 5).

5.3 Implications for the measurement of distributional preferences

This paper operated under a very broad (and perhaps straw-man) null hypotheses that social context would be unrelated to subjects' SVO scores. We can reject that general null hypotheses given the pattern of results reported above.

In resource allocation tasks with role certainty it is clear that strategic considerations are absent because the other subject affected by the DM's behavior is completely passive. Kerschbamer (2015b) argues that role certainty "seems to be the cleanest procedure from a theoretical point of view" (p. 1). We fully agree with this viewpoint and the results from the RCT provide the most straightforward evidence about subject's distributive preferences. Moreover, we show evidence that different subjects respond differently as role uncertainty and strategic interdependence is introduced. This increased social interdependence may be of interest, but it creates complexity and creates multiple confounding factors which make valid interpretations about revealed distributional preferences difficult or impossible to make. Because of this, our conclusions echo Kerschbamer who advocates measuring distributive preferences with role certainty.

When SVO is measured with role uncertainty, the resulting score should be taken as an upper bound to an individuals' distributional preferences. Role uncertainty is clearly a more efficient measurement method but it comes with trade-offs. Experimenters should weigh the benefits of the efficiency advantage with the possibility of inflating scores due to role uncertainty. Luckily for the efficiency-minded researcher the effect is not massive, and for work that only requires a rough approximation of distributive preferences the RUT approach may be sufficient.

We conclude that the measurement of distributional preferences is improved with proper incentives, but the effect is more about variance than an aggregate shift in preferences. The results from our experiment and from Mentzakis and Mestelman (2013) find no significant differences in average SVO given incentives. Nonetheless, we do not agree with Mentzakis and Mestelman (2013) who argue that eliciting SVO without financial incentives comes with no disadvantages. The difference in means (16.46 vs. 19.51), although not statistically significant, points to the presence of other factors, which may inflate SVO if no incentives are used. Moreover, hypothetical decisions significantly inflate variance (see also Camerer and Hogarth, 1999), which reduces statistical power.

The decomposed games technique might be problematic if SVO is measured several times, for example in social ties experiments (e.g., van Dijk et al., 2002). If subjects perceive the game correctly, and further have preferences for fairness or inequality aversion, behavior depends on expectations, and these belief based expectations could develop and change between measurements (Greiff, 2013). To avoid these problems of confounding the measurement of preferences with emergent expectations, we would advise against using the decomposed game method.

This study is silent to which of these social contexts provides the *best* way to measure SVO, i.e., identifying a subject's true score of her distributional preferences. In order to establish evidence to address this question, one would have to measure SVO using a variety of different social contexts and then use these different individual scores to predict other behavior (predictive validity), correlate the SVO scores on other measures (convergent validity), and show how the SVO scores were separate from other constructs like beliefs (divergent validity). Ironically, a confounded measure of SVO (simultaneously measuring both preferences and beliefs) may make better predictions of cooperative choices rather than a pure measure of distributional preferences, as beliefs are also a good predictor of cooperative behavior. It is for this reason that one would have to carefully measure beliefs in tandem with preferences, in order to establish the unique predictive capacity of SVO for cooperative behaviors. All of this is beyond the scope of the current paper, which had the more modest goal of identifying to what degree social contexts affected subject's SVO scores.

6 Conclusion

This article contributes to the small but growing literature on procedural differences and their effects on prosocial behavior. In our experiment we systematically studied the effect of three different methods to measure social value orientation (RCT, RUT, DGT), and included an obvious strategic choice (MGT) setting, all using a within-subjects design. Although laboratory experiments of distributive preferences offer a relatively clean environment, there remains at least some minimal social context in which the choices are embedded, as there is some "other" whose payoff is influenced by the choices of the DM. The role assignments create part of the social context, and we investigated how social context and its salience influences behavior. Overall results show that in the aggregate, distributional preferences are higher when assessed under role uncertainty or in decomposed games, compared to role certainty. Importantly, at the individual level, there are substantial non-systematic differences between treatments.

Our findings imply that, depending on the social context, the measurement of distributional preferences is affected by other factors than the resulting allocation of payoffs. At the aggregate level, the differences between role certainty and role uncertainty are consistent with beliefs about other's intentions or self-presentation concerns. With role uncertainty, subjects might be more generous because the expected cost of expressing generosity is lower, or they might be more generous because they expect the other subject to be generous somehow in response (even though it is a simultaneous choice). At the individual level, the differences between role certainty and decomposed games can be partially accounted for by strategic considerations, which matter if subjects are inequality averse or concerned about fairness. Behavior depends on expectations as well as preferences, hence, the measurement of distributional preferences under these conditions would almost certainly be confounded with beliefs.

The comparison of decomposed games and matrix games adds to the literature on presentation effects (Charness et al., 2004; Requate and Waichman, 2010; Ferraro and Vossler, 2010; Gürerk and Selten, 2011). The comparison of role assignment procedures also complements the studies by Murphy et al. (2011) and Kerschbamer (2015a), which discuss how to identify distributional preferences based on subjects' choices in several resource allocation tasks but do not discuss the issue of role assignment. A better understanding of the influence of social context on behavior is desirable because it leads to a cleaner measurement of distributional preferences. From the perspective of a practitioner, who would like to control for subjects' preferences, a more precise measurement allows for better mitigation, which improves statistical power and helps make predictions more effectively.

The overall pattern of results leads us to the following conclusions: Different people have distributional preferences to different degrees. We know from previous research that these preferences are relatively stable in the absence of information (i.e., in a test-retest sense). What we discover here is that these distributional preferences are sensitive to social contexts and increasing the social interdependence and its salience, causes different people to respond in fundamentally different ways– some people become nicer, some become more selfish. We conjecture that the heterogeneity in response to different social contexts is driven in large part by heterogeneous beliefs about what others are like. DMs homegrown preferences are made more salient by different social contexts, and the result is that DM's distributional preferences, as measured by the SVO Slider Measure, change in different ways as function of these beliefs.

References

Andreoni, James, "Impure altruism and donations to public goods: A theory of warm-glow giving," *The Economic Journal*, 1990, *100* (401), 464–477.

— and John Miller, "Giving according to GARP: An experimental test of the consistency of preferences for altruism'," *Econometrica*, 2002.

Balafoutas, Loukas, Rudolf Kerschbamer, and Matthias Sutter, "Distributional preferences and competitive behavior," *Journal of Economic Behavior and Organization*, June 2012, *83* (1), 125–135.

— , — , Martin Kocher, and Matthias Sutter, "Revealed distributional preferences: Individuals vs. teams," *Journal of Economic Behavior and Organization*, December 2014, *108*, 319–330.

Balliet, Daniel, Craig Parks, and Jeffrey Joireman, "Social value orientation and cooperation in social dilemmas: A meta-analysis," *Group Processes & Intergroup Relations*, JUL 2009, *12* (4), 533–547.

Bardsley, Nicholas, "Dictator game giving: altruism or artefact?," *Experimental Economics*, September 2007, *11* (2), 122–133.

Blanco, Mariana, Dirk Engelmann, and Hans-Theo Normann, "A within-subject analysis of other-regarding preferences," *Games and Economic Behavior*, June 2011, *72* (2), 321–338.

Camerer, Colin F. and Ernst Fehr, "Measuring social norms and preferences using experimental games: A guide for social scientists," in "Foundations of Human Sociality," Oxford University Press, 2004, pp. 55–95.

— and Robin M. Hogarth, "The effects of financial incentives in experiments: A review and capital-labor-production framework," *Journal of Risk and Uncertainty*, 1999, *19* (1-3), 7–42.

Charness, Gary and Brit Grosskopf, "Relative payoffs and happiness: An experimental study," *Journal of Economic Behavior and Organization*, 2001, *45* (3), 301–328.

— , Guillaume R. Frechette, and John H. Kagel, "How robust is laboratory gift exchange?," *Experimental Economics*, 2004, *7* (2), 189–205.

Crosetto, Paolo, Ori Weisel, and Fabian Winter, "A flexible z-Tree implementation of the social value orientation slider measure (Murphy et al. 2011): Manual," *Jena Economic Research Papers*, 2012.

Crumbaugh, Charles M. and Gary W. Evans, "Presentation format other-person strategies and cooperative behavior in prisoner's dilemma," *Psychological Reports*, 1967, *20* (3P1), 895–902.

Dana, Jason, Roberto A. Weber, and Jason Xi Kuang, "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness," *Economic Theory*, September 2007, *33* (1), 67–80.

Engel, Christoph, "Dictator games: A meta study," *Experimental Economics*, May 2011, *14* (4), 583–610.

Engelmann, Dirk and Martin Strobel, "Inequality aversion, efficiency, and maximin preferences in simple distribution experiments," *The American Economic Review*, September 2004, *94* (4), 857–869.

Evans, Gary W. and Charles M. Crumbaugh, "Payment schedule, sequence of choice, and cooperation in the prisoner's dilemma game," *Psychonomic Science*, 1966, *5* (2), 87–88.

Fehr, Ernst and Klaus M. Schmidt, "A theory of fairness, competition, and cooperation," *The Quarterly Journal of Economics*, 1999, *114* (3), 817–868.

_ , Helen Bernhard, and Bettina Rockenbach, "Egalitarianism in young children," *Nature*, August 2008, *454* (7208), 1079–1083.

Ferraro, Paul J. and Christian A. Vossler, "The source and significance of confusion in public goods experiments," *The B.E. Journal of Economic Analysis & Policy*, 2010, *10* (1).

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Fisman, Raymond, Shachar Kariv, and Daniel Markovits, "Individual Preferences for Giving," *The American Economic Review*, December 2007, *97* (5), 1858–1876.

Forsythe, Robert, Joel L. Horowitz, N. E. Savin, and Martin Sefton, "Fairness in simple bargaining experiments," *Games and Economic Behavior*, 1994, *6* (3), 347–369.

Gallo Jr., Philip S., Sandra G. Funk, and Joseph R. Levine, "Reward size, method of presentation, and number of alternatives in a prisoner's dilemma game.," *Journal of Personality and Social Psychology*, November 1969, *13* (3), 239–244.

Greiff, Matthias, "Affective social ties without the need to belong?," *Rationality, Markets and Morals*, 2013, *4*, 20–27.

Griesinger, Donald W. and James W. Livingston, "Toward a model of interpersonal motivation in experimental games," *Behavioral Science*, 1973, *18* (3), 173–188. doi:10.1002/bs.3830180305.

Gürerk, Özgür and Reinhard Selten, "The effect of payoff tables on experimental oligopoly behavior," *Experimental Economics*, December 2011, *15* (3), 499–509.

Hamburger, Henry, "Separable games," *Behavioral Science*, 1969, *14* (2), 121–132.

Hillman, Arye L., "Expressive behavior in economics and politics," *European Journal of Political Economy*, December 2010, *26* (4), 403–418.

Iriberri, Nagore and Pedro Rey-Biel, "The role of role uncertainty in modified dictator games," *Experimental Economics*, 2011.

Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler, "Fairness and the assumptions of economics," *The Journal of Business*, October 1986, *59* (4).

Kamecke, Ulrich, "Rotations: Matching schemes that efficiently preserve the best reply structure of a one shot game," *International Journal of Game Theory*, 1997, *26* (3), 409–417.

Kerschbamer, Rudolf, "The geometry of distributional preferences and a non-parametric identification approach," *Working Papers in Economics and Statistics, University of Innsbruck*, 2013.

— , "The geometry of distributional preferences and a non-parametric identification approach – The equality equivalence test," *European Economic Review*, May 2015, *76* (C), 85–103.

— , "The geometry of distributional preferences and a non-parametric identification approach: The equality equivalence test, Online Appendix," *European Economic Review*, 2015, *76*, 0–9.

Kirchgässner, G, "Towards a theory of low-cost decisions," *European Journal of Political Economy*, 1992, *8* (2), 305–320.

Liebrand, Wim B. G., "The effect of social motives, communication and group size on behaviour in an N-person multi-stage mixed-motive game," *European Journal of Social Psychology*, July 1984, *14* (3), 239–264.

Mentzakis, Emmanouil and Stuart Mestelman, "Hypothetical bias in value orientations ring games," *Economics Letters*, September 2013, *120* (3), 562–565.

Messick, David M. and Charles G. McClintock, "Motivational bases of choice in experimental games," *Journal of Experimental Social Psychology*, 1968, *4* (1), 1–25.

Murphy, Ryan O. and Kurt A. Ackermann, "Social value orientation: Theoretical and measurement issues in the study of social preferences," *Personality and Social Psychology Review*, January 2014, *18* (1), 13–41.

— and — , "Social preferences, positive expectations, and trust based cooperation," *Journal of Mathematical Psychology*, 2015, *67*, 45 – 50.

— , — , and M. J. J. Handgraaf, "Measuring social value orientation," *Judgment and Decision Making*, 2011, *6* (8), 771–781.

Offerman, Theo, Joep Sonnemans, and Arthur Schram, "Value orientations, expectations and voluntary contributions in public goods," *The Economic Journal*, July 1996, *106* (437), 817–845.

Pincus, Jeffrey and V. Edwin Bixenstine, "Cooperation in the decomposed prisoner's dilemma game: A question of revealing or concealing information," *Journal of Conflict Resolution*, 1977.

Pruitt, Dean G., "Reward structure and cooperation: The decomposed prisoner's dilemma game.," *Journal of Personality and Social Psychology*, September 1967, *7* (1p1), 21–27.

Rabin, Matthew, "Incorporating fairness into game theory and economics," *American Economic Review*, 1993, *83* (5), 1281–1302.

Requate, Till and Israel Waichman, ""A profit table or a profit calculator?: A note on the design of Cournot oligopoly experiments," *Experimental Economics*, September 2010, *14* (1), 36–46.

Rigdon, Mary L. and Adam S. Levine, "The role of expectations and gender in altruism," *MPRA Paper 19372*, 2009.

Shafir, Eldar and Amos Tversky, "Thinking through uncertainty: Nonconsequential reasoning and choice," *Cognitive psychology*, 1992, *24* (4), 449–474.

van Dijk, Frans, Joep Sonnemans, and Frans van Winden, "Social ties in a public good experiment," *Journal of Public Economics*, 2002, *85* (2), 275–299.

van Lange, Paul A. M., Ellen de Bruin, Wilma Otten, and Jeffrey A. Joireman, "Development of prosocial, individualistic, and competitive orientations: Theory and preliminary evidence.," *Journal of Personality and Social Psychology*, 1997.

_ , René Bekkers, Theo N. M. Schuyt, and Mark van Vugt, "From games to giving: Social value orientation predicts donations to noble causes," *Basic and Applied Social Psychology*, 2007, *29* (4), 375–384.

A Experimental Instructions

Welcome to the experiment and thank you for your participation. Please read the instructions carefully. Do not talk to your neighbors during the entire experiment. If you have any questions please raise your hand. One of the experimenters will come to you and answer your questions in private. Following these rules is very important. Otherwise the results of this experiment will be scientifically worthless.

Please take your time reading the instructions and making your decisions. You are not able to influence the duration of the experiment by rushing through your decisions, because you always have to wait until the remaining participants have reached their decisions.

The experiment is completely anonymous. Neither during nor after the experiment you will be informed with whom you have interacted. No other participant will be informed about which role you were assigned to and how much you have earned. You will receive a show-up fee of 5 euros for your participation. Depending on your decisions and the decisions of the other participants you can additionally earn between 4 and 12 euros. You will be paid individually, privately, and in cash after the experiment. The expected duration of the experiment is 60 minutes. The exact course of the experiment will be described in the following.

The experiment consists of four parts which are independent from each other. Each part consists of six decision situations. The beginning of a new part of the experiment will be indicated on the screen. The instructions for each part of the experiment are in the colored envelopes. Please open the envelope with the relevant information only if the corresponding part of the experiment begins.

In each round, participants will be randomly matched in pairs, but you will never interact with the same participant twice. In other words, you can be sure that you will never interact with the same participant in several parts of the experiment. Furthermore, the matchings are done in such a way that the actions you take in one round cannot affect the actions of the people you will be paired with in later rounds. This also means that the actions of the participants you are paired with in a given round cannot be affected by your actions in earlier rounds. During the experiment, you will not receive any information about the decisions of the other participants, and other participants will not receive information about your decisions. During the experiment, the payoffs are denoted in points. The exchange rate between points and Euros is 1/50. In other words, for 50 points you receive 1 Euro.

Calculation of your final payoff: After the experiment, a decision situation is randomly selected from each part of the experiment. Each decision situation has the same probability for being selected. That is, in every part each decision situation could be the decision situation, which determined your payoff from this part. Therefore, it makes sense to treat each decision situation as if it is the payoff-relevant decision situation.

The payoff that you get in addition to the 5 euros is the sum of the payoffs in the randomly selected decision situations. At the end of the experiment, you will be informed about the four randomly selected decision situations, the corresponding payoffs, and the sum of the payoffs. There will be a brief questionnaire after the last round is completed. After completion of the questionnaire, you will receive your payoff in cash. Payoffs will be made in private so that other participants won't be informed about your payoff.

The experiment will begin shortly. If you have any questions please raise your hand and wait until someone comes to your place. Please do not talk to the other participants during the entire experiment. Thank you for participating.

Part "Red"

In this part of the experiment, there are two roles, Decider, and Receiver, who will be randomly chosen. At the beginning, you will be informed about your role. In each of the six decision situation, the Decider chooses her most preferred allocation. That is, the decisions of the Decider determines her own payoff and the payoff of the Receiver. The Receiver has no influence on her own payoff and the payoff of the Decider. The Receiver chooses the allocation, that she would have chosen if she were in the role of the Decider. That is, that the decisions of the Receiver are purely hypothetical.

Your payoff from this part is calculated as follows: At the end of this part, a decision situation is randomly drawn and the allocation chosen by the Decider determines Decider's and Receiver's payoffs. You will be informed about the chosen decision situation and the corresponding payoffs at the end of the experiment.

Part "Blue"

In this part of the experiment, there will be six decisions. In each decision, you choose your most preferred allocation. The other participant faces the same decisions and has exactly same information as you. When you and the other participant have taken all decisions, it will randomly be determined whether payoffs are determined by your or the other participants' decisions.

Your payoff from this part is calculated as follows: At the end of this part, a decision situation is randomly drawn. In addition, a fair coin toss determines whether payoffs are determined by your or the other participants' decisions. You will be informed about the chosen decision situation and the corresponding payoffs at the end of the experiment.

Part "Green"

In this part of the experiment, there will be six decisions. In each decision, you choose your most preferred allocation. The other participant faces the same decisions and has exactly same information as you. The payoffs depend of your decisions and on the other participant's decision.

Your payoff from this part is calculated as follows: At the end of this part, a decision situation is randomly drawn. Your payoff is given by the amount you have allocated to yourself, and the amount that the other participant has allocated to you. The payoff of the other participant is given by the amount the other participant has allocated to herself, and the amount that you have allocated to the other participant. You will be informed about the chosen decision situation and the corresponding payoffs at the end of the experiment.

Part "Yellow"

In this part of the experiment, you interact with another participant in six decision situations. In each decision situation, both participants have to make a decision. That is, each participant chooses the option she prefers. The payoffs that are associated with each combination of decisions are displayed in the payoff table.

In the row header there are the options that you can choose. In the column header there are the options that can be chosen by the other participant. The payoffs associated with a particular combination of choices are in the cell, in which the chosen row and column intersect. The entry to the left of the vertical bar within a cell is your payoff. The entry to the right of the vertical bar within a cell is the other participant's payoff.

Your payoff from this part is calculated as follows: A At the end of this part, a decision situation is randomly drawn. Your payoff and the other participant's payoff is determined by the options you and the other participant selected in this decision situations. You will be informed about the chosen decision situation and the corresponding payoffs at the end of the experiment.

B The six matrix games used in MGT

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	170, 170	153, 170	135, 170	118, 170	100, 170
	Option B	170, 153	153, 153	135, 153	118, 153	100, 153
Your options	Option C	170, 135	153, 135	135, 135	118, 135	100, 135
	Option D	170, 118	153, 118	135, 118	118, 118	100, 118
	Option E	170, 100	153, 100	135, 100	118, 100	100, 100

Table 4: Recomposition of task number one into a matrix game.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	100, 100	109, 104	118, 108	126, 111	135, 115
	Option B	104, 109	113, 113	122, 117	130, 120	139, 124
Your options	Option C	108, 118	117, 122	126, 126	134, 129	143, 133
	Option D	111, 126	120, 130	129, 134	137, 137	146, 141
	Option E	115, 135	124, 139	133, 143	141, 146	150, 150

Table 5: Recomposition of task number two into a matrix game.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	150, 150	146, 159	143, 168	139, 176	135, 185
	Option B	159, 146	155, 155	152, 164	148, 172	144, 181
Your options	Option C	168, 143	164, 152	161, 161	157, 169	153, 178
	Option D	176, 139	172, 148	169, 157	165, 165	161, 174
	Option E	185, 135	181, 144	178, 153	174, 161	170, 170

Table 6: Recomposition of task number three into a matrix game.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	150, 150	129, 159	108, 168	86, 176	65, 185
	Option B	159, 129	138, 138	117, 147	95, 155	74, 164
Your options	Option C	168, 108	147, 117	126, 126	104, 134	83, 143
	Option D	176, 86	155, 95	134, 104	112, 112	91, 121
	Option E	185, 65	164, 74	143, 83	121, 91	100, 100

Table 7: Recomposition of task number four into a matrix game.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	150, 150	163, 138	175, 125	188, 113	200, 100
	Option B	138, 163	151, 151	163, 138	176, 126	188, 113
Your options	Option C	125, 175	138, 163	150, 150	163, 138	175, 125
	Option D	113, 188	126, 176	138, 163	151, 151	163, 138
	Option E	100, 200	113, 188	125, 175	138, 163	150, 150

Table 8: Recomposition of task number five into a matrix game.

		The other's options				
		Option A	Option B	Option C	Option D	Option E
	Option A	150, 150	159, 146	168, 143	176, 139	185, 135
	Option B	146, 159	155, 155	164, 152	172, 148	181, 144
Your options	Option C	143, 168	152, 164	161, 161	169, 157	178, 153
	Option D	139, 176	148, 172	157, 169	165, 165	174, 161
	Option E	135, 185	144, 181	153, 178	161, 174	170, 170

Table 9: Recomposition of task number six into a matrix game.

C Histograms

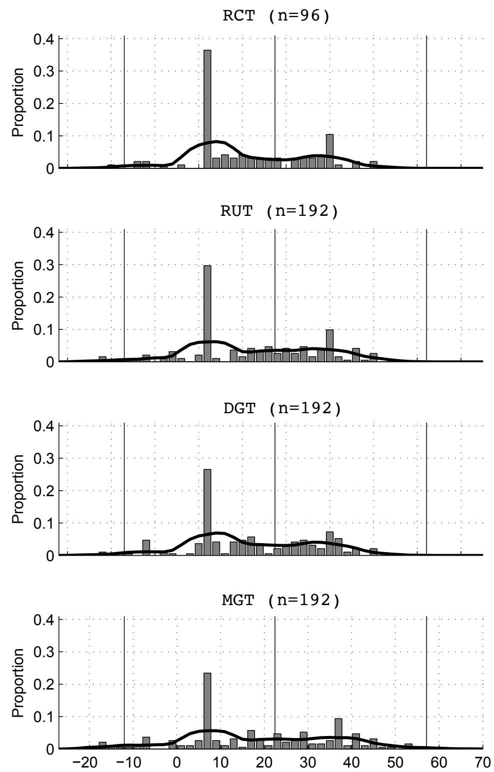

Figure 6: SVO Histograms.

D Tests for Order Effects

In general, there is no systematic relation between SVO angles and the order of when they were assessed (r= -0.03 , p= 0.33).

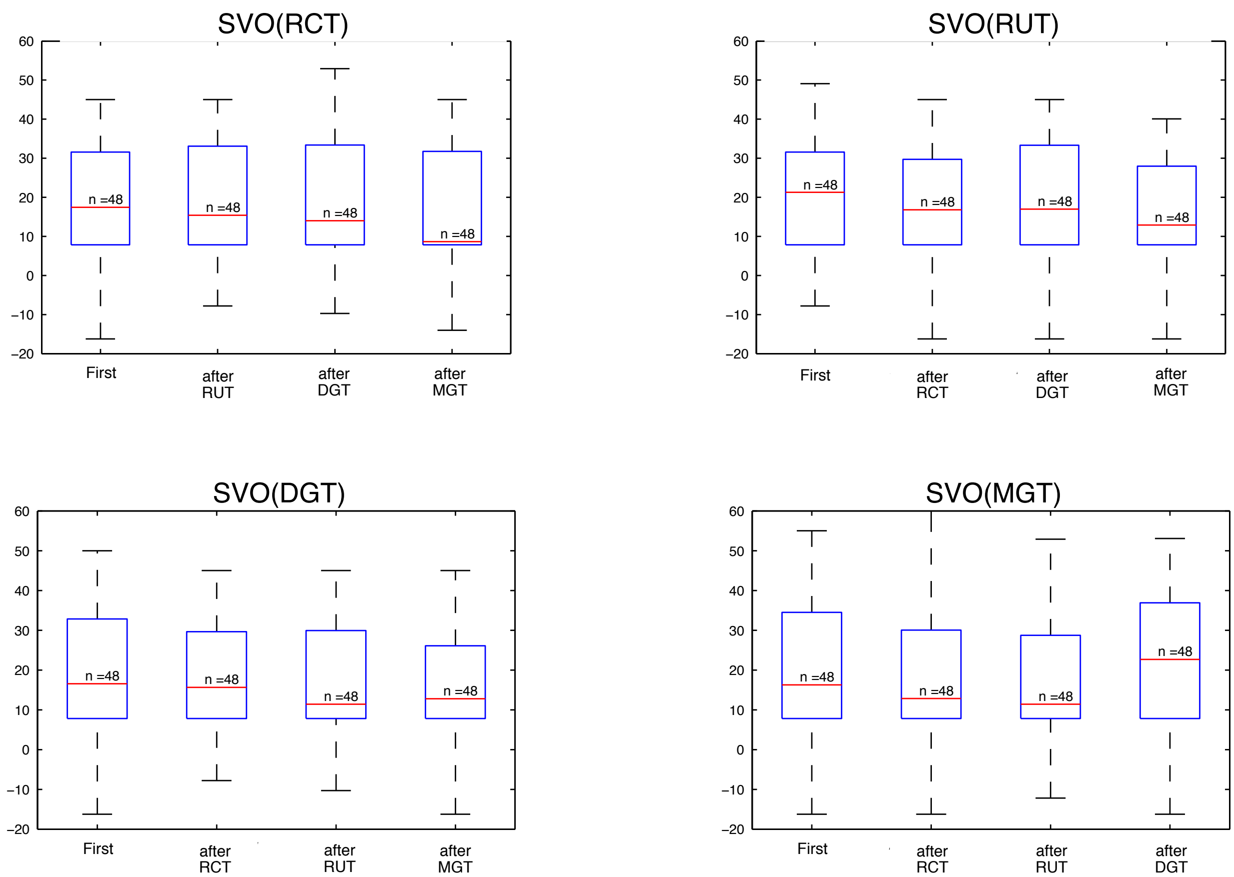

Figure 7: Descriptives of SVO angles per treatment depending on which other treatment was encountered before.