



Status Characteristics and the Provision of Public Goods: Experimental Evidence

Andreas Tutić, Sascha Grehl

Leipzig University

Abstract: We present experimental evidence on the effects of status characteristics in problems involving the provision of public goods. According to Status Characteristics Theory (SCT), status differentials affect performance expectations, which in turn affect the power and prestige order in group tasks. Applied to problems of collective action, SCT suggests several intriguing hypotheses (cf. Simpson, Willer, and Ridgeway 2012). Most importantly, the theory proposes that high-status actors show a greater initiative in and also overall contribute more to the provision of public goods than low-status actors. We put this theoretical claim to a strict experimental test, in addition to other hypotheses and conjectures. In our experimental setup, the volunteer's timing dilemma is used as the group task. Three experimental conditions are implemented, which differ with respect to the way status groups are formed on basis of the type of status characteristic. Our results validate the central hypothesis cited above and also lend support to a conjecture regarding the beneficial effects of heterogeneity in status.

Keywords: volunteer's timing dilemma; social status; public goods; small-group research

THE question of if and to what extent status groups differ in prosocial behavior has a long tradition in the sociological literature (e.g., Mauss [1925] 1954; Homans 1974; Coleman 1990). Because recent advances in experimental social science provide more nuanced tools to differentiate and measure diverse forms of prosocial behavior such as altruistic giving or reciprocating fairness, the status–prosociality nexus has received renewed attention (e.g., Simpson and Willer 2015). Using a wide variety of measures of social status and focusing on plenty of theoretical mechanisms linking social status to prosocial behavior, studies find for instance that status matters regarding donations in the dictator game (e.g., Liebe and Tutić 2010; Piff et al. 2010) as well as the placement of trust in the trust game (e.g., Piff et al. 2010). However, the impact of status hierarchies on one particular core dimension of prosocial behavior—that is, collective action and the voluntary provision of public goods (Olson 1965; Ostrom 1990; Heckathorn 1996)—has so far eschewed extensive experimental inquiry.

It has been just recently that Simpson, Willer, and Ridgeway (2012) elucidated how Status Characteristics Theory (SCT) offers a suitable account for theorizing about the effects of status differentials on collective action. SCT posits that diffuse status characteristics such as gender or ethnicity as well as specific characteristics such as reading ability or mathematical intelligence affect behavior in group tasks via the formation of performance expectations (cf. Berger et al. 1977; Berger, Rosenholtz, and Zelditch 1980; Berger, Wagner, and Zelditch 1985). Simpson et al. (2012) argue that SCT applies to problems of collective action and derive the following three hypotheses: (1) high-status actors take more initiative in contributing towards

Citation: Tutić, Andreas, and Sascha Grehl. 2018. "Status Characteristics and the Provision of Public Goods: Experimental Evidence" *Sociological Science* 5: 1-20.

Received: October 30, 2017

Accepted: November 24, 2017

Published: January 4, 2018

Editor(s): Jesper Sørensen, Gabriel Rossman

DOI: 10.15195/v5.a1

Copyright: © 2018 The Author(s). This open-access article has been published under a Creative Commons Attribution License, which allows unrestricted use, distribution and reproduction, in any form, as long as the original author and source have been credited.

the provision of public goods than low-status actors, (2) high-status actors contribute more towards the provision of public goods than low-status actors, and (3) low-status actors are more eager to match the contributions of high-status actors than vice versa. In addition, Simpson et al. (2012) conjecture that heterogeneity in terms of status characteristics might benefit groups facing problems of collective action since it helps to overcome start-up and free-riding problems.

In this article, we provide experimental evidence on predictions derived from SCT regarding the provision of collective goods. Specifically, we put hypotheses (1) and (2) as well as the conjecture by Simpson et al. (2012) to the test. In our experimental setup, subjects are teamed up in groups of four and confronted with a modified version of the volunteer's timing dilemma (cf. Weesie 1993; Otsubo and Rapoport 2008), in which at least two players have to bear a private cost in order to secure the provision of a public good. Each group is composed of two high-status (so-called stars) and two low-status actors (so-called nonstars). There are three treatments that differ in the way of how subjects are allocated to status groups (i.e. the treatments differ regarding the source of status). In the random treatment, subjects are randomly assigned to the star or nonstar group and, crucially, are informed about the randomness of assignments. In the diffuse treatment, subjects are allocated to status groups based on their subjective social status as measured by a modification of the MacArthur Scale (cf. Adler et al. 2000). In the specific treatment, subjects are told that they have been assigned to status groups based on their performance in a quiz regarding basic understanding of game-theoretic concepts. Yet in fact, they are randomly assigned to status groups.

We find no behavioral differences among status groups in the random treatment. Yet once status groups are composed or allegedly composed based on differences in diffuse or specific status characteristics, high-status actors do show more initiative and do overall contribute more to the provision of public goods than low-status actors. These findings are very much in line with SCT. Also, as conjectured by Simpson et al. (2012), heterogeneity in status characteristics does benefit groups in our experiment in terms of rates of group success in the provision of public goods as well as experimental earnings.

This article extends our knowledge on collective action problems and SCT in three important ways. First, although there is considerable experimental evidence for the validity of core principles of SCT, the larger portion of these findings has been derived from experiments that followed a standardized protocol. This experimental setting involves in a sense fictitious specific characteristics such as contrast sensitivity, meaning-insight ability, relational ability, et cetera as well as specific tasks such as judging the relative size of geometrical figures (Berger et al. 1977: 43–48). Clearly, providing evidence on SCT in other and—from a social-theoretical point of view—more relevant areas of interactive decision-making such as problems of collective action is valuable per se. Second, this is the only study that allows comparing the behavioral effects of different sources of status within a fixed experimental protocol in the context of the voluntary provision of public goods. That is, hitherto scholars have studied the effects of either diffuse status characteristics such as gender (Sell 1997) and scholastic experience (Simpson et al. 2012) or specific characteristics such as performance in general knowledge quizzes (Kumru and

Vesterlund 2010). The question of whether different sources of status affect cooperative behavior in collective action problems in a similar way has consequently eschewed experimental investigation. Hence, our study provides a more reliable basis regarding the extent to which results from prior studies working with diffuse characteristics generalize to specific characteristics and vice versa. Third, our study extends experimental evidence regarding the effects of status characteristics in the voluntary provision of public goods. Hitherto, only “preliminary evidence” for hypothesis (1) has been provided by Simpson et al. (2012: 157), and hypothesis (2) has only been tested and confirmed by Kumru and Vesterlund (2010) and Sell (1997). Most importantly, we provide the only experimental test of the conjecture that heterogeneity in terms of status characteristics might benefit groups facing problems of collective action.

The remainder of this article is organized as follows: In the second section we give a short introduction to SCT. The third section contains a description of our experimental design and procedure as well as testable predictions in empirical terms. In the fourth section we present our experimental findings, and the fifth section discusses possible shortcomings of our approach.

Theory

Put briefly, the core argument of SCT states that whenever a group of actors faces a group task, any status characteristic that is not explicitly dissociated from the group task serves as a basis for performance expectations that in turn affect the power and prestige order of the group (cf. Figure 1)

This statement involves a number of technical concepts and demands some clarification. First of all, any task a group of actors faces qualifies as a group task as long as two scope conditions are met. That is, actors have to be oriented towards success in the task and share a collective orientation in the sense that each actor takes the other actors’ potential performance into account when deciding about her own performance level.

Secondly, the term status characteristic refers to socially significant attributes of actors with differing states, attached to which there are culturally held evaluations and expectations of competence (Correll and Ridgeway 2003: 32). For simplicity, SCT works with binary states, corresponding binary evaluations, and binary expectations of competence—for example, graduate student (positive, competent) and undergraduate student (negative, incompetent). Although all kinds of status characteristics attach evaluations to states, status characteristics differ regarding the range of activities and areas to which the expectations of competence apply. A status characteristic is termed diffuse if these expectations of competence are rather general and not restricted to specific areas. A characteristic is called specific if the expectations of competence are limited to specific areas of expertise. For instance, in many cultures females are believed to lack competence in a wide range of areas such as driving a vehicle, solving mathematical problems, and athletic performance. Hence, gender is a diffuse status characteristic in such cultural environments. In comparison, the attribute “computer expertise” carries far less implications re-

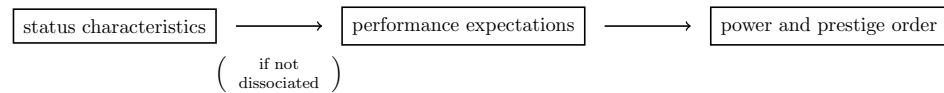


Figure 1: Graphic representation of the core argument of SCT.

garding the areas of competence and hence is an example for a specific status characteristic in most cultures (Correll and Ridgeway 2003: 32).

Thirdly, given the fact that actors want to succeed in the group task and take the behavior of others into account, they search for information regarding the relative competence in performing the group task and also the probable actions of the involved actors and organize them coherently into performance expectations.

Fourthly, the term power and prestige order refers to an important stylized fact that emerged in early research on small groups (cf. Bales 1953; Bales and Slater 1955). In small groups working on a task, the chances to perform, the initiatives to perform, the performance output, as well as the evaluations and rewards for performance are generally highly intercorrelated.

SCT is more concerned with the first arrow in Figure 1, which depicts the causal effect of status characteristics on performance expectations, than with the second arrow, which refers to the causal effect of expectations on behavioral outcomes. Practitioners of SCT (cf. Berger et al. 1977; Simpson and Walker 2002; Simpson et al. 2012) use involved graph-theoretic representations and additional theoretical concepts such as generalized expectation states or abstract task ability to study the details of the causal effect of statuses on expectations (see the online supplement). In contrast, SCT takes a quite simplistic standpoint regarding the second part of the causal chain depicted in Figure 1 (Correll and Ridgeway 2003: 31): “Once developed, performance expectation states (hereafter ‘performance expectations’) shape behavior in a self-fulfilling fashion.”

Irrespective of the technical subtleties of the graph-theoretic account of SCT, the causal chain depicted in Figure 1 implicates that actors who carry positively evaluated states of a diffuse or specific status characteristic obtain a higher position in the power and prestige order, and hence, among other things, show a greater initiative to perform and a higher output in performance towards completion of the group task. Importantly, SCT only theorizes on one particular mechanism by which status characteristics influence behavior in group tasks. That is, SCT is concerned with pure effects of expectation but does not rule out that status characteristics might influence behavior in group tasks via other mechanisms than the formation of performance expectations (cf. Driskell and Mullen 1990; Berger et al. 1977).

Note that perhaps the most interesting aspect of SCT is the fact that the theory predicts effects of statuses on behavior in group tasks that *prima facie* seem completely unrelated to any instrumental task ability involved in performing the task. That is, as long as a status is (as part of the presentation of the group task) not explicitly dissociated from the group task, actors will use this status as a basis for forming performance expectations.

Experimental Design and Procedure

Description

We conducted 12 experimental sessions in spring 2015 for which we recruited a total number of $N = 176$ subjects. Participants were enrolled as university students and recruited through the internet recruitment tool hroot (Bock, Nicklisch, and Baetge 2012). Each session lasted for about one hour and average experimental earnings amounted to 18.93 Euros, which is about twice as much as the average hourly wage for a student in this region. Each session was identical except for treatment conditions. As already indicated, there were three treatment conditions: random ($N = 60$), specific ($N = 60$), and diffuse ($N = 56$). What follows is an overview of the course of the sessions; details of each specific point can be found below.

Once all of the participants arrived, they were asked to complete a questionnaire that among other things measured their respective subjective social status. This measure was used as the diffuse status characteristic in the respective treatment. After completing the questionnaire, subjects took part in a quiz that (allegedly) served later as the specific status characteristic in the respective treatment. Then status groups were formed depending on the treatment conditions. Afterwards, subjects played the volunteer's timing dilemma. A final questionnaire completed the experiment before subjects were paid their experimental earnings.

Beginning of a session. At the beginning of the experiment, subjects were randomly seated at computer terminals within the lab.¹ Because prospect theory (Kahneman and Tversky 1979, 1992) as well as ample experimental evidence (cf. Kahneman, Knetsch, and Thaler 1990) suggest that people generally suffer more from losing a certain amount of money than they enjoy winning the very same amount, we used this so-called endowment effect (Thaler 1980) in order to generate greater incentives with a fixed monetary budget. We therefore endowed subjects with experimental money that they partially lost during the course of the experiment. In order to reinforce this effect, we informed subjects at the start of the experiment that they had earned an amount of 26 Euros, congratulated them for it, and encouraged them to think about what they could buy with that money. However, subjects were also informed that they would probably lose some of the money during the course of the experiment, depending on the decisions made by themselves and by others.

The questionnaire included several questions regarding demographics, such as gender and family situation, as well as questions regarding their educational background and course of study. To obtain the subjective social status of the students, it also included a modified version of the MacArthur Scale of subjective social status (cf. Adler et al. 2000). In contrast to the original MacArthur Scale, which mainly focuses on subjects who are already active on the labor market, our version puts more weight on typical determinants of students' social status, such as network of friends and career prospects. As in the original version, the scale consisted of a 10-rung ladder on which subjects were asked to place themselves. It was accompanied by the following instructions:

Think of this ladder as representing where students stand in Germany. At the top of the ladder are the students who are the best off—those who study the most desired academic subjects at the most respectable universities, obtain the best grades, have the best career prospects, and who are integrated into an attractive network of friends and acquaintances. At the bottom are the students who are the worst off—those who study the least desired academic subjects at the least respectable universities, obtain the worst grades, have no or dim career prospects, and who are isolated or integrated into an unattractive network of friends and acquaintances. The higher you are on this ladder, the closer you are to the students at the very top; the lower you are, the closer you are to the students at the very bottom. Where would you place yourself on this ladder?

Note that this measure of subjective status was used as a diffuse status characteristic in one of our treatments.

Quiz. The quiz was presented to the subjects as a “test of analytic skills in decision situations” and consisted of 15 questions concerning basic game-theoretical concepts such as dominance or best-response (see the online supplement for more details). These concepts were explained via examples in advance. Subjects were told that “the concepts presented are of highly practical use for yourself and others” and “your answers to these questions are pivotal for the future course of the study.” The aspired benefit of using these statements is twofold: On the one hand, these statements are supposed to underline the importance of the quiz and hence increase the legitimacy of using the quiz to form status groups in the respective treatment. On the other hand, we wanted to amplify the illusion that the points that were credited after the quiz and the number of correct answers were identical. In fact, points achieved in this quiz were totally random and had no relation to the number of correct answers.² Note that quiz points serve as a specific status characteristic in one of our treatments.

Status assignment ceremony. Following the quiz, we conducted the status assignment ceremony. Subjects were split up into two groups of equal number, either called stars (high status) or nonstars (low status). The basis for the split was publicly announced. In the random treatment, each subject drew an envelope containing a note stating that the participant was either a star or a nonstar. In the diffuse treatment, subjects were split according to their self-assessment on the modified MacArthur Scale. Therefore, subjects holding a subjective status that was above the median were assigned to the group of stars, whereas subjects that were below the median were declared to be nonstars. Subjects whose self-assessment equaled the median were randomly distributed to either group (in a manner that both groups would be equally sized). In the specific treatment, the points achieved in the quiz were used to determine group membership while following the same principles as in the diffuse treatment.

Once every subject had been informed of which group they belonged to, it was announced that the stars would be seated in a more comfortable place and would be served free soft drinks and chocolates, whereas nonstars would stay at their original places and get nothing at all. Stars and nonstars were then asked to temporarily

leave the lab, with stars being the first ones to leave. Outside the lab, both groups received written instructions concerning the volunteer's timing dilemma³ and were, depending on status group membership, either allowed to take a seat (stars) or forced to stand (nonstars). Meanwhile, inside the lab the places were rearranged so that stars would sit on one side of the lab where cabins were of a better material and more spacious as compared to those of nonstars. After the rearrangement, stars (and after them nonstars) were welcomed inside the lab and guided to their new places. As soon as all subjects were seated, drinks and chocolate were offered to the stars.

Volunteer's timing dilemma. Before starting the volunteer's timing dilemma (henceforth: VTD), subjects were asked to answer five control questions on the VTD at their computer terminal. After having selected an answer, the correct answer and an explanation were displayed to the subject, irrespective of whether the subject answered the question correctly or incorrectly. At this point, subjects were encouraged to ask questions at any time. Once the round of control questioning was completed and all upcoming questions were answered, the VTD began.

Subjects played one practice round followed by 15 regular rounds.⁴ At the beginning of each round, groups composed of two stars and two nonstars were randomly assigned anew. Importantly, each subject was aware that her group was composed of two stars and two nonstars. Then a timer started counting down from 60 seconds and caused all subjects to lose 1 cent per second until either the timer reached 0 and all subjects lost an additional 40 cents or until at least two subjects voluntarily bore a private cost of 40 cents each in order to stop the timer and end the round earlier.⁵ Over the course of a round, subjects were able to choose between two options by pressing one of two buttons labeled "Option A" and "Option B," respectively. Whereas choosing option A was equivalent to volunteering, choosing option B corresponded to not volunteering (that is, free riding). Once a subject had chosen an option, this choice was final and could not be reversed. If a subject did not choose any option before the timer had reached 0, option B was automatically chosen. During a round, subjects were completely unaware of the actions of their group members, but once the round was completed (i.e., as soon as two subjects contributed or the timer passed the 60-second mark), they were informed about the other group members' choices and respective payoffs. Because subjects were unaware of the actions of their group members during the actual play of a round and because of the fact that groups were randomly assigned anew each round, our setup does not allow for a test of the third hypothesis by Simpson et al. (2012). A formal display of the payoff function can be found in the online supplement.

Note that each group playing the VTD was composed of two stars and two nonstars. At the same time, two players were required to contribute towards the provision of the public good. The particular design of our study aimed at creating a test scenario, which ensures that our findings cannot be reduced to a pure focal point effect. The latter term refers to the stylized fact that in games involving some form of coordination, any asymmetry among the players may serve as a basis for the formation of expectations and hence behavior. Diekmann and Przepiorka (2016) demonstrated strong focal point effects in the volunteer's dilemma. This suggests that in a setup in which only one contribution is required and one star as well as

three nonstars are involved, we would expect that the star contributes simply due to a focal point effect. Note that in our design, there are two focal points (i.e., both stars or both nonstars contribute). Whereas the focal point argument provides no ground for choosing one outcome over the other, SCT supplies the sharp prediction that the stars contribute.

End of a session. Finally, each subject filled out a second questionnaire, which included a standard MacArthur Scale regarding the social status of the students' family as well as items on the educational level of the subject's parents, general happiness, religious beliefs, attitudes concerning social inequality, and prior knowledge in game theory. Once all the participants completed the questionnaire, they were paid, thanked, and dismissed.

Hypotheses. As already indicated, Simpson et al. (2012) establish that SCT applies to problems involving the voluntary provision of public goods. That is, essential scope conditions of SCT are fulfilled in classic problems of collective action such as n-person prisoner dilemmas (e.g., Hardin 1971; Hamburger 1973), linear public good games (e.g., Croson 2007; Chaudhuri 2011), or volunteer's dilemmas (cf. Diekmann 1985; Przepiorka and Diekmann 2013; Diekmann and Przepiorka 2016). In this kind of interaction situation, it is very reasonable to assume that actors are indeed oriented towards success in generating enough supply of the collective good and that actors indeed take the contributions of other actors into account while reasoning about their own level of contribution.

Because SCT applies to problems of collective action, the theoretical core argument depicted in Figure 1 supplies, among other things, two testable implications: (1) High-status actors take more initiative in contributing towards the provision of public goods than low-status actors. (2) High-status actors contribute more towards the provision of public goods than low-status actors.⁶ In addition, Simpson et al. (2012) conjecture that heterogeneity in terms of status characteristics might benefit groups facing problems of collective action because it helps to overcome startup and free-riding problems. Keep in mind that these predictions only apply to situations in which the status characteristic under consideration is not explicitly dissociated from the group task.

For clarity and future reference, we restate these hypotheses in the context of our experimental setting. The first of our central dependent variables will be called "contribution rate." This variable measures the proportion of subjects relative to the respective status group who de facto contribute (i.e., choose option A). Note that this variable somehow mixes willingness to contribute and initiative to contribute, as subjects who would be willing to contribute but wait too long (i.e., until two other team members contribute) are counted as noncontributors. Still, this variable is very close to overall performance output. Hence, hypothesis (2) by Simpson et al. (2012) translates into the following:

Individual Contribution Hypothesis (ICH):

Stars show a higher rate of contribution than nonstars in the specific and the diffuse treatment.

Because being a star or a nonstar is random in the random treatment and subjects were aware of this fact (i.e., the status characteristic is explicitly dissociated from

the group task), no performance expectations should be attached to either being a star or a no-star. We therefore state the following hypothesis:

Dissociated Contribution Hypothesis (DCH):

Rates of contribution do not differ between stars and nonstars in the random treatment.

Besides contribution rates, we will also study the willingness to contribute at any point over the course of the round in order to operationalize the initiative to contribute (i.e., hypothesis [1] by Simpson et al. 2012). That is, we will look at the pace of contribution:

Individual Initiative Hypothesis (IIH):

Stars contribute faster than nonstars in the specific and the diffuse treatment.

Analogously to DCH, we will formulate a hypothesis regarding initiative in the random treatment:

Dissociated Initiative Hypothesis (DIH):

Stars and nonstars contribute with the same pace in the random treatment.

Finally, the conjecture by Simpson et al. (2012) that status-differentiated groups are more productive will be put to a test. To do so, we will be looking at a measure of efficiency and a measure of effectiveness. First, the average loss of experimental endowments because of lag in production of the public good will be used as a measure of inefficiency. Second, the rate of successfully generated public goods (i.e., the proportion of groups in which at least two people contribute) will be considered as a measure of effectiveness. This provides two final hypotheses:

Group Efficiency Hypothesis (GEcyH):

Average loss of experimental endowments is lower in the specific and the diffuse treatment than in the random treatment.

Group Effectiveness Hypothesis (GEssH):

The rates of successfully provided public goods are higher in the specific and the diffuse treatment than in the random treatment.

Results

Figure 2 shows contribution rates by status group and treatment. We observe that contribution rates between status groups do not differ significantly in the random condition ($p < 0.538$),⁷ which confirms DCH. In stark contrast, the contribution rates of stars significantly exceed the contribution rates of nonstars in both the diffuse as well as in the specific treatment ($p < 0.002$ and $p < 0.018$, respectively), confirming ICH. Interestingly, these differences among treatments can be mainly attributed to an increased contribution rate of stars in both the diffuse and the

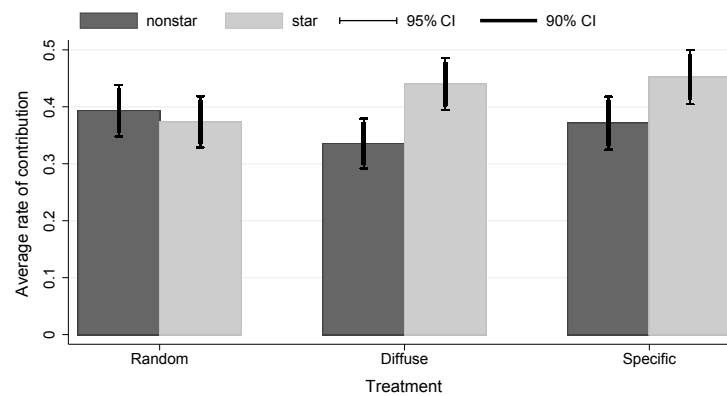


Figure 2: Mean rates of contribution by status group and treatment.

specific treatment compared to the contribution rate of stars in the random condition ($p < 0.043$ and $p < 0.019$, respectively). Although the contribution rate of nonstars differs weakly significantly between the random and diffuse treatment ($p < 0.072$), it does not differ between the random and specific treatment ($p < 0.507$). Note that neither the contribution rates of stars nor the contribution rates of nonstars differ between the diffuse and specific treatment ($p < 0.715$ and $p < 0.270$, respectively).

In order to investigate how the pace of contributions is affected by treatment and status group, we estimated three Kaplan–Meier survival models (see Figure 3). These plots measure the fraction of participants who refrain from contributing over the course of a round. At the beginning of a round (i.e., when the analysis time equals 0), the fraction of subjects who are not contributing equals 1; over the course of a round, some subjects choose to contribute and the fraction decreases. With an increase in the pace of contribution comes a faster decrease in the fraction of noncontributing participants, and therefore the curve drops earlier.⁸ We observe that descriptively, stars contribute slower than nonstars in the random treatment, although this difference is not significant ($p < 0.160$, log-rank test), which confirms DIH. Contrary to this, stars contribute significantly faster than nonstars in the diffuse treatment ($p < 0.001$, log-rank test). Similarly, in the specific treatment, the contribution pace of stars exceeds the contribution pace of nonstars, with weak significance ($p < 0.089$, log-rank test). These findings support IHH. Finally, stars in the diffuse treatment do not show significantly more initiative in contributing towards the public good than stars in the specific treatment; however, nonstars are somewhat more reluctant in the specific treatment ($p < 0.935$ and $p < 0.075$, respectively).

We now turn to the heterogeneity conjecture formulated by Simpson et al. (2012). First, we observe that subjects lose on average approximately 68 cents more in the random treatment than in both other treatments ($p < 0.022$, t-test), supporting GEcyH (see Figure 4 [a]). Also, as visualized in Figure 4 (b), the public good is produced less often in the random treatment than in both other treatments, and this descriptive finding reaches weak statistical significance ($p < 0.088$), which confirms GEssH.

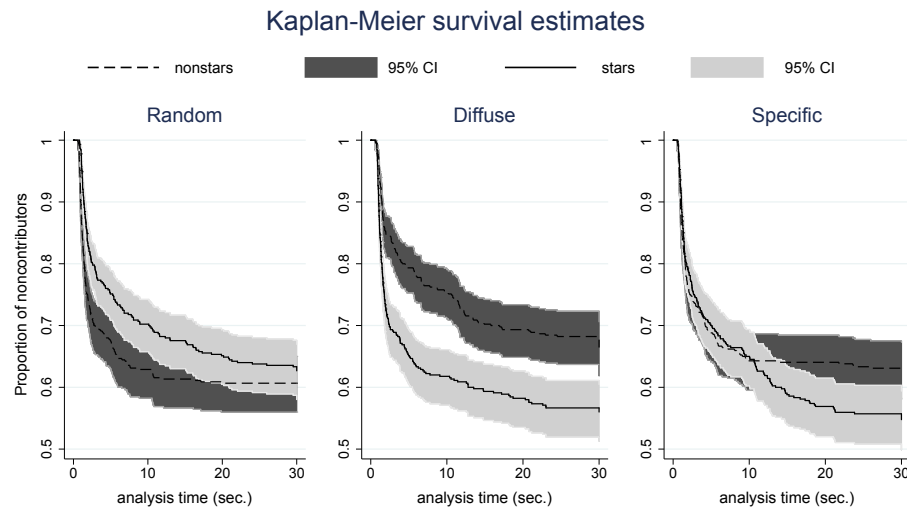


Figure 3: Kaplan–Meier survival estimates of noncontribution by status group and treatment (failure = contribution).

Finally, in order to summarize and consolidate our findings, we estimate six multivariate regressions (see Table 1). In models 1 through 3 we estimate the probability that a subject contributes with three random effects logit regressions.⁹ In models 4 through 6, survival time (i.e., the number of seconds from the start of the round until a subject contributes) is estimated by means of three random effects linear regressions.¹⁰ Models 1 and 4 are the most basic and depict the treatment effect and the effect of being a star in either the random or the nonrandom treatment. Note that this specification and in particular the inclusion of the aggregated treatment effect (variable “Nonrandom treatment?”) allows for a clean statistical test of ICH and IIH (variable “Star in nonrandom treatment?”) as well as DCH and DIH (variable “Star in random treatment?”). Models 2 and 5 add our two intrinsic status characteristics as explanatory variables—that is, subjective social status (as measured via the modified MacArthur Scale) and performance in the quiz (as measured by the number of correctly answered questions). In models 3 and 6 we add control variables. On the one hand, these included the standard demographic variables gender, age, and net income (measured via an ordinal scale).¹¹ On the other hand, we include variables that turned out to have a significant influence on the dependent variables in bivariate correlations.¹²

The variables we chose to include were two dummy variables, one measuring whether the subject had stated that she was trying to maximize her payoff and the other one measuring whether the subject held an academic degree. Adding to that we included a variable depicting the number of semesters that the subject had already been studying for. Unless otherwise specified, independent variables are dichotomous, with the value 0 for no and the value 1 for yes.

Most importantly, we find that stars in a nonrandom treatment have a significantly higher probability to contribute than nonstars in a nonrandom treatment, which confirms ICH. Similarly, they tend to contribute faster towards the public

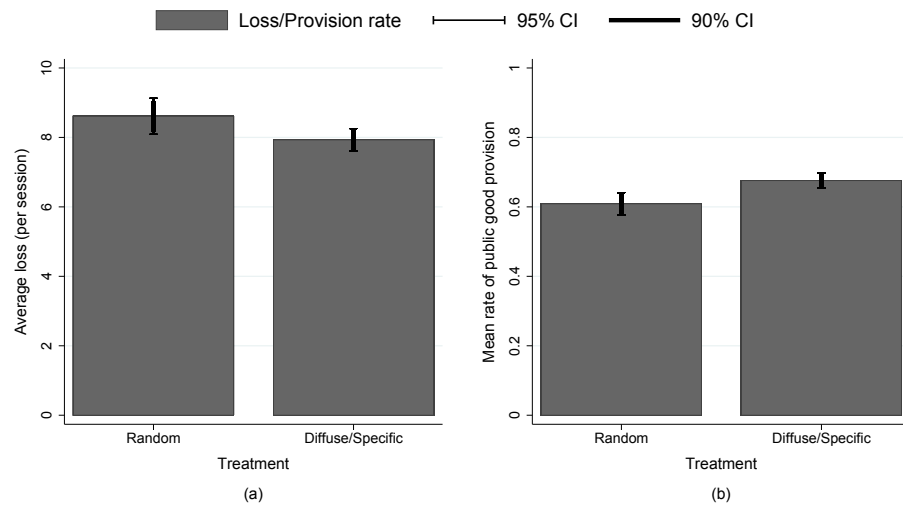


Figure 4: Average loss and mean rate of public good provision by treatment.

good, confirming IIIH. Also, we do not find any behavioral differences between stars and nonstars in the random treatment, which corroborates DCH and DIH. In addition, aggregating over both status groups, we find that neither the probability of contribution nor the pace of cooperation differ between the random and the nonrandom treatment in our study. Please note that none of our hypotheses suggest such an aggregated treatment effect. Notably, the diffuse status characteristic (i.e., subjective social status as measured via the modified MacArthur Scale) has no intrinsic influence on performance besides affecting performance expectations, as models 2 through 3 as well as models 5 through 6 corroborate. Neither the probability to contribute nor the survival time are affected by ratings on the modified MacArthur Scale. Also, the performance in the quiz has no significant influence on any of the two dependent variables.

Models 3 and 6 display that neither gender nor net income have any considerable influence on the contribution rate and pace. One variable that does have a statistically significant influence is age: with an increase in age, the likelihood of contribution increases and the pace of contribution speeds up. Further, we observe that students with more semesters of study and students who already have an academic degree tend to contribute significantly less often. Finally, those subjects who describe themselves as eager to maximize their own material payoffs show a tendency towards free-riding.

Discussion

The article at hand presents experimental evidence on how status characteristics influence performance expectations and hence performance outputs in problems involving the provision of public goods. Our results are clear-cut: we find that positively evaluated status groups contribute more and with a faster pace than

Table 1: Random effects logit (for contribution probability) and random effects linear (for contribution pace) regressions.

	<i>contribution rate (odds ratios)</i>			<i>contribution pace (in seconds)</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Nonrandom treatment?	0.782 (0.225)	0.777 (0.226)	0.640 (0.177)	3.178 (3.142)	3.175 (3.189)	5.156* (3.058)
Star in random treatment?	0.890 (0.293)	0.855 (0.280)	0.664 (0.211)	2.343 (3.607)	2.796 (3.607)	5.349 (3.542)
Star in nonrandom treatment?	1.668 [†] (0.396)	1.675 [†] (0.410)	1.674 [†] (0.380)	-5.228 [†] (2.594)	-5.127* (2.684)	-4.956 [†] (2.511)
Subjective social status (1–10)		0.965 (0.070)	0.968 (0.067)		0.214 (0.800)	0.234 (0.768)
Correct quiz answers (0–15)		0.920 (0.049)	0.928 (0.049)		1.021* (0.590)	0.890 (0.591)
Male?			1.042 (0.205)			-0.215 (2.182)
Net income (1–6)			0.965 (0.105)			0.299 (1.208)
Age (years)			1.122 [†] (0.035)			-1.151 [†] (0.346)
Semesters of study			0.895 [†] (0.032)			1.139 [†] (0.387)
Academic degree?			0.326 [†] (0.073)			12.299 [†] (2.453)
Try to maximize payoff?			0.465 [†] (0.133)			9.034 [†] (3.187)
Constant	0.585 [†] (0.136)	1.613 (1.082)	0.730 (0.734)	38.08 [†] (2.551)	27.30 [†] (7.446)	36.62 [†] (11.035)
Observations	2,640	2,640	2,535	2,640	2,640	2,535

Note: standard errors in parentheses.

* $p < 0.1$; [†] $p < 0.05$.

negatively evaluated status groups in volunteer's timing dilemmas. This supports two of the most important implications of applying SCT to problems of collective action. We also find positive evidence for an interesting conjecture by Simpson et al. (2012), according to which status differentials might have a positive effect on group success in problems of collective action. In our setting, this holds true with respect to experimental earnings as well as the proportion of groups succeeding in producing the public good.

Compared to previous experimental research on the workings of status characteristics in collective action problems, one distinct advantage of our design lies in the comparison of the effects of diffuse and specific status characteristics within a fixed experimental protocol. Prima facie, Figures 2 and 3 suggest that descriptively, status groups show more pronounced behavioral differences if based on a diffuse

than on a specific characteristic. However, closer inspection revealed that in terms of statistical significance, the source of status seems to have a very subtle effect. Comparing status groups across treatments, we found that neither the rate nor the pace of contributions differ between the diffuse and the specific treatment, with the sole exception that nonstars show somewhat less initiative in contributing towards the public good if this status derives from a specific characteristic than if the status stems from a diffuse characteristic. This finding suggests that the current practice of working with either diffuse or specific status characteristics in experimental research on SCT is rather unproblematic. On a side note, this finding also sheds light on a longstanding controversy regarding the relative strength of effects of diffuse and specific statuses on the power and prestige order. Whereas the original SCT states that diffuse and specific characteristics should carry equal weight in determining the power and prestige order, Simpson and Walker (2002) argued that a conceptually more consistent reformulation of SCT lends support to the prediction that diffuse characteristics have a greater impact than specific characteristics. Because we observe a subtle yet definitive greater impact of diffuse characteristics on the pace of contributions, our results favor the reformulated version by Simpson and Walker (2002).

A second advantage of our design is that it provides convincing evidence on the mechanism by which status characteristics influence performance outputs. That is, status characteristics influence performance outputs via forming the basis of performance expectations. Three observations from our study support this claim. First, our specific treatment documents pure effects of performance expectations by separating assignment to status groups from any other property of the actors by randomization. Hence, all observed differences in performance outputs between the status groups are purely due to differential performance expectations. Second, in the random treatment we do not observe any differences in performance outputs, which is very well in line with SCT because being a star or a nonstar is explicitly dissociated from the group task in this treatment, and hence status provides no basis for performance expectations. Finally, although our observations in the diffuse treatment are more questionable on methodological grounds, they still are informative regarding pure effects of differential performance expectations. Methodologically, in the diffuse treatment we encounter the problem that any property of the subjects that is correlated with the modified MacArthur Scale of subjective social status might bias the effects of performance expectations. However, as models 2 and 3 as well as 5 and 6 indicate, subjective social status and observable behavior in the VTD are de facto uncorrelated in our data (see additional analyses in the online supplement).

Though our results speak to the power of status characteristics in influencing performance outputs in group tasks, there is an important caveat in generalizing the results. That is, our design used an extensive award ceremony and differential treatment of subjects (chocolate, lemonades, etc.) when forming the status groups. In terms of SCT, this procedure has a twofold effect. First, it ensures that the status characteristics that serve as the basis of discrimination between status groups become salient. And second, being a star as well as getting drinks and chocolate is intrinsically rewarding. As argued by an extension of SCT to reward expectations

and structures (Berger et al. 1985; Webster and Hysom 1998), actors generally expect that high-status actors obtain greater rewards than low-status actors. Hence, providing these benefits to stars but not to nonstars should theoretically bolster the differential performance expectations; Hysom (2009) corroborates this reasoning with experimental evidence. Of course, we implemented this feature of our design to ensure that subjects got the impression that at least during the duration of the experiment, being a star or a nonstar had some meaning besides being a mere label displayed on a computer screen. Because this feature was present in all three experimental conditions, we have no way of isolating or estimating its effects. Hence, this leaves open the question of to what extent our results generalize to status characteristics outside the lab that come without such award ceremonies and differential treatment. Although this marks a limitation of our study, it has to be kept in mind that many established status characteristics such as gender, ethnicity, and age do de facto come with differential treatment and rewards in our societies. For instance, wages and income (i.e., the most important form of material rewards) generally rise with age (cf. Hellerstein, Neumark, and Troske 1999; Hall and Farkas 2008), and discrimination of members of ethnic minorities (cf. Pager, Bonikowski, and Western 2009; Blommaert, Coenders, and van Tubergen 2013; Bursell 2014) as well as the gender pay gap (cf. Aisenbrey and Brückner 2008; Auspurg, Hinz, and Sauer 2017) are realities on labor markets. In addition, the differential treatment of status groups is often part of our everyday culture and subject to strong and sanctioned norms of conduct. To cite an illuminating facet, studies have shown that teachers judge as well as respond differently to performance outputs by boys and girls during class (cf. Jones and Wheatley 1990; Tiedemann 2002). In this sense, the award ceremony and differential treatment of our subjects just mimics fundamental aspects of social reality.

Although the results of existing experimental research look promising regarding the empirical validity of SCT applied to problems of collective action, the data basis still is rather slim and the question of to what extent these findings are reliable remains. Hence, future research on the application of SCT to problems of collective action should address the question of whether these observations are robust regarding variations in essential parameters of our design. These variations should cover different types of public goods (i.e., different “technologies” that translate inputs into outputs; cf. Sandler 1992) as well as different diffuse and specific status characteristics.

Notes

- 1 The whole experiment was programmed and conducted with the software z-Tree (Fischbacher 2007).
- 2 All subjects in the specific treatments were debriefed after the completion of the study.
- 3 Although the text was the same for all subjects, instructions of stars were decorated with an additional golden star.
- 4 In order to avoid endgame effects, subjects were not informed about how many rounds they would play.

- 5 Strictly speaking, in a volunteer's dilemma it is only necessary for one participant to cooperate. In our experiment we use a generalization of the volunteer's dilemma in which two volunteers are required. The volunteer's dilemma is a special case of a step-level public good game. The characterizing feature of a step-level public good is that it is provided if and only if a certain threshold of contributions towards the public good is met. Failure in meeting the threshold results in no or an inefficient provision of the public good. Also, contributions in excess of the threshold are inefficient. Real-world examples for step-level goods are, for instance, fundraising for community projects such as building a bridge or a fence (cf. Andreoni 1998).
- 6 The reader might wonder why we are interested in the initiative to contribute on top of total contributions. This interest is related to the fact that in many real-life public good problems, there is a certain time limit for the provision of the public good or the value of the public good degenerates over time. For instance, if a person is in urgent need of medical help because of an emergency, her condition might deteriorate until help is provided, with possibly irreversible effects. Note that the VTD models a situation in which the value of contributions towards the provision of the public good shrinks over time. Additionally, Simpson et al. (2012) relates the initiative to contribute to the so-called "start-up problem" in collective action. That is, many collective action problems are characterized by a "critical mass" incentive structure, in which it is beneficial to contribute towards the provision of the public good provided that enough contributions by other actors are made (Oliver, Marwell, and Teixeira 1985). In situations like these, the initiative to make contributions is vital for the provision of public goods because initial contributors might trigger contributions by other actors.
- 7 Unless otherwise specified, χ^2 tests are used.
- 8 Because noncontribution rates only change very marginally after 30 seconds, the plots are truncated, and the final drops at 30 seconds represent the change of the noncontribution rates at the end of the whole round (i.e., after 60 seconds).
- 9 Additionally, we ran clustered logit regressions, which yielded the same results.
- 10 Additionally, we ran clustered Cox regressions as well as several random effects parametric survival models; these models provided the very same results (including exponential, lognormal, and Weibull survival distribution). Also, using logarithmic survival times as the dependent variable in models 4 through 6 does not affect our findings.
- 11 Specifically, net income was measured in steps of 300 Euros, i.e., 0 to 300, 301 to 600, etc.
- 12 Because we lack theoretical reasons to include specific control variables, we opted to approach the task of selecting control variables on purely empirical grounds. Note that because of randomization, controlling for confounds is pointless with regard to the research interest of this article. Still, the reported effects might be interesting per se.

References

- Adler, Nancy E., Elissa S. Epel, Grace Castellazzo, and Jeannette R. Ickovics. 2000. "Relationship of Subjective and Objective Social Status with Psychological and Physiological Functioning: Preliminary Data in Healthy, White Women." *Health Psychology* 19:586–92. <https://doi.org/10.1037/0278-6133.19.6.586>.
- Aisenbrey, Silke, and Hannah Brückner. 2008. "Occupational Aspirations and the Gender Gap in Wages." *European Sociological Review* 24:633–49. <https://doi.org/10.1093/esr/jcn024>.

- Andreoni, James. 1998. "Toward a Theory of Charitable Fund-Raising." *Journal of Political Economy* 106:1186–213. <https://doi.org/10.1086/250044>.
- Auspurg, Katrin, Thomas Hinz, and Carsten Sauer. 2017. "Why Should Women Get Less? Evidence on the Gender Pay Gap from Multifactorial Survey Experiments." *American Sociological Review* 82:179–210. <https://doi.org/10.1177/0003122416683393>.
- Bales, Robert F. 1953. "The Equilibrium Problem in Small Groups." In *Working Papers in the Theory of Action*, edited by Talcott Parsons, Robert F. Bales, and E. A. Shils, pp. 111–61. Glencoe: Free Press.
- Bales, Robert F., and Philip E. Slater. 1955. "Role Differentiation in Small Decision Making Groups." In *Family, Socialization, and Interaction Process*, edited by Talcott Parsons and Robert F. Bales, pp. 259–306. Glencoe: Free Press.
- Berger, Joseph, M. Hamit Fisek, Robert Zane Norman, and Morris Zelditch, Jr. 1977. *Status Characteristics and Social Interaction: An Expectation-States Approach*. New York: Elsevier.
- Berger, Joseph, Susan J. Rosenholtz, and Morris Zelditch. 1980. "Status Organizing Processes." *Annual Review of Sociology* 6:479–508. <https://doi.org/10.1146/annurev.so.06.080180.002403>.
- Berger, Joseph, David G. Wagner, and Morris Zelditch, Jr. 1985. "Introduction: Expectation States Theory - Review and Assessment." In *Status, Rewards, and Influence: How Expectations Organize Behaviour*, edited by Joseph Berger and Morris Zelditch, Jr., pp. 1–72. San Francisco: Jossey-Bass.
- Blommaert, Lieselotte, Marcel Coenders, and Frank van Tubergen. 2013. "Discrimination of Arabic-Named Applicants in the Netherlands: An Internet-Based Field Experiment Examining Different Phases in Online Recruitment Procedures." *Social Forces* 92:957–82. <https://doi.org/10.1093/sf/sot124>.
- Bock, Olag, Andreas Nicklisch, and Ingmar Baetge. 2012. "hroot: Hamburg Recruitment and Organization Online Tool." *WiSo-HH Working Paper Series No. 1*.
- Bursell, Moa. 2014. "The Multiple Burdens of Foreign-Named Men – Evidence from a Field Experiment on Gendered Ethnic Hiring Discrimination in Sweden." *European Sociological Review* 30:399–409. <https://doi.org/10.1093/esr/jcu047>.
- Chaudhuri, Ananish. 2011. "Sustaining Cooperation in Laboratory Public Goods Experiments: A Selective Survey of the Literature." *Experimental Economics* 14:47–83. <https://doi.org/10.1007/s10683-010-9257-1>.
- Coleman, James Samuel. 1990. *Foundations of Social Theory*. Cambridge, Mass: Belknap Press.
- Correll, Shelley J., and Cecilia L. Ridgeway. 2003. "Expectation States Theory." In *Handbook of Social Psychology*, edited by John Delamater, pp. 29–53. New York: Kluwer Academic/Plenum Publishers.
- Croson, Rachel T. A. 2007. "Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games." *Economic Inquiry* 45:199–216. <https://doi.org/10.1111/j.1465-7295.2006.00006.x>.
- Diekmann, Andreas. 1985. "Volunteer's Dilemma." *Journal of Conflict Resolution* 29:605–10. <https://doi.org/10.1177/0022002785029004003>.

- Diekmann, Andreas, and Wojtek Przepiorka. 2016. "Take One for the Team! Individual Heterogeneity and the Emergence of Latent Norms in a Volunteer's Dilemma." *Social Forces* 94:1309–333. <https://doi.org/10.1093/sf/sov107>.
- Driskell, James E., and Brian Mullen. 1990. "Status, Expectations, and Behavior: A Meta-Analytic Review and Test of the Theory." *Personality and Social Psychology Bulletin* 16:541–53. <https://doi.org/10.1177/0146167290163012>.
- Fischbacher, Urs. 2007. "z-Tree: Zurich Toolbox for Ready-made Economic Experiments." *Experimental Economics* 10:171–78. <https://doi.org/10.1007/s10683-006-9159-4>.
- Hall, Matthew, and George Farkas. 2008. "Does Human Capital Raise Earnings for Immigrants in the Low-Skill Labor Market?" *Demography* 45:619–39. <https://doi.org/10.1353/dem.0.0018>.
- Hamburger, Henry. 1973. "N-Person Prisoner's Dilemma." *Journal of Mathematical Sociology* 3:27–48. <https://doi.org/10.1080/0022250X.1973.9989822>.
- Hardin, Russell. 1971. "Collective Action as an Agreeable N-Prisoners' Dilemma." *Behavioral Science* 16:472–81.
- Heckathorn, Douglas D. 1996. "The Dynamics and Dilemmas of Collective Action." *American Sociological Review* 61:250–77. <https://doi.org/10.2307/2096334>.
- Hellerstein, Judith K., David Neumark, and Kenneth R. Troske. 1999. "Wages, Productivity, and Worker Characteristics: Evidence from Plant-Level Production Functions and Wage Equations." *Journal of Labor Economics* 17:409–46. <https://doi.org/10.1086/209926>.
- Homans, George C. 1974. *Social Behavior: Its Elementary Forms*. Oxford: Harcourt Brace Jovanovich.
- Hysom, Stuart J. 2009. "Status Valued Goal Objects and Performance Expectations." *Social Forces* 87:1623–48. <https://doi.org/10.1353/sof.0.0160>.
- Jones, M. Gail, and Jack Wheatley. 1990. "Gender Differences in Teacher-Student Interactions in Science Classrooms." *Journal of Research in Science Teaching* 27:861–74. <https://doi.org/10.1002/tea.3660270906>.
- Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler. 1990. "Experimental Tests of the Endowment Effect and the Coase Theorem." *Journal of Political Economy* 98:1325–48. <https://doi.org/10.1086/261737>.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–92. <https://doi.org/10.2307/1914185>.
- Kahneman, Daniel, and Amos Tversky. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5:297–323. <https://doi.org/10.1007/BF00122574>.
- Kumru, Cagri S., and Lise Vesterlund. 2010. "The Effect of Status on Charitable Giving." *Journal of Public Economic Theory* 12:709–35. <https://doi.org/10.1111/j.1467-9779.2010.01471.x>.
- Liebe, Ulf, and Andreas Tutić. 2010. "Status Groups and Altruistic Behaviour in Dictator Games." *Rationality and Society* 22:353–80. <https://doi.org/10.1177/1043463110366232>.

- Mauss, Marcel. [1925] 1954. *The Gift: Forms and Functions of Exchange in Primitive Societies*. London: Cohen and West.
- Oliver, Pamela, Gerald Marwell, and Ruy Teixeira. 1985. "A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action." *American Journal of Sociology* 91:522–56. <https://doi.org/10.1086/228313>.
- Olson, Mancur. 1965. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press. <https://doi.org/10.1017/CB09780511807763>.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Political Economy of Institutions and Decisions. Cambridge and New York: Cambridge University Press.
- Otsubo, Hironori, and Amnon Rapoport. 2008. "Dynamic Volunteer's Dilemmas over a Finite Horizon: An Experimental Study." *Journal of Conflict Resolution* 52:961–84.
- Pager, Devah, Bart Bonikowski, and Bruce Western. 2009. "Discrimination in a Low-Wage Labor Market: A Field Experiment." *American Sociological Review* 74:777–99. <https://doi.org/10.1177/000312240907400505>.
- Piff, Paul K., Michael W. Kraus, Stéphane Côté, Bonnie Hayden Cheng, and Dacher Keltner. 2010. "Having Less, Giving More: the Influence of Social Class on Prosocial Behavior." *Journal of Personality and Social Psychology* 99:771–84. <https://doi.org/10.1037/a0020092>.
- Przepiorka, Wojtek, and Andreas Diekmann. 2013. "Individual Heterogeneity and Costly Punishment: A Volunteer's Dilemma." *Proceedings of the Royal Society B* 280:20130247. <https://doi.org/10.1098/rspb.2013.0247>.
- Sandler, Todd. 1992. *Collective Action: Theory and Applications*. Ann Arbor: University of Michigan Press.
- Sell, Jane. 1997. "Gender, Strategies, and Contributions to Public Goods." *Social Psychology Quarterly* 60:252–65. <https://doi.org/10.2307/2787085>.
- Simpson, Brent, and Henry A. Walker. 2002. "Status Characteristics and Performance Expectations: A Reformulation." *Sociological Theory* 20:24–40. <https://doi.org/10.1111/1467-9558.00149>.
- Simpson, Brent, and Robb Willer. 2015. "Beyond Altruism: Sociological Foundations of Cooperation and Prosocial Behavior." *Annual Review of Sociology* 41:43–63. <https://doi.org/10.1146/annurev-soc-073014-112242>.
- Simpson, Brent, Robb Willer, and Cecilia L. Ridgeway. 2012. "Status Hierarchies and the Organization of Collective Action." *Sociological Theory* 30:149–66. <https://doi.org/10.1177/0735275112457912>.
- Thaler, Richard H. 1980. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization* 1:39–60. [https://doi.org/10.1016/0167-2681\(80\)90051-7](https://doi.org/10.1016/0167-2681(80)90051-7).
- Tiedemann, Joachim. 2002. "Teachers' Gender Stereotypes as Determinants of Teacher Perceptions in Elementary School Mathematics." *Educational Studies in Mathematics* 50:49–62. <https://doi.org/10.1023/A:1020518104346>.
- Webster, Murray, and Stuart J. Hysom. 1998. "Creating Status Characteristics." *American Sociological Review* 63:351–78. <https://doi.org/10.2307/2657554>.

Weesie, Jeroen. 1993. "Asymmetry and Timing in the Volunteer's Dilemma." *Journal of Conflict Resolution* 37:569–90. <https://doi.org/10.1177/0022002793037003008>.

Acknowledgements: Financial support by the German Research Foundation (DFG TU 409/1) and research assistance by Maximilian Lutz are gratefully acknowledged.

Andreas Tutić: Institute of Sociology, Leipzig University.

E-mail: andreas.tutic@sozio.uni-leipzig.de.

Sascha Grehl: Institute of Sociology, Leipzig University.

E-mail: sascha.grehl@uni-leipzig.de.