# The B.E. Journal of Economic Analysis & Policy

### Contributions

### The Source and Significance of Confusion in Public Goods Experiments

Paul J. Ferraro<sup>\*</sup> Christian A. Vossler<sup>†</sup>

\*Georgia State University, pferraro@gsu.edu †University of Tennessee, cvossler@utk.edu

#### **Recommended** Citation

Paul J. Ferraro and Christian A. Vossler (2010) "The Source and Significance of Confusion in Public Goods Experiments," *The B.E. Journal of Economic Analysis & Policy*: Vol. 10: Iss. 1 (Contributions), Article 53. Available at: http://www.bepress.com/bejeap/vol10/iss1/art53

Copyright ©2010 The Berkeley Electronic Press. All rights reserved.

### The Source and Significance of Confusion in Public Goods Experiments\*

Paul J. Ferraro and Christian A. Vossler

#### Abstract

Economists use public goods experiments to develop and test theories of individual preferences and institutional design. Previous work demonstrates many participants in public goods experiments contribute to the public good out of confusion. We design experiments to provide insights on the consequences and causes of confusion. We establish that confusion amounts to more than statistical noise and does not dissipate with repetition (i.e. learning). Confused subjects use experimental parameters and the behavior of other players as cues, which confounds treatment effects and traditional strategies to identify other-regarding preferences through exogenous parameter changes and the modeling of reactions to other subjects' decisions. We argue that confusion stems from an inaccurate understanding of game incentives ("failure of game form recognition"), which is a consequence of the framing and inadequate payoff information in standard instructions. Modified instructions can substantially reduce confusion, and, in turn, change the distribution of contributions to the public good.

**KEYWORDS:** public goods, identification, other-regarding preferences, experimental design, confusion

<sup>\*</sup>Research supported by the Georgia State University Office of Sponsored Programs. Thanks to Krawee Ackaramongkolrotn and Steve Cotten for software development, James Cox, Susan Laury and the GSU Research Initiation Grant for funding, Kwaw Andam for data entry, and James Cox, Douglas Davis, Catherine Eckel, Joseph Henrich, Susan Laury, Howard Margolis, Michael Mc-Kee, Ragan Petrie, participants at the 74th Annual Meeting of Southern Economic Association, the 2004 North American Regional Meetings and the 2006 International Meetings of the Economic Science Association and two anonymous referees for useful comments.

#### I. INTRODUCTION

The Voluntary Contributions Mechanism (VCM) game is central to experimental research on the private provision of public goods. Public economists use the VCM to test their theories, behavioralists use the VCM to develop theories of individual preferences, and institutionalists and policy-oriented economists use the VCM to explore how changes in the rules affect collective outcomes.

The typical linear VCM experiment places subjects in a social dilemma. Subjects are given an endowment of "tokens" to be divided between a private exchange and a group exchange. Tokens placed in the group exchange yield a return to each group member, regardless of their own investment level (i.e., the group exchange is a pure public good). If the marginal return from investing a token to the group exchange is less than the value of a token kept in the private exchange, but the sum of the marginal returns to the group is greater than the value of a token kept, the unique Nash equilibrium is for all players to contribute zero tokens to the group exchange. The Pareto-dominant, welfare-maximizing outcome, however, is realized when everyone contributes their entire endowment to the group exchange.

Over 30 years of VCM experiments have resulted in the following stylized facts (Davis and Holt, 1993; Ledyard, 1995; Holt and Laury, 2008):

(SF-1) Average contributions to the public good are a significant portion of endowments.

(a) In single-shot settings, average contributions are between 40-60% of endowments.

(b) In repeated-round settings, average contributions start in the same range, but decline over rounds and remain significantly different from zero in the last round (even in experiments that last 50 or more rounds).

(SF-2) Increases in the Marginal per Capita Return (MPCR) increase contributions.

(SF-3) Increasing group size, at least for low MPCRs and small group sizes, increases contributions.

Standard game theory that assumes purely self-interested players does not explain these stylized facts. Alternative theories and supporting empirical evidence have been offered, but there is no consensus explanation of observed behavior (Ledyard). At least some results are consistent with other-regarding behaviors: warm-glow altruism (e.g., Palfrey and Prisbey, 1997), pure altruism (e.g. Goeree et al., 2002), conditional cooperation (e.g., Fischbacher et al., 2001; Croson et al. 2005), or a combination of these (e.g., Cox and Sadiraj, 2007).<sup>1</sup> The

<sup>&</sup>lt;sup>1</sup> "Pure altruism" exists when an individual's utility function is a function of his own payoff and the payoffs of his group members. "Warm-glow altruism" (also called "impure altruism"; Andreoni, 1990) exists when an individual gains utility from the simple act of contributing to a

decline in contributions over repeated rounds has been attributed to reductions in confusion (Andreoni, 1995; Palfrey and Prisbrey; Houser and Kurzban, 2002; Carpenter, 2004), or to the revocation of conditional cooperation (Fischbacher et al., and others who use the 'strategy method').

This study focuses on the role of confusion. Our main motivation for this mirrors Andreoni (1995), who argues that the results of experiments designed to test theories of social giving are difficult to interpret if confusion is a primary source of cooperation. We use the term "confusion" to characterize behavior that stems from subjects' inability to discern the relationships between the choices made and the game's incentives (what Chou et al., 2009, call "game form recognition" and Plott and Zeller, 2005, call "subject misperceptions"). Confused players include those who are unaware of the opportunity to free-ride on others' contributions in the VCM game. An informed player may choose to forgo this opportunity, but a confused player is unaware of this opportunity. Similar to previous studies, which find as much as fifty percent of contributions stem from confusion (Andreoni, 1995; Palfrey and Prisbrey; Houser and Kurzban, 2002), we find substantial evidence of confusion. By investigating the structure of this confusion, however, we demonstrate that confusion affects internal validity. In particular, confusion confounds treatment effects (e.g. MPCR effects) and attempts to identify the relative importance of various motives in public goods experiments.

In order to adequately investigate the source and significance of confusion behavior in public goods experiments as it pertains to identification of the stylized facts, we conduct four experiments. The experiments are summarized in Table 1. Experiment 1 uses the design of Goeree et al., which is a static experiment that allows one to explore how contributions correspond with changes in the MPCR and group size (i.e. SF-2 and SF-3). We find that confusion confounds strategies to identify other-regarding preferences through the use of exogenous parameter changes. In this experiment, confused subjects behave like pure altruists. Experiments 2 and 3 are static and dynamic experiments, respectively, that relate to SF-1. They shed light on the two motives put forth to explain the standard decay in repeated-round experiments: the revocation of conditional cooperation and learning. Experiment 2 applies the increasingly used 'strategy method' design of Fischbacher et al., which allows one to identify conditional cooperators and other player "types". We demonstrate that many confused individuals in these experiments are erroneously classified as conditional cooperators. Experiment 3 is a standard repeated-round VCM game (e.g. Isaac et al., 1984), which provides an opportunity for learning. In this experiment we find little evidence of learning and

publicly spirited cause. "Conditional cooperation" (also called "strong reciprocity") exists when an individual tries to match or condition her contributions on the contributions of her group members (Keser and van Winden, 2000).

#### Ferraro and Vossler: Confusion in Public Goods Experiments

Experiment	Experimental Design	Instructions	Sample Size	Main Findings
Number	(Environment)			
1	Goeree et al., 2002	Standard	96	50% of contributions stem from confusion;
	(Static)			Half of measured altruism stems from confusion
2	Fischbacher et al., 2001 (Static)	Standard	40	Confused players behave like conditional cooperators
3	Isaac et al., 1984 (Dynamic)	Standard	240	50% of contributions stem from confusion; Confused players behave like conditional cooperators; Little evidence of learning
4	Isaac et al., 1984 (Dynamic)	Modified	88	Alternative framing and a complete payoff matrix substantially reduce confusion and change the distribution of contributions

#### Table 1. Summary of Experiments

more evidence that some confused individuals behave like conditional cooperators. Overall, both Experiment 2 and 3 suggest that confusion confounds strategies to identify other-regarding preferences through the modeling of reactions to other subjects' decisions.

Our results question the internal validity of public goods experiments, in the sense that some stylized facts of public good experiments may be artifacts of the instructions and designs commonly used by experimentalists. To be clear, we are not saying that stylized facts do not exist. Indeed, we still find evidence of other-regarding preferences when we control for confusion. However, we cannot draw inferences for a large proportion of subjects in these studies, and in the absence of confusion, the stylized facts may or may not be considerably different.

Nevertheless, our study presents encouraging results and guidance for researchers. In particular, results from Experiments 1-3 and narratives from a post-experiment focus group indicate that confusion arises mainly from (1) experimenter demand effects related to the "investment" language of traditional instructions and (2) a failure to see the opportunity to free-ride on others' contributions while correctly perceiving that own payoffs increases with others' contributions (i.e., they fail to see the social dilemma of the experimental design). In response these observations, we modified standard repeated-round VCM instructions in Experiment 4. Our modified instructions use associative framing and a complete payoff matrix, which we find substantially decreases confusion as well as changes the distribution of contributions. These modified VCM instructions offer a vehicle for further investigation. Our empirical evidence also provides strong support for our claim that the virtual-player method, which we rely upon to identify confusion behavior, neutralizes other-regarding preferences and self-interested strategic play.

#### **II. THE VIRTUAL-PLAYER METHOD**

The virtual-player method discriminates between confusion and other-regarding behavior in single-round public goods experiments, and discriminates between confusion and other-regarding behavior or self-interested strategic play in repeated-round experiments. The method relies on three important features: (1) the introduction of nonhuman, virtual players (i.e., automata) that are programmed to execute pre-determined contribution sequences made by human players in an otherwise comparable treatment; (2) a split-sample design where each participant is knowingly grouped with either humans (the "all-human treatment") or with virtual players (the "virtual-player treatment"); and (3) a procedure that ensures that human participants understand how the virtual players are programmed. In the repeated-game context, the method also ensures that

subjects in the two treatments see the same history of play by their group "members". Johnson et al. (2002) use virtual players in a sequential bargaining game, but their virtual players are programmed to play a subgame perfect equilibrium strategy rather than as humans have played in previous experiments. Houser and Kurzban (2002) use virtual players in a VCM game, but as we describe in Section V, their design differs in several important ways that affect the inferences they can draw about confusion. Ferraro et al. (2003) also use virtual players, but in a game design that does not allow one to draw inferences about the structure of confusion.

Here is an excerpt from the virtual-player instructions for Experiment 1 (Section III):

"In this experiment you will be asked to make a series of choices about how to allocate a set of tokens. You will be in groups for this experiment. However, you will <u>not</u> be grouped with others in the room. Your group will consist of yourself and "Virtual Players." These Virtual Players are <u>not</u> human and their decisions have already been determined. Your decisions will thus have absolutely <u>no effect</u> on how the Virtual Players behave. To assure you that the decisions of the nonhuman Virtual Players have indeed been determined already and will not change during the experiment, we have envelopes in which the investment decisions of the nonhuman Virtual Players in your group are printed on a piece of paper. We have placed these envelopes on your desk. AFTER the experiment is over, you may open your envelope and confirm that it contains the decisions made by the nonhuman Virtual Players in your group. PLEASE DO NOT OPEN THE ENVELOPE UNTIL THE EXPERIMENT IS COMPLETED."

Our other experiments use similar language (see online appendices for instructions). Using virtual players allows us to use traditional experimental instructions and thus test hypotheses without changing the nature of the game. The instructions emphasize the nonhuman nature of the virtual players and the exact way in which virtual player decisions are programmed. To avoid misunderstandings in payoff calculation examples and practice questions, we are careful to place words such as "earn" between quotes when describing payoffs for the virtual players and then immediately insert further explanation such as, "Of course, because the Virtual Player is not real, it does not actually receive any earnings."

When a subject knowingly plays a linear VCM game matched with predetermined contribution sequences unaffected by her decisions, otherregarding preferences and self-interested strategic play no role in decisions. Subjects gain no utility by being altruistic to pieces of paper in envelopes. Strategic play in a repeated-game has no impact on the predetermined decisions of virtual players. The random assignment of participants to an all-human or a virtual-player group allows the researcher to net out confusion contributions by subtracting contributions by (human) participants in the virtual-player treatment from all-human treatment contributions. In single-round experiments where the decisions of other players are not known *ex ante*, the use of virtual players should have no effect on human contributions nor should they confound any comparison between all-human and virtual-player treatments. Thus, randomly selecting the contribution sequence of any previous human participant, with replacement, as the contribution sequence for a virtual player ensures comparability.

In repeated-round public goods games where group contributions are announced after each period, one must exercise additional control because the history of play may affect contributions. The additional control comes by ensuring that each human in the all-human treatment has a human "twin" in the virtualplayer treatment: each twin sees exactly the same contributions by the other members of his group in each round – the only difference is that the participant in the virtual-player treatment knows he is grouped with pre-programmed contributions sequences. Thus, for example, say subject H1 plays with H2, H3 and H4 in the all-human treatment session. Subject V1 in the virtual-player session plays with 3 virtual players, one programmed with the contributions sequence of human subject H2, one with the sequence of subject H3, and one with the sequence of subject H4. This design ensures that we can treat the individual as the observational unit, rather than just the group, without making the additional assumption that the history of play has no effect on contributions.

Throughout the following analyses, we interpret the contributions in virtual-player treatments as stemming purely from confusion. We realize that the source of contributions under virtual player conditions is open to debate, and we thus have accumulated direct and indirect evidence in support of this interpretation. For ease of exposition, we present this evidence after presenting our main results. In particular, we provide evidence that nearly all subjects understood the nature of the virtual players and they behaved as if profit maximization was their only motive in virtual-player treatments.

#### **III. EXPERIMENT 1: ALTRUISM, MPCR AND GROUP SIZE**

#### **Experimental Design**

In order to detect contributions stemming from pure and warm-glow altruism, Goeree et al. (hereafter GHL) create a variant of the one-shot linear VCM game and empirically model contributions using a quantal response equilibrium model of noisy decision-making. In each of ten decision tasks, each subject is given an endowment of 25 tokens and allocates them between a private and a group exchange. Each decision is made without feedback, and the internal  $(m_1)$  and external rates  $(m_e)$  of return, and group size (n) vary across tasks. For any task, the payoff function for individual *i* is

$$[1] \qquad \pi_i = \omega - y_i + m_{\mathrm{I}} * y_i + m_{\mathrm{e}} * Y_i$$

where  $y_i$  denotes own contributions to the group exchange,  $Y_i$  is contributions from other group members to the group exchange, and  $\omega$  is the endowment. The internal rate of return is the marginal return to oneself from a token contributed to the group exchange. The external rate of return is the marginal return to other players from one's contribution to the group exchange. Group size is either two or four players. The internal rate of return is always lower than the value of a token in the private exchange, and thus subjects have a dominant strategy to allocate zero to the group exchange. Subjects are paid for only one decision, chosen at random *ex post*.

In the typical VCM game, all players receive the same return from the public good (i.e.,  $m_I = m_e$ ). Varying the returns and group size is a strategy to identify pure altruism in their empirical model: participants exhibiting pure altruism should increase their contributions when the external return or the group size increases.<sup>2</sup> If considerable contributions are observed, but they show little correlation with external return and group size, the conjecture is that contributions are largely attributable to warm-glow. We further hypothesize that individuals who are confused about the incentives in the game will believe the variation in the returns and group size is a signal about the payoff-maximizing allocation of tokens.

We create an augmented GHL design, in which we include 15 decision tasks that allow us to generate greater variability in the MPCR and group size. We also increase the financial returns from private and group exchange allocations (GHL earnings are so low that an unrelated experiment was needed to augment earnings). Endowments are 50 tokens per task, private exchange returns are 6 cents/token, internal returns from the public good are 1, 2, 4 or 6 cents, and external returns from the public good are 1, 2, 4, 6 and 12 cents.

Experiment instructions are presented orally and in writing (see online appendix A). The all-human treatment instructions are from GHL, with minor changes. The virtual-player instructions are similar with the exception that they emphasize that participants are matched with pre-determined contributions. As in

<sup>&</sup>lt;sup>2</sup> Palfrey and Prisbey also use exogenous changes in parameters to identify altruism.

GHL, participants make decisions via paper and pencil and must answer a series of practice questions before making their decisions. A post-experiment questionnaire is given to collect basic demographic information as well as to assess understanding of the experimental design and decision tasks. The same author moderated all of the sessions.

GHL assign subjects to two- and four-member groups by selecting marked ping-pong balls after all decisions are made. We pre-assign participants to twoand six-member groups based on their subject ID number. This pre-assignment is important for virtual-player sessions because we give each subject a sealed envelope containing information on the aggregate contributions of the virtual players before the subjects make their decisions.

Ninety-six undergraduate volunteers participated in the computerized experiment at the University of Tennessee Experimental Economics Laboratory. For each treatment, there were four sessions consisting of twelve people who were visually isolated, and who were not aware of the identity of other group members. Four distinct orderings of the 15 decision tasks were used, and for each treatment there was one session associated with each ordering. As in all experiments, no subject had ever participated in a public goods experiment. Earnings averaged \$14.98, and the experiment lasted no more than one hour.

#### Results

Figure 1 presents mean contributions (as a percentage of endowment) in the two treatments for each of the fifteen decision tasks. Looking at the all-human treatment, the pattern of contributions in relation to design factors is quite similar to the GHL study, with contributions increasing with respect to external return, group size, and MPCR. Albeit at a lower level, subjects in the virtual-player treatment alter their contributions based on the same stimuli as subjects in the all-human treatment: the two response patterns are parallel. Contributions are nearly equal only for the two decision tasks in which the MPCR is equal to one: most subjects in the virtual-player sessions who contributed zero for all other decision tasks contributed full endowments when it cost them nothing to do so (a clear indication that they understood the payoff incentives).

Excluding treatments where the MPCR equals one, average public good contributions are 24.1% of endowment in the all-human treatment. This figure is 12.0% in the virtual-player treatment. The ratio of virtual-player contributions to all-human contributions provides an estimate of the proportion of all-human treatment contributions that stem from confusion: 49.9%. This estimate of approximately half of contributions to the public good resulting from confusion is consistent with previous results (Andreoni, 1995; Houser and Kurzban, 2002;





Figure 1. Experiment 1, Mean Contributions per Decision Task

**Result 1.** In virtual-player and all-human treatments of Experiment 1, contributions to the public good increase with an increase in the Marginal Per Capita Return and group size (for small MPCRs).

Decision tasks 1, 2, 3 and 4 and tasks 10, 11, 12 and 13 increase the returns from the public good while holding group size constant and equating internal and external returns from the public good. The corresponding MPCRs

within each set of four tasks are: 0.17, 0.33, 0.67 and 1. From Figure 1, and consistent with SF-2, one can easily see that contributions increase with an increase in MPCR in the human treatment. However, this increase in contributions occurs in *both* treatments: about 30 tokens from the lowest to highest MPCR. The correlation between the MPCR and contributions is nearly indistinguishable between the treatments:  $\rho = 0.97$  in the all-human treatment and  $\rho = 0.93$  in the virtual-player treatment.

Turning to group-size effects, note that tasks 1 through 6 are identical to 10 through 15 with the exception that the group size for the latter set is six rather than two. For the human treatment, focusing on tasks with "small" MPCRs (consistent with SF-3) – 1, 2 and 5 (n = 2) and 10, 11 and 14 (n = 6) – average contributions increase from 10.5% to 21.0% of endowment with an increase in group size from two to six. For the virtual-player treatment, contributions increase from 4.1% to 11.9%. The ratio of the change in virtual-player contributions to the change in all-human contributions suggests that nearly 75% of the group-size effect observed in the all-human treatment may be due to confusion.

# **Result 2**. Experiment 1 subjects matched with virtual players behave as if they are motivated by pure altruism.

To formally quantify the magnitude of pure altruism and warm-glow, GHL consider theoretical specifications for individual utility and estimate utility function parameters using a logit equilibrium model (see GHL for model and estimation details). We estimate logit equilibrium models with our data for the purpose of additional comparisons. The "Altruism" model considers pure altruism. The "combined" model considers both pure altruism and warm-glow. The parameter  $\alpha$  is a measure of pure altruism, the parameter *g* measures warm glow, and  $\mu$  is an error parameter. While  $\mu$  measures dispersion and does not indicate the magnitude of confusion contributions, GHL argue that statistical significance of this parameter reflects decision error. Estimated coefficients and standard errors are presented in Table 2.

Our all-human treatment results mirror those of GHL. In particular, we find that pure altruism is a statistically significant motive. Further, estimates of  $\mu$  are statistically different from zero for each specification.

Note, however, that the altruism parameter is also positive and statistically different from zero in the virtual-player treatment in which the contributions of "others" are just numbers on pieces of paper in an envelope on the subjects' desks. There is no altruism in such an environment. Altruism parameters are statistically different across models, but are of similar magnitude. Concentrating on the altruism model, the  $\alpha$ 's in the two treatments suggest that about half (53.6%) of the altruism detected in the all-human treatment is noise generated by

confused subjects using the experimental parameters as cues to guide payoff-maximizing contributions.<sup>3,4</sup>

Confused subjects thus use changes in the parameters across decision tasks as a cue of how to vary choices. In contrast to the white noise picked up in the logit equilibrium model's noise parameter, the confusion we identify systematically varies with the parameters designed to identify pure altruism. Thus confused individuals mimic altruists. In the next section, we apply the virtualplayer design to another single-shot VCM environment: the strategy-method design that has led some to conclude that conditional cooperation is the main motivation for contributions.

	All-Huma	n Treatment	Virtual-Player Treatment						
	Altruism	Combined	Altruism	Combined					
α	0.069*	0.054*	0.037*	0.038*					
	(0.009)	(0.015)	(0.006)	(0.010)					
g	-	0.380	-	-0.013					
		(0.300)		(0.209)					
μ	40.846*	38.116*	20.741*	20.786*					
	(2.205)	(2.887)	(0.911)	(1.165)					
Log-L	-2380.507	-2379.800	-2001.342	-2001.340					
N	720	720	720	720					

 Table 2. Experiment 1, Estimated Logit Equilibrium Models

Note: standard errors in parentheses.

\* indicates parameter is statistically different from zero at the five percent level.

Test results (tests allow  $\mu$  to vary across treatments). Altruism model:  $\alpha_V = \alpha_H$   $\chi^2(1) = 8.720, p = 0.003$ Combined model:  $\alpha_V = \alpha_H$  and  $g_V = g_H$   $\chi^2(2) = 10.068, p = 0.007$ 

<sup>&</sup>lt;sup>3</sup> GHL's estimates of  $\alpha$  range from ten to fourteen cents.

<sup>&</sup>lt;sup>4</sup> Our results do <u>not</u> stem from our modest changes to the GHL design. We also conducted an exact replication of the GHL instructions with all-human and virtual-player treatments (Cotten et al., 2007). The pattern of contributions in relation to design factors is similar to both the GHL results and our current results.

# IV. EXPERIMENT 2: CONFUSION MIMICS CONDITIONAL COOPERATION

#### **Experimental Design**

Fischbacher et al. (hereafter, FGF) use a variant of the strategy-method in a oneshot VCM (p.397) "that directly elicits subjects' willingness for conditional cooperation." Their innovation is the use of a "contribution table," which asks subjects to consider possible average contribution levels of the other group members, and state how much they would contribute to the public good conditional on each level. Each subject must indicate an unconditional contribution (traditional VCM design) <u>and</u> fill out a contribution table. Within each group of four players, the total contributions to the public good are determined by the unconditional contribution of three players and the relevant conditional response from the fourth player's contribution table.

Importantly, FGF claim (p.398) that having subjects "answer 10 control questions that tested their understanding of this public good problem...indicates that the subjects understood the mechanics and the implications of the above payoff function." In other words, FGF suggest their subjects are not confused about the incentives in the game. FGF classify 50% of the subjects as conditional cooperators, 14% as "hump-shaped" contributors who conditionally cooperate up to about half of endowment and then decline, 30% as free riders, and the rest as "other patterns." Other studies on conditional cooperation (Keser and van Winden; Fischbacher and Gächter, 2004; Burlando and Guala, 2005; Houser and Kurzban, 2005; Chaudhuri et al., 2006; Chaudhuri and Paichayontvijit, 2006) find higher rates of conditional cooperators compared to free riders, but similar frequencies of these player "types."<sup>5</sup>

To test if the use of control questions mitigates confusion and if the FGF design measures conditional cooperation, we conducted an experiment identical to FGF with the exception that subjects were matched with virtual players. We use the same control questions FGF claim ensures proper understanding of the game's incentives (see instructions in online appendix B). Forty undergraduate volunteers, split into two sessions, participated in the computerized experiment at the Georgia State University Experimental Center Laboratory. Earnings averaged \$28.27, and the experiment lasted less than 1.5 hours.

<sup>&</sup>lt;sup>5</sup> A recent study using the design in a repeated-round context (Fischbacher and Gächter, 2004) claims the majority of contributions are motivated by conditional cooperators with no evidence of pure or warm-glow altruism. They claim confusion accounts for few contributions to the public good ("at most 17.5 percent," p.3). They also argue that most of the decay in contributions is not from learning but from the interaction among free riders and conditional cooperators who revoke their cooperation once they realize they are among people who are not "norm abiders."

Ferraro and Vossler: Confusion in Public Goods Experiments



Figure 2. Experiment 2, Average Own Contribution by Average Contribution of Other Members (diagonal = perfect conditional)

The B.E. Journal of Economic Analysis & Policy, Vol. 10 [2010], Iss. 1 (Contributions), Art. 53

**Result 3.** In Experiment 2, a large fraction of subjects behave as if they are conditionally cooperating with automata programmed with predetermined contributions.

Our results are strikingly similar to FGF despite the fact that our subjects knew that they were not grouped with human beings or with agents whose choices would respond to their decisions; in particular, our Figure 2 looks similar to FGF's Figure 1 (p.400). Using FGF's criteria for classifying subjects, we are forced to classify 53% of our sample as conditional cooperators, 23% as free-riders, 13% as hump-shaped contributors, and the rest as "other patterns." The FGF practice questions thus do not mitigate confusion.

Confused players behave in similar fashion to conditional cooperators. If the experimenter asks them how much they would contribute if the other group members invested X tokens, the experimenter will see a high correlation between subject answers and X. Our results offer insights on why recent studies on conditional cooperation find little confusion and a lot of conditional cooperation. The confused subjects of previous studies (Andreoni, 1995; Palfrey and Prisbey; Houser and Kurzban, 2002; Ferraro et al.) have been reclassified as conditional cooperators, "hump-shaped" contributors and "others."<sup>6</sup>

#### V. EXPERIMENT 3: DYNAMICS OF REPEATED-ROUND VCM GAMES

#### **Experimental Design**

We use the archetypal repeated-round, linear VCM game instructions. Group size is four individuals who remain (anonymously) matched for a single treatment. Each subject is given an endowment of 50 laboratory tokens per round (US \$0.50). The MPCR is constant and equal to 0.50, thus making free-riding the dominant strategy and contributing the entire endowment the socially optimal strategy. These attributes of the experiment are common knowledge.

Instructions are presented both orally and in writing (see online appendix C). Subjects receive a payoff table that displays the payoff from the group exchange for every possible aggregate amount of tokens invested in the group exchange. Every subject answers a series of practice questions that tests their understanding of payoff calculations. No subject can proceed until all the questions are answered correctly. After each round, subjects receive information on their investment in the group exchange, the aggregate investment of the other group members, their payoff from the group exchange, and their payoff from their

<sup>&</sup>lt;sup>6</sup> The same problem is inherent in designs that identify conditional cooperation through a positive correlation between own contributions and beliefs about the contributions of others.

private exchange. On the decision screen is a "Transaction History" button, through which subjects can, at any time, observe the outcomes from previous rounds of the experiment. The same author moderated all of the sessions. In the all-human treatment, subjects play 25 rounds of the game. Each subject knows that she will be playing 25 rounds with the same three players. To prevent individuals from discerning the identity of other group members, group assignment is random and five groups participate simultaneously in the sessions (subjects are separated by dividers).

The virtual-player treatment is identical to the all-human treatment with one exception: each human is aware that she is paired with three nonhuman, virtual players and that each virtual player plays a predetermined contribution sequence. Subjects are informed that this contribution sequence is the same sequence of contributions produced by a human player in a previous all-human treatment. They are told that a computer scours a database of observations of human contributions in a previous all-human session and then picks at random (without replacement) a set of three human subjects from a group as the "identity" of the three virtual players. As with all our experiments, subjects are provided these contribution sequences on paper sealed in an envelope at their desk and reminded that the reason we provide this envelope is to prove to them that there is no deception: the virtual players behave exactly as the moderator explained they do.

In each session subjects face two experimental conditions, with 25 rounds of play in each. We designate participants in the first 25-rounds of a session as "T" (for "inexperienced") and participants in the second 25-rounds as "E" ("experienced"). At the beginning of each session, however, subjects are unaware that they would be playing an additional 25 rounds after the first 25 rounds. They simply begin with the instructions for the first 25 rounds. After the first 25 rounds are over, subjects are informed that there will be another 25 rounds.

Overall, with both inexperienced and experienced subject groups playing in the all-human (designated as "H") and virtual-player ("V") treatments, we have four experimental conditions that will be used to make inferences about the dynamics of subject behavior in the repeated-round VCM game:

- 1) **HI**: Participants are *inexperienced*, and play in all-human groups for 25 rounds.
- 2) **VI**: Participants are *inexperienced*, and play in virtual-player groups for 25 rounds.
- 3) **HE**: Participants are *experienced*, and play in all-human groups for 25 rounds.
- 4) **VE**: Participants are *experienced*, and play in virtual-player groups for 25 rounds.

The H*I* condition is the standard linear VCM game about which we wish to draw inferences about the subjects' motives. To do so, we contrast H*I* with V*I*, V*E* and H*E*. Subjects in a V*I* (V*E*) treatment observe the same history of contributions as subjects in a corresponding H*I* (H*E*) condition: each subject in H*I* (H*E*) has a "twin" in V*I* (V*E*). The only difference between H*I* (H*E*) and V*I* (V*E*) is that the humans in V*I* (V*E*) are playing with virtual players. Collecting data corresponding with all four experimental conditions requires running a sequence of three experiments. The H*I* data come from the first experiment, the V*I* and H*E* data from the second experiment, and the V*E* data from the third experiment. The experiments were run in this order because H*I* data are needed for the V*I* condition.<sup>7</sup>

Two-hundred and forty undergraduate volunteers participated in the computerized experiment at the Georgia State University Experimental Center Laboratory. Eighty were assigned to each experimental condition. Earnings averaged \$33.14, and the experiment lasted less than 1.5 hours.

Houser and Kurzban's (2002) (hereafter, HK) design is similar to ours, but there are three important differences. First, aggregate computer contributions to the public good in HK are three-fourths of the average aggregate contribution observed for that round in the human condition. Thus, the identification of contributions attributable to confusion in their design relies on the assumption that contributions in a given round are independent of the history of group contributions. If they are not, individual subjects are not independent observations and merely presenting all computer condition subjects with average aggregate contributions from the human condition thwarts important dynamics. Keser and van Winden, Ashley et al. (2003), and Carpenter find that contributions *are* history-dependent.

Second, and related to the role of the history of contributions, HK's computer condition changes the standard VCM game beyond simply grouping a human with automata. Human subjects in the computer condition observe their group members' aggregate contribution *before* they make their decision in a round (as opposed to after they make their decision, as in the human condition). If the history of contributions affects both confused and other-regarding subjects, then such a change in design can also affect the comparability of the two treatments. Third, HK do not attempt to discriminate among different kinds of other-regarding preferences or confusion behaviors, whereas we present in the next subsection a microeconometric model to undertake this discrimination.<sup>8</sup>

<sup>&</sup>lt;sup>7</sup> The last 25 rounds of the first experiment and the first 25 rounds of the third were with virtual players, and these data are not included in the analysis.

<sup>&</sup>lt;sup>8</sup> Our sample size is also much larger; HK have only 20 subjects in their all-human sample.

#### **Microeconometric Model**

We analyze the contributions data at the individual level using a partial adjustment model similar to that of Mason and Philips (1997) to analyze cooperation in a common property resource experiment. The model in particular is intended to capture potential learning of the dominant strategy ("hill-climbing") through inclusion of a profit feedback variable, and conditional cooperation or herding, by conditioning current contributions on the past contributions of other group members. Specifically, the econometric model is:

$$[2] y_{it} = \alpha + \beta_1 y_{i,t-1} + \beta_2 y_{i,t-2} + \lambda [(y_{i,t-1} - Y_{i,t-1} / (n-1)] + \gamma [D_{i,t-1}(\pi_{i,t-1} - \pi_{i,t-2})] + \varepsilon_{it}$$

where  $y_{it}$  and  $Y_{it}$  denote own and group contributions, and  $\pi_{it}$  denotes earnings for participant *i* in round *t*;  $\varepsilon_{it}$  is a mean-zero error term that captures the analyst's uncertainty about the specification of individual behavior. One and two-period lags of own contributions are included to capture dynamics present in the data that are not quantified by other components of the model. Inherent in the model is the standard assumption that warm-glow and pure altruism do not diminish over time (e.g., Palfrey and Prisbrey); warm-glow or pure altruism is thus depicted by the relationship between contributions and a constant term

The next-to-last covariate captures conditional cooperation in the allhuman treatment, and herding - a label that we use to describe the situation where a player uses the actions of others as an indication of profit-maximizing behavior - in all-human and virtual player treatments. In particular, the term allows contributions to be conditioned on the past deviations of own contributions from the average of contributions from other group members. For the herder, a negative (positive) deviation is a signal that she is contributing less (more) than average and should thus increase contributions. A conditional cooperator should behave in a similar manner to a herder: she increases her contribution if the average group member is contributing more than her, and decreases contributions when she perceives she is giving too much relative to others. Note that while conditional cooperator and herder behavior may look the same, the motivation for the behavior is different. For the herding subject, the average contribution from others is a signal of how the subject should behave; for the conditional cooperator, the average contribution of others is a signal of whether the other players are normabiders or they are taking advantage of the subject. Larger estimates of the coefficient  $\lambda$  with all-human versus virtual-player treatment data indicate that conditional cooperator behavior is an important motive.

The last covariate is a profit feedback variable that captures the interaction between a change in past contributions and the associated change in profit. In particular,  $D_{i,t-1}$  is an indicator variable that equals 1 if the subject increases

contributions from round t-2 to t-1, equals -1 if contributions decrease between rounds t-2 and t-1, and equals 0 when contributions are unchanged. The profit feedback mechanism directs the hill climber to continue to increase (decrease) contributions if they increased (decreased) last period and earned more money, or directs her to adjust contributions in the opposite direction when their last adjustment yielded lower earnings. No profit feedback is provided when contributions or profits do not change between rounds t-2 and t-1.

In estimating the unknown parameters of our model, it is important to account for the characteristics of our dependent variable as well as be suitable for panel data. Contributions are, by construction, non-negative integers, and there are a preponderance of zeros and small values. The "count data" nature of contributions lends itself well to a Poisson estimator. The standard arguments for discrete choice models motivate using a Poisson over OLS: OLS predicts negative values and the discrete nature of the data causes OLS errors to be heteroskedastic. We account for unobserved, individual-specific heterogeneity (i.e., the panel structure of the data) by using a Poisson quasi-MLE. Specifically, this is the Poisson MLE coupled with White's (1982) robust covariance estimator, a.k.a. the "sandwich" estimator, adjusted for clustering at the individual level. This estimator is robust (i.e., consistent) to a variety of misspecifications, including distributional misspecification and unspecified autocorrelation. Though we provide motivation for using the Poisson estimator, note that Poisson and OLS estimators produce qualitatively consistent results for this data.

Although our econometric model includes one and two-period lags of the dependent variable as explanatory factors, the number of lags to include (i.e., how backwards looking subjects are) is largely an empirical question. We estimated models (available upon request) that included up to five-period lags. Inferences drawn from these alternative specifications are similar to those presented below.<sup>9</sup>

#### Results

Figure 3 presents average contributions by round as a percentage of endowment for each of the four experimental conditions. Table 3 presents the estimated coefficients from the econometric model corresponding to each experimental condition.

Comparing all-human and virtual-player contribution rates with inexperienced subjects represents the distinction between contributions stemming from other-regarding motives versus those due to confusion in the standard VCM

<sup>&</sup>lt;sup>9</sup> In a similar vein, we investigated more general specifications that allowed the slopes on the "feedback" and group behavior variables to depend on whether the deviations were positive or negative. We failed to reject our more parsimonious specification using conventional tests.

Ferraro and Vossler: Confusion in Public Goods Experiments



Figure 3. Experiment 3, Mean Contributions per Round

Published by The Berkeley Electronic Press, 2010

Dependent variabl	e is $y_{it}$ ( <i>i</i> 's cor	ntribution to the p	ublic good in rou	and $t$ )	
Variable	Parameter	All-Human, inexperienced	Virtual- Player, inexperienced	All-Human, experienced	Virtual- Player, experienced
Intercept	α	1.8516** (0.0738)	1.2482** (0.1353)	1.3803** (0.1001)	0.9379** (0.1235)
<i>y</i> <sub><i>i,t</i>-1</sub> [subject contributions in round <i>t</i> -1]	$eta_1$	0.0337** (0.0028)	0.0376** (0.0047)	0.0497** (0.0037)	0.0462** (0.0059)
<i>y<sub>i,t-2</sub></i> [subject contributions in <i>t-2</i> ]	$\beta_2$	0.0141** (0.0016)	0.0298** (0.0034)	0.0150** (0.0024)	0.0353** (0.0038)
$y_{i,t-1} - Y_{i,t-1}/(n-1)$ [deviation from average contributions of other group members in <i>t</i> -1]	λ	-0.0153** (0.0023)	-0.0072* (0.0040)	-0.0256** (0.0035)	-0.0143** (0.0039)
$D_{i,t-1}(\pi_{i,t-1} - \pi_{i,t-2})$ [profit "feedback" mechanism]	γ	0.0044** (0.0018)	0.0062* (0.0033)	-0.0003 (0.0037)	0.0039 (0.0045)
Log-L		-13,751.38	-12,538.29	-13,530.07	-10,134.32
Ν		1840	1840	1840	1840

Table 5. Experiment 5, Dynamic roisson mouels of mutvicular denay	ſał	Гa	ľa	at	D	e	3.	ŀ	X	bei	rin	ner	1t	3.	D	<b>y</b>	na	mic	P	oi	SSO	n N	Ло	bde	els	of	lı	ıdi	vi	dua	al	B	eha	avi	İC	r
---	-----	----	----	----	---	---	----	---	---	-----	-----	-----	----	----	---	----------	----	-----	---	----	-----	-----	----	-----	-----	----	----	-----	----	-----	----	---	-----	-----	----	---

*Note:* robust standard errors are in parentheses. \* and \*\* indicate that parameters are statistically different from zero at the 5% and 1% level, respectively. Consistent with our theoretical hypotheses, these are one-sided tests.

game. Mean contributions to the public good in H*I*, which represents the standard VCM game where inexperienced subjects play with other human subjects over repeated rounds, start at 50.1% of endowment in round 1, and steadily decline to 14.1% by round 25. This parallels the standard finding in the literature of 40 to 60% contributions in the initial period followed by a steady decline (Davis and Holt).

In comparison, VI contributions start at 28.5% and fall to 9.8% by round 25. On average, subjects contribute 32.5% and 16.8% of all endowments to the public good in the all-human and virtual-player treatments, respectively. Dividing VI contributions by HI contributions suggests that 51.6% of the total contributions in the standard VCM game stem from confusion; the remaining 48.4% are attributable to other-regarding behavior. Statistical tests indicate that public good contributions are statistically different, and *higher*, in the all-human treatment at the 5% level both on average and in 24 of 25 rounds (see online appendix D).

In their related study, HK find that, on average, 54% of the total contributions in their all-human treatment are attributable to confusion. Focusing on our first ten rounds, the length of the HK experiment, our figure is 53%. Although these summary statistics are quite close, note that HK find that the rate of contributions decline in the all-human treatment is statistically *slower* than the virtual-player treatment. This suggests a larger fraction of the observed contributions is attributable to other-regarding preferences as the experiment progresses (and *less* is due to confusion). In contrast, our rate of decline is statistically different and is about 1.8 times *faster* for the all-human treatment, indicating that other-regarding behavior declines over rounds.<sup>10</sup>

Ledyard (p.146) conjectures that confused subjects, when asked to invest an amount between zero and their entire endowment, might simply split their endowment approximately half-half. Our first-round data support his conjecture: in HI, 31 subjects chose a contribution between 20 and 30 tokens and in VI, 29 subjects chose a contribution between 20 and 30 tokens. Note that in HI, 11 subjects contributed their entire endowment, while none did in the VI, suggesting that most full-endowment contributors are not confused. Results from Experiment 4 will corroborate this: improved instructions make endowment splitting rare, but not full endowment contributions.

<sup>&</sup>lt;sup>10</sup> We regress mean contributions (%) on a constant and an indicator variable for the experiment round. To facilitate hypothesis tests, this is done within a time-series cross-section modeling framework (see Greene 2003, p. 320-333) whereby each treatment is a cross-sectional unit observed over a 25 period time horizon. This framework allows for treatment-specific heteroscedasticity, first-order autocorrelation, and correlation across units. The estimated relationships for the HI and VI conditions are: [HI] contributions = 49.05 – 1.27\*round; [VI] contributions = 25.84 – 0.69\*round. A likelihood ratio test rejects the hypothesis of equal slope coefficients for two experiment conditions [ $\chi^2(1)=13.43$ , p<0.01].

Although experienced subjects in the all-human treatment contribute less than inexperienced subjects (HI vs. HE), a finding consistent with the literature, the relationships observed between virtual-player and all-human treatments with inexperienced subjects are robust to experience (HE vs. VE). That is, there is statistical evidence that contributions stemming from other-regarding behavior are significant and are *decreasing* over rounds. In particular, other-regarding preferences account for 51%, 47%, and 25% of total contributions in rounds 1, 10, and 25, respectively. The rate of decline is approximately 1.6 times *faster* for the all-human treatment.<sup>11</sup>

**Result 4.** Contribution rates in Experiment 3 are similar across inexperienced and experienced subjects in the virtual-player treatment. Thus, there is no evidence that increasing awareness of the dominant strategy drives the decay of contributions over time.

If a substantial component of the decay in contributions across rounds stems from subjects becoming aware of (i.e., "learning") the dominant strategy of zero contributions (SF-1), the contributions from inexperienced subjects in VIshould be significantly higher than contributions from experienced subjects in VE. The data do not support this implication. Average contributions are 16.8% and 11.9% of endowment with inexperienced and experienced subjects, respectively. Using a Kolmogorov-Smirnov Test where the average contribution across rounds from an individual is used as an independent observation, these averages are not statistically different at the 5% level. Inexperienced subject contributions are only statistically higher than experienced subject contributions in the first three rounds and in round 22 (see online appendix D). While this pattern suggests that a few inexperienced subjects may have indeed (quickly) learned the dominant strategy, overall learning effects appear to be minimal. Analysis of the numbers of free-riders (\$0 contribution) by round yields a similar conclusion: in Round 1, Round 2, Round 24 and Round 25 of VI, there were 22, 27, 49 and 50 free-riders; the corresponding numbers in VE are 38, 31, 46 and 58.

An alternative explanation for the decay in virtual-player contributions is that confused subjects are simply herding on the observed downward trend in virtual player contributions (which reflect behavior in past all-human sessions). We test this alternative hypothesis directly using our econometric model.

<sup>&</sup>lt;sup>11</sup> Using the framework outlined in footnote 8, the estimated relationships for the HE and VE conditions are: [HE] contributions = 31.26 - 0.79\*round; [VE] contributions = 18.19 - 0.48\*round. A likelihood ratio test rejects the hypothesis of equal slope coefficients for these experiment conditions [ $\chi^2(1)$ =4.63, p=0.03].

Ferraro and Vossler: Confusion in Public Goods Experiments

**Result 5**. In Experiment 3, the majority of the decline in contributions in the virtual-player treatment with inexperienced or experienced subjects arises from herding behavior.

All parameters of the estimated models have the expected sign and are statistically significant at the 5% level, with the exception of the parameter on the profit feedback variable, which is only significant for inexperienced subjects. The lack of statistical significance on the feedback variable with experienced subjects is consistent with our expectation that most of the "hill-climbing" or "reinforcement learning" would dissipate over repeated rounds. In the interest of determining whether there is a cut-off point during the experiment where the average reinforcement learning that takes place becomes negligible, we generalized our virtual-player model for inexperienced subjects in Table 3 by allowing a structural break with respect to the feedback variable. This investigation yields an interesting result: we fail to reject the hypothesis that contributions due to reinforcement learning are statistically different from zero in periods 9-25 (we fail to reject this hypothesis for the all-human treatment as well). Thus, the main driving force behind the decay in virtual-player contributions is herding behavior. The next section reports focus group results that further elucidate this herding behavior.

For both experienced and inexperienced subjects, the estimate of  $\lambda$  is statistically *larger* (in absolute value) in the all-human treatment than in the corresponding virtual-player treatment at the 5% level [inexperienced: z = 1.76, p = 0.04; experienced: z = 2.15, p = 0.02]. Thus, this suggests conditional cooperation is a statistically significant motive for contributions in the all-human conditions ( $\lambda^{H} > \lambda^{Herd}$ ).

The above findings suggest that history matters: contributions of group members in period t-1 influence individual contributions in period t. Herders look to history for a signal on how they should behave in a confusing situation. Conditional cooperators look to history to infer whether they are playing with "norm abiders" and thus whether they should continue to cooperate or begin to revoke their cooperation. Thus, analysts who model individual behavior in public goods experiments must appropriately account for the dynamics associated with repeated group interactions in order to make valid inferences.

Finally, given the standard assumption that warm-glow and pure altruism do not decay over rounds, we can use the difference between all-human and virtual-player contributions in the last round as an upper bound on warm-glow/pure altruism contributions. The average inexperienced subject contributes, at most, 4.23% of their endowment due to warm-glow and pure altruism. For experienced subjects, this figure is 2.3%. Putting this into another perspective, at most just 13.0% and 11.0% of observed contributions across rounds could be

attributed to warm-glow and pure altruism for inexperienced and experienced subjects, respectively.

Thus, in the absence of punishment opportunities, the co-existence of free riders, conditional cooperators and herders leads to substantial decline in contributions to the public good. The initial contribution behavior, rather than the payoff outcome, starts a cascade of declining contributions through the revocation of cooperation by disappointed conditional cooperators and the herding on the downward trend by confused players.

#### **Focus Group**

To explore subject motivations in more depth, we paid subjects in our last session (n=20) an additional \$10 to serve as a focus group to provide written and, afterwards, one-by-one oral feedback to the experimenters. These subjects had just completed playing 50 rounds in virtual-player groups. As summarized below, the subjects reveal misperceptions of the game form that match behavior in the experiments. More detailed results are contained in online appendix E.

First, all subjects stated in writing and then confirmed orally that they were playing with nonhuman agents with pre-determined decisions that could not be affected by their actions. Second, subjects were asked, "In your opinion, what is the point of this experiment?" Almost all subjects wrote something about observing how people make investments (e.g., "Observe investment speculation on a personal and group level," "To create a computer program that can predict a person's investment decisions. Possibly to predict fluctuations in the stock market."). In other words, many subjects believe they are playing a sort of stock market game with incentives that differ substantially from the public good setting experimentalists believe they are studying.

Third, subjects responded to the following question: "How did you determine how many tokens to invest in the Group Exchange in the early rounds of the experiment?" They were offered four multiple choice responses: (A) The choice was clear from the instructions; (B) I invested different amounts and watched how my payoff changed; (C) I observed how many tokens the Virtual Players invested and altered my decision accordingly; and (D) Other (please specify). Subjects were instructed they could choose more than one response. Only 30% answered A. Fifty-five percent answered B and 65% answered C (only one subject chose D).

After the written responses were completed, the moderator asked each subject orally for more detail on how he or she made decisions in the early rounds of the experiment. Only 25% of the subjects said that the payoff-maximizing strategy was clear from the instructions. Ten percent of subjects reported having no idea about what was going on and simply chose contribution levels at random.

Another 10% reported attempting, without success, to vary their contributions and infer a pattern. Twenty percent reported depending solely on the behavior of the virtual players to determine their own contribution. Thirty-five percent of the subjects reported a mix of beginning with a split of their endowment, followed by watching what the virtual players were doing and by attempting to infer if there was any pattern to earnings, followed quickly by abandoning any attempt to infer a pattern and instead herding along with the virtual players. Only one of these subjects reported finally "getting it" and changing his behavior for the second set of 25 rounds.

Fourth, the moderator was aware that two-thirds (67.5%) of the subjects correctly answered "0" to the post-experiment question about the payoffmaximizing Group Exchange contribution. When the moderator asked each subject (first in writing and then orally) why he or she wrote down zero to this question, but generally did not invest zero in the Group Exchange, two general responses were heard: (1) one had to come up with an answer to the postexperiment question and given the virtual players were contributing at zero or near zero in the final rounds, an answer of "0" seemed like the best answer; and (2) the question was asking about the "risk-free investment decision." References to "risk" were common, orally and in writing, among self-reported herders. As explained in more detail in online appendix E, these subjects understood that higher group payoffs were engendered when all members contributed, but they mistakenly thought that this outcome maximized their own earnings. In writing, about half the respondents indicated that the best response depended on what the virtual players were choosing (e.g., "More money could be made in the group investment versus not investing at all. In the previous rounds, the virtual players were on a gradual increase in investing in the group. So I wanted to get more money," "The virtual players invested in the group exchange and it was profitable for me to get in on the money," "[I invested in group exchange] because I thought the robots would be doing the same," "[In invested in group exchange] because I thought in addition to my individual exchange investment, the group investment would increase my earnings potential.").

The focus group results thus suggest that approximately one in three subjects in this session correctly understood the incentives. The other two-thirds either found the incentives undecipherable and herded on virtual player contributions (sometimes after unsuccessful hill-climbing exercise) or erroneously believed they were playing a risky investment game, where zero contributions are seen as a risk-dominant strategy but not as free-riding. In other words, a substantial proportion of subjects begin and end the experiment without recognizing the tension between the privately-optimal strategy of free-riding and the socially-optimal strategy of contributing to the public good. Based on results in the first three experiments and the focus group, we infer that this failure of game form recognition arises because of (1) experimenter demand effects associated with the standard "investment" language used in the instructions (i.e., subject ask themselves why would two accounts be provided if the goal was not to determine the optimal allocation across accounts or simply to test their understanding of the aphorism "never put all your eggs in one basket?"); and (2) an inability to determine how to maximize payoffs (i.e. participants do not see that own payoffs rise by lowering own contributions, *ceteris paribus*). In the next section, we test this hypothesis directly in the context of an experimental design similar to that of Experiment 3.

#### **VI. EXPERIMENT 4: CONFUSION REDUCTION**

#### **Experimental Design**

Levitt and List (2007) argue that experimentalists should anticipate the types of biases common to the lab, and design experiments to minimize such biases. To address experimenter demand effects associated with the standard "investment" language used in the instructions, we substitute it with "donation" language. More precisely, the instructions describe the decision task as deciding how many tokens to "keep" and how many tokens to "donate" instead of describing it as deciding how many tokens to invest in individual and group exchanges. Also, instead of stating that all group members "earn" tokens from the group exchange, the instructions state that donated tokens are "shared equally between all members of your group." To facilitate the ability of subjects to recognize the social dilemma inherent in the public goods game, we offer subjects a complete payoff table. As we note below, associative framing and complete payoff information have been used before in public goods experiments. However, they have not been used to study the role of confusion nor have they been used concurrently.<sup>12</sup>

In the context of the repeated-round VCM, we predict that if these changes to the instructions reduce confusion, few subjects would contribute to the public good when playing with non-human players. Based on previous published conjectures that endowment-splitting reflects confusion (Ledyard), we also predict that endowment splitting would decline dramatically under the modified instructions (note that endowment splitting can also be considered a stylized fact of the published literature). Based on our results from Experiment 3, we also predict that, under the modified instructions, hill-climbing will be less evident and

<sup>&</sup>lt;sup>12</sup> We also explored the framing and payoff matrix effects in isolation with small samples. When used in isolation, either approach reduces confusion by about half based on our metrics. Our overall results suggest that the two instructional changes are best used in tandem.

more subjects will correctly identify the payoff-maximizing strategy in a postexperiment questionnaire in which subjects are paid for the correct answer.

Framing, or context, may be a source of confusion if it creates associations that cause subjects to misperceive the game's incentives. Although the goal of many experimentalists is to eliminate context, Loewenstein (1999, p. F30) notes that "the context-free experiment is, of course, an elusive goal." Loewenstein (p. F31) further argues that we should create "context that is similar to the one in which economic agents will actually operate." In the VCM game, Cookson (2000) finds greater average contributions (50.7% of endowment versus 33.0%) when he makes one small change in the framing of the standard instructions: instead of framing the decision as dividing tokens between an "individual exchange" and a "group exchange," the revised instructions frame the decision as "donating" tokens to the group, with everyone receiving an equal "share" of the donation returns. Similar results and conclusions were drawn by Rege and Telle (2004).

Standard linear VCM instructions provide subjects with information on returns from the group exchange given total investment in the group exchange. A subject must make additional computations to determine that, holding the contributions of others' constant, *individual* earnings are maximized by investing nothing in the group exchange. Saijo and Nakamura (1995) provide a payoff matrix that shows own payoffs corresponding with combinations of own and group investment in the group exchange. They find that average contributions are substantially closer to the Nash equilibrium when subjects have the payoff matrix than when they do not.

We believe that instructions should display both own and other payoffs, as the latter may be important for subjects who have other-regarding preferences. Similar to Charness et al. (2004) in a gift-exchange game, and Oxoby and Spraggon (Forthcoming) in a nonpoint pollution tax experiment, we design a "complete" payoff matrix that displays own payoffs and average payoffs of other group members for combinations of own contributions and the average contributions of other group members.<sup>13</sup> This payoff matrix is shown in Figure 4.

The rest of the design is similar to that of Experiment 3. Group size is four and the MPCR is 0.5. Each subject is given an endowment of 10, which makes the size of the payoff matrix manageable. The experiment begins with 25 rounds under the all-human treatment, whereby each participant remains anonymously matched with three other participants. These rounds are followed by five founds under the virtual-player treatment, where each participant is matched with three virtual players. The contributions sequences for the virtual players are taken from

<sup>&</sup>lt;sup>13</sup> Charness et al. find that providing a "comprehensive" payoff table leads to significantly lower cooperation, and Oxoby and Spraggon find more individuals playing the dominant strategy.

#### The B.E. Journal of Economic Analysis & Policy, Vol. 10 [2010], Iss. 1 (Contributions), Art. 53

Your Tokens in				Average Number	r of Tokens Put ir	n Group Exchang	e by <u>Other 3 Gr</u>	oup Members				
Group Exchange ↓	0 1		2	3	4	5	6	7	8	9	10	
0	( <b>\$0.60</b> , \$0.60)	( <b>\$0.69</b> , \$0.63)	( <b>\$0.78</b> , \$0.66)	( <b>\$0.87</b> , \$0.69)	( <b>\$0.96</b> , \$0.72)	( <b>\$1.05</b> , \$0.75)	( <b>\$1.14</b> , \$0.78)	( <b>\$1.23</b> , \$0.81)	( <b>\$1.32</b> , \$0.84)	( <b>\$1.41</b> , \$0.87)	( <b>\$1.50</b> , \$0.90)	
1	( <b>\$0.57</b> , \$0.63)	( <b>\$0.66</b> , \$0.66)	( <b>\$0.75</b> , \$0.69)	( <b>\$0.84</b> , \$0.72)	( <b>\$0.93</b> , \$0.75)	( <b>\$1.02</b> , \$0.78)	( <b>\$1.11</b> , \$0.81)	( <b>\$1.20</b> , \$0.84)	( <b>\$1.29</b> , \$0.87)	( <b>\$1.38</b> , \$0.90)	( <b>\$1.47</b> , \$0.93)	
2	( <b>\$0.54</b> , \$0.66)	( <b>\$0.63</b> , \$0.69)	( <b>\$0.72</b> , \$0.72)	( <b>\$0.81</b> , \$0.75)	( <b>\$0.90</b> , \$0.78)	( <b>\$0.99</b> , \$0.81)	( <b>\$1.08</b> , \$0.84)	( <b>\$1.17</b> , \$0.87)	( <b>\$1.26</b> , \$0.90)	( <b>\$1.35</b> , \$0.93)	( <b>\$1.44</b> , \$0.96)	
3	( <b>\$0.51</b> , \$0.69)	( <b>\$0.60</b> , \$0.72)	( <b>\$0.69</b> , \$0.75)	( <b>\$0.78</b> , \$0.78)	( <b>\$0.87</b> , \$0.81)	( <b>\$0.96</b> , \$0.84)	( <b>\$1.05</b> , \$0.87)	( <b>\$1.14</b> , \$0.90)	( <b>\$1.23</b> , \$0.93)	( <b>\$1.32</b> , \$0.96)	( <b>\$1.41</b> , \$0.99)	
4	( <b>\$0.48</b> , \$0.72)	( <b>\$0.57</b> , \$0.75)	( <b>\$0.66</b> , \$0.78)	( <b>\$0.75</b> , \$0.81)	( <b>\$0.84</b> , \$0.84)	( <b>\$0.93</b> , \$0.87)	( <b>\$1.02</b> , \$0.90)	( <b>\$1.11</b> , \$0.93)	( <b>\$1.20</b> , \$0.96)	( <b>\$1.29</b> , \$0.99)	( <b>\$1.38</b> , \$1.02)	
5	( <b>\$0.45</b> , \$0.75)	( <b>\$0.54</b> , \$0.78)	( <b>\$0.63</b> , \$0.81)	( <b>\$0.72</b> , \$0.84)	( <b>\$0.81</b> , \$0.87)	( <b>\$0.90</b> , \$0.90)	( <b>\$0.99</b> , \$0.93)	( <b>\$1.08</b> , \$0.96)	( <b>\$1.17</b> , \$0.99)	( <b>\$1.26</b> , \$1.02)	( <b>\$1.35</b> , \$1.05)	
6	( <b>\$0.42</b> , \$0.78)	( <b>\$0.51</b> , \$0.81)	( <b>\$0.60</b> , \$0.84)	( <b>\$0.69</b> , \$0.87)	( <b>\$0.78</b> , \$0.90)	( <b>\$0.87</b> , \$0.93)	( <b>\$0.96</b> , \$0.96)	( <b>\$1.05</b> , \$0.99)	( <b>\$1.14</b> , \$1.02)	( <b>\$1.23</b> , \$1.05)	( <b>\$1.32</b> , \$1.08)	
7	( <b>\$0.39</b> , \$0.81)	( <b>\$0.48</b> , \$0.84)	( <b>\$0.57</b> , \$0.87)	( <b>\$0.66</b> , \$0.90)	( <b>\$0.75</b> , \$0.93)	( <b>\$0.84</b> , \$0.96)	( <b>\$0.93</b> , \$0.99)	( <b>\$1.02</b> , \$1.02)	( <b>\$1.11</b> , \$1.05)	( <b>\$1.20</b> , \$1.08)	( <b>\$1.29</b> , \$1.11)	
8	( <b>\$0.36</b> , \$0.84)	( <b>\$0.45</b> , \$0.87)	( <b>\$0.54</b> , \$0.90)	( <b>\$0.63</b> , \$0.93)	( <b>\$0.72</b> , \$0.96)	( <b>\$0.81</b> , \$0.99)	( <b>\$0.90</b> , \$1.02)	( <b>\$0.99</b> , \$1.05)	( <b>\$1.08</b> , \$1.08)	( <b>\$1.17</b> , \$1.11)	( <b>\$1.26</b> , \$1.14)	
9	( <b>\$0.33</b> , \$0.87)	( <b>\$0.42</b> , \$0.90)	( <b>\$0.51</b> , \$0.93)	( <b>\$0.60</b> , \$0.96)	( <b>\$0.69</b> , \$0.99)	( <b>\$0.78</b> , \$1.02)	( <b>\$0.87</b> , \$1.05)	( <b>\$0.96</b> , \$1.08)	( <b>\$1.05</b> , \$1.11)	( <b>\$1.14</b> , \$1.14)	( <b>\$1.23</b> , \$1.17)	
10	( <b>\$0.30</b> , \$0.90)	( <b>\$0.39</b> , \$0.93)	( <b>\$0.48</b> , \$0.96)	( <b>\$0.57</b> , \$0.99)	( <b>\$0.66</b> , \$1.02)	( <b>\$0.75</b> , \$1.05)	( <b>\$0.84</b> , \$1.08)	( <b>\$0.93</b> , \$1.11)	( <b>\$1.02</b> , \$1.14)	( <b>\$1.11</b> , \$1.17)	( <b>\$1.20</b> , \$1.20)	

#### EARNINGS EACH ROUND

Rows represent YOUR decisions (the allocation of YOUR tokens).

Columns represent the average number of tokens placed in the Group Exchange by your three group members.

The first number within each cell (in bold font) is YOUR total earnings from BOTH the Group Exchange and the Individual Exchange.

The second number is the average earnings of each of the other 3 group members from BOTH the Group Exchange and the Individual Exchange.

#### Figure 4. Experiment 4, Complete Payoff Matrix

http://www.bepress.com/bejeap/vol10/iss1/art53

the HI condition of Experiment 3 (scaled by 1/5 to account for differences in the endowment).

A pilot experiment indicated that students needed a detailed explanation to comprehend the complete payoff table. In response, we made the instructions longer to draw the subjects' attention to specific cells in the payoff matrix. Our aim was to make subjects aware that, *ceteris paribus*: (1) own payoffs decreased (increased) as own contributions increased (decreased); (2) own payoffs decreased (increased) as contributions from others decreased (increased); and (3) other's payoffs decreased (increased) as own contributions decreased (increased). Instructions were presented orally and in writing (see online appendix F).

Eighty-eight undergraduate volunteers participated in the computerized experiment at the University of Tennessee Experimental Economics Laboratory. The sample was split between two treatments. The first treatment used standard instructions (56 subjects). The second treatment used the modified instructions (32 subjects). Earnings averaged \$27.50. Sessions lasted an average of 68 and 84 minutes, respectively, in standard and modified instruction treatments.

#### Results

The main results from the experiment are illustrated in Figure 5 and Figure 6. First, as seen in Figure 5, average contributions for the all-human rounds do not vary much across treatments and they display the standard contributions decay through repetition. Mean contributions across the 25 rounds are 34.2% of endowment with standard instructions and are 28.7% with the modified instructions. The similarity of mean contributions across treatments is not unexpected. As mentioned above, the two instruction modifications are likely to have changed contributions in opposite directions: providing a complete payoff table leads to a reduction in average contributions whereas the associative framing increases contributions. However, when we examine the heterogeneity of subjectspecific contributions and contributions in the virtual-player treatment, we see striking differences. Figure 6(a) and Figure 6(b) display a histogram of firstperiod contributions for the all-human and virtual-player conditions, respectively.

# **Result 6**. The use of modified instructions with associative framing and a complete payoff table substantially reduces the number of confused subjects in the repeated-round VCM experiment.

Recall that Ledyard has suggested that the high frequency of "endowmentsplitting" observed among VCM experimental subjects is indicative of confusion (a conjecture supported by results in Experiment 3). Figure 6(a) shows that, with the modified instructions, there is much less endowment-splitting: 9.4% versus



The B.E. Journal of Economic Analysis & Policy, Vol. 10 [2010], Iss. 1 (Contributions), Art. 53

Figure 5. Experiment 4, Mean Contributions per Round



#### Ferraro and Vossler: Confusion in Public Goods Experiments

#### (a) All-Human Condition



(b) Virtual-Player Condition

#### Figure 6. Experiment 4, Round 1 Contributions

Published by The Berkeley Electronic Press, 2010

26.8% (Fisher exact test, p=0.04). Instead, there is a higher proportion of subjects who either free-ride or contribute their full endowment with the modified instructions: 40.6% versus 23.2% (p=0.07).

Figure 6(b) shows that in the first period with virtual players, 84.4% of subjects contribute zero tokens in the modified instruction treatment, whereas it is just 41.1% of subjects with standard instructions.<sup>14</sup> Using non-zero contributions as a measure of confused participants, the modified instructions serve to reduce confusion by over 70%, from 58.9% to 16.6%, and this difference is statistically significant (p < 0.01). As a second metric of confusion, similar to our other experiments, participants were asked to identify the profit-maximizing strategy for the virtual-player condition. With standard instructions, similar to Experiment 3, 32.1% of participants answered incorrectly. No participants in the modified instructions treatment answered incorrectly, and this difference between treatments is statistically significant (p < 0.01). Finally, we analyzed the all-human condition data using the same model applied to Experiment 3 data. We find statistically significance evidence of "hill climbing" with the standard instruction treatment [ $\gamma$ =0.005, std. err.= 0.002], similar to Experiment 3, but not with the modified instruction treatment [ $\gamma$ =0.003, std. err.=0.003]. This is yet another piece of evidence that suggests the modified instructions reduce confusion, as players who understand the incentives of the game do not need to search for the profitmaximizing strategy.

#### **VII. Rival Explanations**

Rival explanations of our results must not only explain the confirmation of our predictions in Experiment 4, but also the results from the first three experiments and the subjects' written and oral responses in the focus group. There are at least four potential rival explanations, the first three of which are also noted in Houser and Kurzban (2002). Subjects may have contributed to the public good when playing with virtual agents (1) because they wished to express altruism towards the experimenter; (2) because of social pressure from being observed by the experimenter; (3) because subjects did not understand (or forgot) that the virtual players were non-human; or (4) because subjects knew the virtual players were

<sup>&</sup>lt;sup>14</sup> In Experiment 3, we also ran 25 rounds of the virtual-player treatment (V*E*) after H*I* to ensure total earnings were the same across sessions (not reported). With this design, only 35% of subjects contributed zero tokens in the first round of the virtual-player treatment. Note also that for all designs in Experiments 3 and 4 that use standard instructions, we observe a "restart" effect in that the contributions go up in the first round of the new treatment (virtual-player or all-human) in comparison to the last round of the old treatment, but this restart effect is not evident with the modified instructions.

non-human but they played as if the virtual players were humans (perhaps, as noted by one referee, because subjects play a lot of video games and cannot distinguish reality from fantasy, or simply because subjects are altruistic even if other players are nonhuman).

We find the first two rival explanations deficient for several reasons. Most importantly, altruism towards the experimenter or social pressure is at odds with the dramatic drop in contributions when subjects play with virtual players under the modified instructions in Experiment 4, for which the language about virtual players remained unchanged. If anything, the use of the word "donations" in the modified instructions would seem to encourage altruism towards the experimenter or increase social pressure. Second, in our focus groups, the subjects who contributed to the public good with virtual players, and who had no ability to coordinate their stories, told similar stories about guessing the optimal investment strategy, about herding on virtual-player contributions because they could not infer the optimal investment strategy, or about seeing the opportunity for coordination while failing to make any mention of the opportunity to free-ride. If they instead contributed because of altruism towards the experimenter or social pressure, it would seem far easier and better for one's self-image to have simply stated that they contributed to the public good so that the experimenter did not have to spend too much of his research funds. No participant offered such a motive.

Third, altruism or social pressure would also have to explain why there were so many incorrect answers to our (paid) dominant-strategy question in post-experiment questionnaires for the Experiments 1, 3 and 4 (with standard instructions), but not with the modified instructions in Experiment 4. After the experiments, subjects were asked to identify the payoff-maximizing level of contributions in the virtual treatment, and were <u>paid</u> for correct answers (see online appendices). Since Experiment 1 included decision tasks with an MPCR of 1, we asked for the profit-maximizing contribution associated with the last decision task in this experiment, which had an MPCR < 1. In Experiment 1, 28.1% of respondents gave an incorrect answer. This may be best interpreted as a lower-bound estimate as players may have only realized the dominant strategy after being asked.<sup>15</sup> Using as an upper-bound the percentage who did not free-ride in all rounds except when MPCR=1, as much as 64.6% did not understand incentives. Thirty percent of the subjects failed to identify the dominant strategy in Experiment 3 after 50 rounds of play (for these players, mean response is 28

<sup>&</sup>lt;sup>15</sup> We also asked participants in the all-human treatment to answer the dominant-strategy question and 25% answered incorrectly. These results provide further evidence that the virtual player design does not induce confusion over and above that already present in the all-human setting.

tokens and median response is 25 tokens).<sup>16</sup> Again, this is a lower-bound measure, especially given that many subjects herded to zero contributions by Round 50. Similar to Experiment 3, 32.1% of participants answered incorrectly in Experiment 4 with standard instructions. However, <u>no</u> participants in the modified instructions treatment answered incorrectly.

So then, what motivated behavior in virtual-player groups? The evidence suggests that it was profit-maximization. Using a Wilcoxon matched-pairs signedranks test, we fail to reject the hypothesis that stated profit-maximizing contributions and actual contributions are equal for the selected decision task in Experiment 1 [z=0.65, p=0.51], or, using average contributions from the last five virtual-player rounds, in Experiment 3 [z=0.51, p=0.61]. Given that participants in Experiment 1 received no information on virtual player decisions prior to answering the dominant strategy question, some additional insight can be gleaned through further analysis of these responses. The correlation coefficient for stated and actual contributions is large and statistically different from zero for all subjects [ $\rho=0.51$ , p<0.01] as well as for the subset of subjects who failed to identify the dominant strategy [ $\rho$ =0.52, p<0.01]. Further, the vast majority of virtual-player treatment contributions, 70.0%, come from subjects who failed to identify the profit-maximizing level of contributions. That this percentage is not 100% is not alarming, given that some subjects confused during the experiment may simply guess the correct answer or may have only realized the correct answer after being asked about it. The statistics together suggest that subjects in the virtual-player treatment were attempting to maximize profits and were not instead driven by non-monetary motives.<sup>17</sup>

We use similar arguments to eliminate the last two rival explanations. If subjects did not understand they were playing with non-human players, or did understand but nevertheless played as if they were matched with other humans, we should <u>not</u> see a dramatic change in contributions with modified instructions in Experiment 4. These instructions use the same language to describe virtual players as used in Experiments 1-3. Moreover, references to donations and information about relative payoffs should, if anything, feed the supposition that subjects were instead matched with other humans or were expected to play as if they were. Yet we observe that few subjects contribute to the virtual players with our modified instructions (but many contribute when grouped with humans).

 $<sup>^{16}</sup>$  We did not ask this question in the first three sessions (n=60). We added the question only after being surprised by how many individuals were contributing in the last round of the virtual-player treatment.

<sup>&</sup>lt;sup>17</sup> Rival explanations based on altruism toward the experimenter or social pressure from the experimenter would also have to explain why altruistic subjects believe the experimenter would have preferred to receive the wrong answer rather than pay out for the correct one.

Moreover, we asked all subjects "True or False" questions about the nature of the virtual group members (e.g., "The members of your group were human beings who received money from your investment in the Group Exchange") and about the exogeneity of the decisions made by the group members (e.g., "You were able to affect how much the Virtual Players invested in the Group Exchange by changing your investment."). In Experiments 1 and 3, we asked these questions after subjects had made their decisions. Because of concerns that post-experiment questionnaires do not necessarily capture predecision understanding, we asked these questions *before* subjects made their decisions in Experiments 2 and 4 (no subject was allowed to continue if they answered the question incorrectly).

In Experiment 1, three virtual-player treatment participants (6%) answered that their group members were human<sup>18</sup> and three (6%) answered that they could affect the decisions of their virtual player group members. For Experiment 3, these same figures are just 3% and 1%.<sup>19</sup> Thus, evidence from the questionnaire strongly suggests that participants understood the role of virtual players (and our results do not pivot on the inclusion/exclusion of the few participants who answered the questions incorrectly). One might discount data from questionnaires, but the two rival explanations would require one to assume that subjects understand the game form but not the nature of the other players, and then when asked in a post-experiment questionnaire, subjects instead affirm they understand the nature of the virtual players but fail to answer the question of their own-payoff maximizing contribution correctly (even though they would be paid for correct answers).

Finally, the last two rival explanations also require subjects in the focus groups to hide, en masse and without explicit coordination, the fact that they played as if the virtual agents were human or did not care that they were nonhuman, and instead offer arguably more embarrassing explanations (i.e., not understanding, mimicking robots, or believing that their payoff increased with their own and others' contributions). Finally, with regard to an explanation arguing that subjects played as if they were in a video game, only two of our four experiments were played on a computer (in Experiment 2, the virtual agents were just pieces of paper in the subject's envelope). Importantly, one of the computerized experiments was Experiment 4, in which few subjects contribute when grouped with virtual agents. As in any scientific study, we cannot prove

<sup>&</sup>lt;sup>18</sup> These subjects may have believed the question was asking about the source of the virtual player contributions, which was human, rather than the nature of the virtual players, which was nonhuman.

<sup>&</sup>lt;sup>19</sup> In Experiment 3 (virtual-player only), one subject answered the "human" T/F question incorrectly and when prompted to reread the question, changed his answer. Another subject answered the "pre-determined" T/F question incorrectly and when prompted to reread the question, changed her answer.

with certainty that our results do not arise from an alternative explanation, but we believe we have carefully designed our experiments to reject the most plausible rival explanations.

#### VIII. CONCLUSION

Decision errors, confusion, and noisy behavior are familiar concepts in experimental economics. However, as suggested by Hey (2005, p.325): "...the source and possible nature of the noise are rarely explicitly discussed.... If one makes the wrong assumptions about the ... noise, then one usually makes wrong inferences from the data." Although public goods experiments have identified "confusion" – behavior that stems from subjects' inability to discern the form of the game in which they are playing – as a substantial source of contributions, most studies either ignore this confusion or treat it as no more than (random) statistical noise that disappears through repetition.

To elucidate the structure of confusion, we use experimental designs combined with microeconometric models that place structure on the decisions of confused players. We demonstrate that confused subjects use experimental parameters and (in repeated games) the behavior of other subjects as behavioral cues. Confusion behaviors mimic other-regarding behaviors and thus confound analyses by, for example, distorting estimates of altruism or the effects of increasing the MPCR. This confusion does <u>not</u> disappear with repetition, as suggested by some, nor is it picked up in the noise parameters of quantal response models.

We then show that modifying the standard instructions used in linear public goods experiments can substantially reduce confusion, which can, in turn, lead to changes in the distribution of contributions: an increase in free-riding and in full-endowment contributions, and a decrease in the splitting of endowments across private and group exchanges. The latter has been suggested by Ledyard as indicative of confusion. In addition to the dramatic reduction in endowmentsplitting, two other metrics support our contention that subjects in the modifiedinstructions treatment are less confused. In stark contrast to subjects in the standard-instruction treatment, most subjects in the modified-instruction treatment free-ride when grouped with non-human, pre-programmed players (i.e., virtual players). Moreover, unlike subjects in the standard-instruction treatment, all subjects in the modified-instruction treatment identify, on a post-experiment questionnaire, free-riding as the payoff-maximizing strategy when playing with virtual players. These latter two results also corroborate our claim, and those of others (e.g. Houser and Kurzban, 2002), that virtual players can be used to distinguish between confusion and other motives for contributing in public goods

experiments. Given the evidence from this study, two important questions emerge.

How would behavior in public goods experiments change if participants were no longer confused? Although our study provides some insight, much more investigation of the impact of our modified instructions in the repeated-round VCM and other experimental designs, with larger sample sizes, is warranted. Our modified instructions required about 15 minutes of additional explanation and exercises. Focus groups and questionnaires are likely to prove valuable for further instruction development that can be adapted to a variety of public goods and related games.

Do subject misperceptions affect the external validity of laboratory public good experiments? Although this study focuses on internal validity, the pervasiveness of confusion in the laboratory is a methodological shortcoming, and as such can be argued to decrease the external validity of laboratory experiments (Loewenstein; Levitt and List). Two recent studies appear to support this notion. Benz and Meier (2008) find an overall positive, but weak correlation of individuals' donations across laboratory and field settings. Further, they find many individuals that never contribute in a field setting contribute substantially in the laboratory. Laury and Taylor (2008) report that subjects who are more 'altruistic' in the GHL experiment are *less* likely to contribute money to a naturally-occurring public good. Assuming similar rates of confusion between the Laury and Taylor study and our own, our evidence suggests that those confused individuals who look like altruists in the experiment may not be altruistic in a setting absent of confusion.

On a final note, given the simplicity of the VCM game and the ubiquity of abstract instructions that disassociate experiment decisions from the decisionsettings we are interested in studying, we believe our results have important implications for the burgeoning use of laboratory methods to test economic theories. When using abstract instructions it is likely that subjects will introduce their own context in order to make understandable the decision task. When subjects are confused about incentives and the parameters are changing as they make decisions, they infer that the parameter changes must be a signal that their decisions ought to be changing. Experiments using within-subject designs are more likely to experience this confound. In conclusion, we believe that an important area for future research in experimental economics will be to identify the source and nature of the noise in experimental games and to develop ways to reduce this noise when it appears to be an artifact of the experimental design rather than part of the decision process being studied. The B.E. Journal of Economic Analysis & Policy, Vol. 10 [2010], Iss. 1 (Contributions), Art. 53

#### REFERENCES

- Andreoni, J. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal* **100**(401): 464-477.
- Andreoni, J. 1995. Cooperation in Public-Goods Experiments: Kindness or Confusion? *American Economic Review* **85**(4): 891-904.
- Ashley, R., S. Ball, and C. Eckel. 2003. Analysis of Public Goods Experiments Using Dynamic Panel Regression Models. Working Paper, Department of Economics, Virginia Tech.
- Burlando, R.M. and F. Guala. 2005. Heterogeneous Agents in Public Goods Experiments. *Experimental Economics* 8: 35-54.
- Carpenter, J. 2004. When in Rome: Conformity and the Provision of Public Goods. *Journal of Socio-Economics* **33**(4): 395-408.
- Cason, T.N. and D. Friedman. 1999. Learning in Laboratory Markets with Random Supply and Demand. *Experimental Economics* **2**(1): 77-98.
- Charness, G., G.R. Frechette and J.H. Kagel. 2004. How Robust is Laboratory Gift Exchange? *Experimental Economics* **7**: 189-205.
- Chaudhuri, A., S. Graziano and S. Maitra. 2006 Social Learning and Norms in a Public Goods Experiment with Inter-generational Advice. *Review of Economic Studies* **73**(2): 357-380.
- Chaudhuri, A. and T. Paichayontvijit. 2006. Conditional cooperation and voluntary contributions to a public good. *Economics Bulletin* **3**(8): 1–14.
- Chou, E., M. McConnell, R. Nagel, and C.R. Plott. 2009. The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person 'guessing' game. *Experimental Economics* 12(2): 159-179.
- Cookson, R. 2000. Framing Effects in Public Goods Experiments. *Experimental Economics* **3**(1): 55-79.
- Cotten, S., P.J. Ferraro and C.A. Vossler. 2007. Can Public Goods Experiments Inform Policy? Intepreting results in the presence of confused subjects. In *Environmental Economics, Experimental Methods*, edited by T. Cherry, S. Kroll and J. Shogren. Routledge, pp.194-211.
- Cox, J.C. and V. Sadiraj. 2007. Social Preferences and Voluntary Contributions to Public Goods. *Public Finance Review* 35(2): 311-322.
- Croson, R., E. Fatas, and T. Neugebauer. 2005. Reciprocity, Matching and Conditional Cooperation in Two Public Goods Games. *Economic Letters* 87(1): 95-101.
- Davis, D.D. and C.A. Holt. 1993. *Experimental Economics*. Princeton, N.J.: Princeton University Press.

Ferraro and Vossler: Confusion in Public Goods Experiments

- Erev, I. and A.E. Roth. 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review* **88**(4): 848-881.
- Ferraro, P.J., D. Rondeau, and G.L. Poe. 2003. Detecting Other-regarding Behavior with Virtual Players. *Journal of Economic Behavior and Organization* **51**: 99-109.
- Fischbacher, U., Gächter S., Fehr E., 2001. Are people conditionally cooperative? Evidence from a Public Goods Experiment. *Economics Letters* **71**: 397-404.
- Fischbacher, U. and S. Gächter. 2004. Heterogeneous Motivations and the Dynamics of Free Riding in Public Goods. Working Paper, Institute for Empirical Research in Economics, University of Zurich.
- Goeree, J., C. Holt, and S. Laury. 2002. Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior. *Journal of Public Economics* 83: 255-276.
- Greene, W.H. 2003. *Econometric Analysis*, fifth edition. Upper Saddle River, N.J.: Prentice Hall.
- Hey, J.D. 2005. Why We Should Not be Silent about Noise. *Experimental Economics* 8: 325-345.
- Holt, C.A. and S.K. Laury. 2008. Theoretical Explanations of Treatment Effects in Voluntary Contributions Games. In C.R. Plott and V.L. Smith (eds.), *Handbook of Results in Experimental Economics*. Amsterdam: North Holland, pp 846-855.
- Houser, D. and R. Kurzban. 2002. Revisiting Kindness and Confusion in Public Goods Experiments. *American Economic Review* **92**(4): 1062-1069.
- Houser, D. and R. Kurzban. 2005. An experimental investigation of cooperative types in human groups: A complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences* **102**(5): 1803-1807
- Isaac, R.M., J. Walker, and S. Thomas. 1984. Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations. *Public Choice* **43**: 113-149.
- Johnson, E. J., C. Camerer, S. Sen and T. Rymon. 2002. Detecting Failures of Backward Induction: monitoring information search in sequential bargaining. *Journal of Economic Theory* **104**: 16-47.
- Keser, C., and F. van Winden. 2000. Conditional Cooperation and Voluntary Contributions to Public Goods. *Scandinavian Journal of Economics* **102**(1): 23-39.
- Laury, S.K. and L.O. Taylor. 2008. Altruism Spillovers: Are Behaviors in Context-free Experiments Predictive of Altruism Toward a Naturally Occurring Public Good? *Journal of Economic Behavior and Organization* 65(1): 9-29.

- Ledyard, J. 1995. Public Goods: A Survey of Experimental Research. In J.H. Kagel and A.E. Roth eds, *Handbook of Experimental Economics*. Princeton: Princeton University Press, pp. 111-194.
- Levitt, S.D. and J.A. List. 2007. What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World? *Journal of Economic Perspectives* **21**(2): 153-174.
- Loewenstein, G. 1999. Experimental Economics from the Vatange-Point of Behavioural Economics. *Economic Journal* **109**: F25-F34.
- Mason, C.F. and O.R. Phillips. 1997. Mitigating the Tragedy of the Commons through Cooperation: An Experimental Evaluation. *Journal of Environmental Economics and Management* **34**: 148-172.
- Oxoby, R.J. and J. Spraggon. Forthcoming. Ambient-Based Policy Instruments: The Role of Recommendations and Presentation. *Agricultural and Resource Economics Review*.
- Palfrey, T.P. and J.E. Prisbrey. 1997. Anomalous Behavior in Public Goods Experiments: How Much and Why? *American Economic Review* 87: 829-846.
- Plott, C. and K. Zeiler. 2005. The Willingness to Pay–Willingness to Accept Gap, the 'Endowment Effect,' Subject Misconceptions, and Experimental Procedures for Eliciting Valuations. *American Economic Review* 95: 530-545.
- Rege, M. and K. Telle. 2004. The Impact of Social Approval and Framing on Cooperation in Public Good Situations. *Journal of Public Economics* 88(7-8): 1625-1644.
- Saijo, T. and H. Nakamura. 1995. The Spite Dilemma in Voluntary Contribution Mechanism Experiments. *Journal of Conflict Resolution* **39**(3): 535-560.
- Roth, A.E. and I. Erev. 1995. Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior* **8**: 164-212.
- White, H. 1982. Maximum Likelihood Estimation of Misspecified Models. *Econometrica* **50**(1): 1-25.