The Effects of Single-Player Coalitions on Reward Divisions in Cooperative Games

Greg d'Eon, Kate Larson, and Edith Law

Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada
{greg.deon, kate.larson, edith.law}@uwaterloo.ca

Abstract. Cooperative game theory is a useful theoretical tool for describing reward divisions in cooperative scenarios. However, empirical work on transferable utility games has focused on the class of zero-normalized games, where players cannot earn any reward without forming a coalition. In this paper, we study how humans divide rewards when acting as impartial decision makers in cooperative games. We construct games that systematically vary the values of the single-player coalitions, and we use crowdsourced experiments to find how they impact people's reward divisions. Our results show that increasing the individual players' values has a significant impact on people's reward divisions, moving them in directions that are sometimes unrelated to the Shapley value. These reward divisions do not always satisfy the null-player or additivity axioms. Based on our results, we propose and evaluate descriptive models for human reward divisions.

Keywords: Cooperative game theory · Shapley value · Behavioural game theory.

1 Introduction

Cooperative game theory studies coalitional games, where a group of players can form coalitions, and the game describes how much reward every possible coalition could earn. One of the main questions in this field is: after the players have formed a coalition, how should the reward be divided among them? The most well-known division method is the Shapley value [15], which is the unique division that satisfies four fairness axioms of efficiency, symmetry, null players, and additivity. This value also has a simple interpretation: the amount of reward that each player receives is the average amount of value that they bring to the group. Alternative solution concepts [4, 8, 11–13] have also proposed modifications to these axioms that attempt to capture social or psychological aspects of cooperation.

Empirical studies in this field have focused on understanding how humans think about these games. The majority of this work has emphasized the bargaining process, developing detailed models that describe how humans form coalitions and make counter-offers in these games. Little research has examined how humans divide rewards in these games when they are impartial to the outcome. The most relevant work is [3], who studied the rewards that impartial "decision makers" gave to "recipients". Their results showed that people tended to select convex combinations of equal splits and the Shapley values, which only break the null player axiom. However, [3] only used zero-normalized games, where individual players cannot earn any rewards without forming coalitions.

In this paper, we investigate how humans select reward divisions in cooperative games. In particular, we focus on understanding whether humans place more importance on the values of the 1- or 2-player coalitions in 3 player games. To do this, we experiment with two sets of games where we fix the Shapley values and vary the marginal contributions in the 1- and 2-player coalitions. In the first experiment, we control the sizes of these marginal contributions, creating games where the rank-ordering of the players is implied by the solo coalitions, the pair coalitions, or both. In the second experiment, we take this idea further by constructing games where the solo and pair values imply different orderings of the players. Our results show that the 1-player coalitions' values have a larger impact on people's selected rewards than the 2-player coalitions. We use our data to identify conditions where people emphasize the individual values, test whether their divisions adhere to Shapley's axioms, and propose descriptive models for these rewards by evaluating existing solution concepts from the literature.

1.1 Related Work

The earliest study of human behaviour in cooperative games is [7], where 4 to 7 participants bargained face-to-face. This work found that players tended to split their rewards equally, and powerful players rarely took full advantage of their position. However, the main focus was on the bargaining process, such as the speed of the negotiations and the resulting coalition structures.

Most of the experimental work following [7] is characterized by two features. First, it places an emphasis on the bargaining procedure, with participants discussing coalitions and reward divisions while acting as players in the games. Second, it focuses on *zero-normalized* games, where players cannot earn any reward without forming a coalition. We refer to [6] and [9] for comprehensive surveys of this work. More recent studies have continued to focus on bargaining with more restrictive protocols [1, 10] or with computer agents [16].

One experiment [5] is notable for using non-zero-normalized games; they found that the Shapley value is a good fit when all 3 players form a coalition. However, their analysis also includes situations where only two players formed a coalition, making it difficult to evaluate the accuracy of the Shapley values.

[3] is the most relevant experiment to our work. In their experiment, three "recipients" earned baskets of items by answering trivia questions. These items increased in value when combined with other recipients' baskets (for example, making a pair of shoes). Then, impartial "decision makers" chose how to divide rewards between the recipients based on their baskets' values. They concluded that humans select convex combinations of the equal split and the Shapley value. To our knowledge, their work is the first where the participants dividing the rewards are impartial to the divisions. However, their games were zero-normalized, as the recipients' baskets were worthless alone.

2 Values for Cooperative Games

We begin by describing cooperative game theory concepts that we use to motivate our experiments.

2.1 Cooperative Games

A *transferable utility game* $G = (N, f)$ consists of a set of players $N = \{1, 2, \ldots, n\}$ and a characteristic function $f : 2^N \to \mathbb{R}$. This characteristic function assigns a reward $f(C)$ to each coalition $C \subseteq N$; we add the requirement $f(\emptyset) = 0$. In this paper, we restrict our attention to transferable utility games with $n = 3$ players, so we often refer to the characteristic function f as a "game". Also, we often write the set $\{i\}$ as i and the set $\{i, j\}$ as ij – for example, $C \cup i$ means $C \cup \{i\}$.

A player i's *marginal contribution* to a coalition $C \subseteq N \setminus i$ is $mc(i, f, C) = f(C \cup i) - f(C)$. Each marginal contribution is the amount of reward that the player brings by joining a coalition. Players i and j are *symmetric* if $mc(i, f, C) = mc(j, f, C)$ for all $C \subseteq N \setminus ij$, and player i is a *null player* if $mc(i, f, C) = 0$ for all $C \subseteq N \setminus i$. A game is *monotonic* if all marginal contributions are non-negative.

2.2 Values

A *value* is a function $v : \mathbb{R}^{2^N} \to \mathbb{R}^N$ that assigns a reward $v_i(f)$ to each of the players i in the game f. We will focus on *efficient* values, where $\sum_i v_i(f) = f(N)$ – all of the reward is allocated. Perhaps the simplest value is the *equal division value* $ED(f)$, where each player receives an equal fraction of the total:

$$ED_i(f) = \frac{f(N)}{n}.$$

The most celebrated value is the Shapley value [15], which is the unique value $Sh(f)$ that satisfies four axioms:

- *Symmetry*: if players i and j are symmetric in f, then $Sh_i(f) = Sh_j(f)$.
- *Efficiency*: the players' rewards sum to $f(N)$: $\sum_i Sh_i(f) = f(N)$.
- *Null players*: if player i is a null player in f, then $Sh_i(f) = 0$.
- *Additivity*: if f and g are two games, define a new game $(f + g)(C) = f(C) + g(C)$ for all coalitions C. Then, $Sh_i(f + g) = Sh_i(f) + Sh_i(g)$.

This value can be computed by rewarding each player the amount of value they bring to a coalition, averaged over all possible orders of building the coalitions:

$$Sh_i(f) = \sum_{C \subseteq N \setminus i} \frac{|C|!(n - |C| - 1)!}{n!} mc(i, f, C).$$

A number of modifications to the Shapley values have been proposed. One is the family of *egalitarian Shapley values* [2, 4], which is the set of convex combinations of the equal division and Shapley values

$$Sh^{\alpha}(f) = \alpha Sh(f) + (1 - \alpha) ED(f).$$

The parameter α is a measurement of equality: $\alpha = 0$ gives an equal division, while $\alpha = 1$ gives the Shapley value. Another is the *solidarity value* [11]

$$Sol_i(f) = \sum_{C \ni i} \frac{(n - |C|)!(|C| - 1)!}{n!} A^f(C)$$

where $A^f(C) = \frac{1}{|C|} \sum_{i \in C} mc(i, f, C)$ is the average marginal contribution of any player to C. Nowak suggests that the solidarity value captures some subjective psychological aspects, while the Shapley value is the "pure economic" solution.

2.3 Procedural Values

The family of *procedural values* [8,13] generalizes the egalitarian Shapley values. A procedural value $P^s(f)$ is parameterized by a tuple $s = (s_1, s_2, \ldots, s_{n-1})$. Each term in this tuple is a measure of equality: when a player joins a coalition of size k, they keep a fraction s_k of their marginal contribution, splitting the remaining fraction $(1 - s_k)$ equally among the other players. To simplify calculations, we denote $s_0 = s_n = 1$. These procedural values are

$$P^s(f) = \sum_{C \subseteq N \setminus i} \frac{|C|!(n - |C| - 1)!}{n!} \left[s_{|C|+1} f(C \cup i) - s_{|C|} f(C) \right].$$

To help understand the effects of varying the s parameters, we describe a method for decomposing a value into several components. First, we define the games f^k for $1 \leq k \leq n$ as

$$f^k(C) = \begin{cases} f(C), & |C| = k, \\ 0, & |C| \neq k. \end{cases}$$

Then, we define $d^k(f) = Sh(f^k)$. Each of these d^k vectors represents the differences between the players' marginal contributions in the coalitions of size k. Note that $d^n(f) = ED(f)$, and for $k < n$, $\sum_i d_i^k = 0$ – adding d^k to a value preserves efficiency. For example, consider the game $f(\emptyset) = f(2) = f(3) = 0$, $f(1) = f(13) = 30$, and $f(12) = f(23) = f(123) = 60$. This game has $d^1(f) = [10, -5, -5]$, $d^2(f) = [-5, 10, -5]$, and $d^3 = ED(f) = [20, 20, 20]$. This decomposition allows the procedural values for any game to be written as a vector sum; for a 3-player game,

$$P^s(f) = ED(f) + s_1 d^1(f) + s_2 d^2(f).$$

We also overload notation and write $d^{Sh}(f) = Sh(f) - ED(f)$ so the egalitarian Shapley values are

$$Sh^\alpha(f) = ED(f) + \alpha \, d^{Sh}(f).$$

We use procedural values to design our games and interpret our results in this paper for two reasons. First, the family of procedural values includes all of the values described previously: $ED(f)$ has $s_k = 0$, $Sh(f)$ has $s_k = 1$, $Sh^\alpha(f)$ has $s_k = \alpha$, and $Sol(f)$ has $s_k = \frac{1}{k+1}$. Second, [3] found the egalitarian Shapley values to be a good model for their results; procedural values are a natural way to generalize this idea to non-zero-normalized games.

3 Method

3.1 Games

For our experiments, we constructed games with identical Shapley values where we divided the players' contributions between the d^1 and d^2 vectors in different ways. First, we chose Shapley values that represent different rank-orderings of the players. We refer to these Shapley values as 1-Worse ($Sh = [25, 25, 10]$), 1-Better ($Sh = [30, 15, 15]$), Distinct ($Sh = [30, 20, 10]$), and 1-Null ($Sh = [40, 20, 0]$). Then, we chose d^1 and d^2 vectors in two different ways.

Experiment 1: For the 1-Worse, 1-Better, and Distinct Shapley values, we created 3 games by placing the players' marginal contributions in the Solo values or the Pair values, or splitting them between Both coalition sizes. These games have

$$\begin{aligned} \text{Solo:} \quad & d^1(f) = d^{Sh}(f); \quad d^2(f) = 0 \\ \text{Both:} \quad & d^1(f) = d^2(f) = \frac{d^{Sh}(f)}{2} \\ \text{Pair:} \quad & d^1(f) = 0; \quad d^2(f) = d^{Sh}(f) \end{aligned}$$

After fixing d^1 and d^2, the average values of the solo and pair coalitions are still unconstrained; we arbitrarily set them to make the games monotonic. We also added a purely additive game with $Sh = [10, 20, 30]$ and a symmetric game with $Sh = [20, 20, 20]$. These 11 games are listed in Table 1.

Experiment 2: For the 1-Worse and 1-Better Shapley values, we selected d^1 vectors that do not point towards these values. For 1-Worse, we used vectors of the form

$$d^1_{\text{1-Worse}} = [2x, -x, -x]$$

and for 1-Better, we used vectors of the form

$$d^1_{\text{1-Better}} = [x, x, -2x].$$

For each Shapley value, we selected 6 games using these vectors. In 3 of these games (Zeros2, Zeros5, and Zeros10), we gave values of 0 to some of the players and values of 2, 5, or 10 to the others. In the other 3 games (Sum30, Sum45, and Sum60), we chose values that summed to 30, 45, or 60.

We also selected four games with 1-Null Shapley values. In these games, we gave player 2 a reward of 0 (1-Null-Zeros) or had the individual rewards sum to 40, 50, or 60 (1-Null-Sum40, 1-Null-Sum50, and 1-Null-Sum60)). We also included the Symmetric game again. All 17 games are listed in Table 2.

3.2 Experiment

Participants: We hired participants from Mechanical Turk. For experiment 1, we posted human intelligence tasks (HITs) with the title "Divide rewards in fictional scenarios (10 mins)" with a payment of $1.25 USD. For experiment 2, we

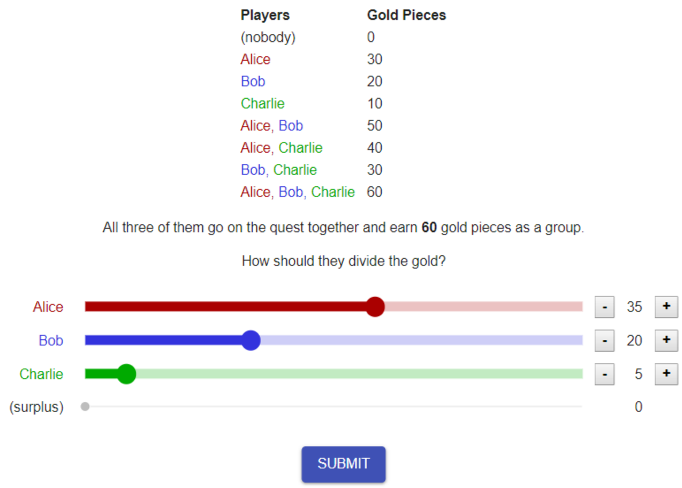

Fig. 1. The task interface. Participants were presented with a tabular representation of the game and asked to divide the total reward between the three players. The "submit" button was only enabled when the entire reward was allocated.

changed these values to 15 minutes and $1.75 USD. We required workers to have at least 1000 approved HITs with a 95% or higher approval rate. We restricted the HIT to workers located in the United States, and we used Mechanical Turk's qualification system to ensure that workers could only accept the HIT once.

Task: During the experiment, participants were presented with a series of scenarios about three fictional characters – Alice, Bob, and Charlie – playing a video game online. Each of these scenarios was associated with a cooperative game, which describes how many gold pieces every coalition could earn by working together. We displayed this information in a colour-coded table, which listed every combination of players and the amount of gold that the group could earn. Then, we told workers that the three characters all chose to work together, and we asked how the gold should be divided. Workers entered their responses by adjusting three sliders and clicking the submit button. The interface disabled the submit button as long as there was a surplus, only allowing efficient responses to be submitted. The experiment interface is shown in Figure 1.

Procedure: After workers accepted the HIT, they filled out a consent form and completed a brief tutorial. In this tutorial, we described the interface and asked comprehension questions about the reward displays. Then, workers completed several rounds of the task, with each round corresponding to one of the games above. We randomized the order of the games. We also randomly labelled players 1, 2, and 3 as Alice, Bob, and Charlie in each game. Finally, workers received a confirmation code and submitted the HIT.

4 Results

In both experiments, a total of 100 workers completed the HIT. A number of workers submitted low-quality answers (for example, [30, 30, 0] in the SYMMETRIC game). To remove these workers, we filtered out 21 workers that spent less than 5 seconds on any scenario. We also removed workers that repeatedly submitted nonsensical answers, such as $[1, 1, 58]$ in DISTINCT-BOTH. This filtering process left us with 75 workers in experiment 1 and 74 workers in experiment 2. We confirmed that this criteria was appropriate by checking the SYMMETRIC games: after filtering, the most extreme reward in this game was [20, 22, 18].

4.1 Experiment 1

The rewards that each participant submitted for each game are plotted in Figure 2. Each of these plots shows the distribution of selected rewards, along with the equal division (red) and the Shapley value (blue). There are several key features to note about these plots.

First, in all games, most rewards are close to the line between the equal division and the Shapley value. On this line, there are a few key points where most rewards land. The most common is the equal division, which was picked by at least 25 of the 75 participants in each game. The Shapley value also appears frequently. Other common points include rewards half or double the distance from the equal division to the Shapley value.

Second, the main difference between the games is the distance from each of the rewards to the equal division. For all three Shapley values, the SOLO games have the most extreme rewards, while the BOTH and PAIR divisions are generally more equal. For instance, in 1-BETTER-SOLO, 14 participants submitted rewards close to $[40, 10, 10]$; in 1-BETTER-BOTH, only 3 such rewards remained.

We confirmed this trend using non-parametric statistical tests. For each division, we calculated the L_1 distance to the equal division, and we compared these distances using Holm-Bonferroni-corrected Wilcoxon signed-rank tests. For all three Shapley values, we found a significant difference between the SOLO and PAIR games ($p < 0.001$) and between the the BOTH and PAIR games ($p < 0.01$). We also found a significant difference between the 1-BETTER-SOLO and 1-BETTER-BOTH games ($p < 0.001$). These results confirm that people gave more equal divisions in the PAIR games and more unequal rewards in the SOLO games.

4.2 Experiment 2

Each participant submitted a reward division for all 17 games. We split this data across two figures. Figure 4 shows the rewards for the 1-WORSE and 1-BETTER games; Figure 3 has the 1-NULL games.

These rewards show striking differences from the data in Experiment 1. First, the majority of the rewards that participants selected are not affine combinations of equal divisions and the Shapley values. Further, in the 1-WORSE and 1-BETTER games, the Shapley values are quite uncommon. In fact, in four of the

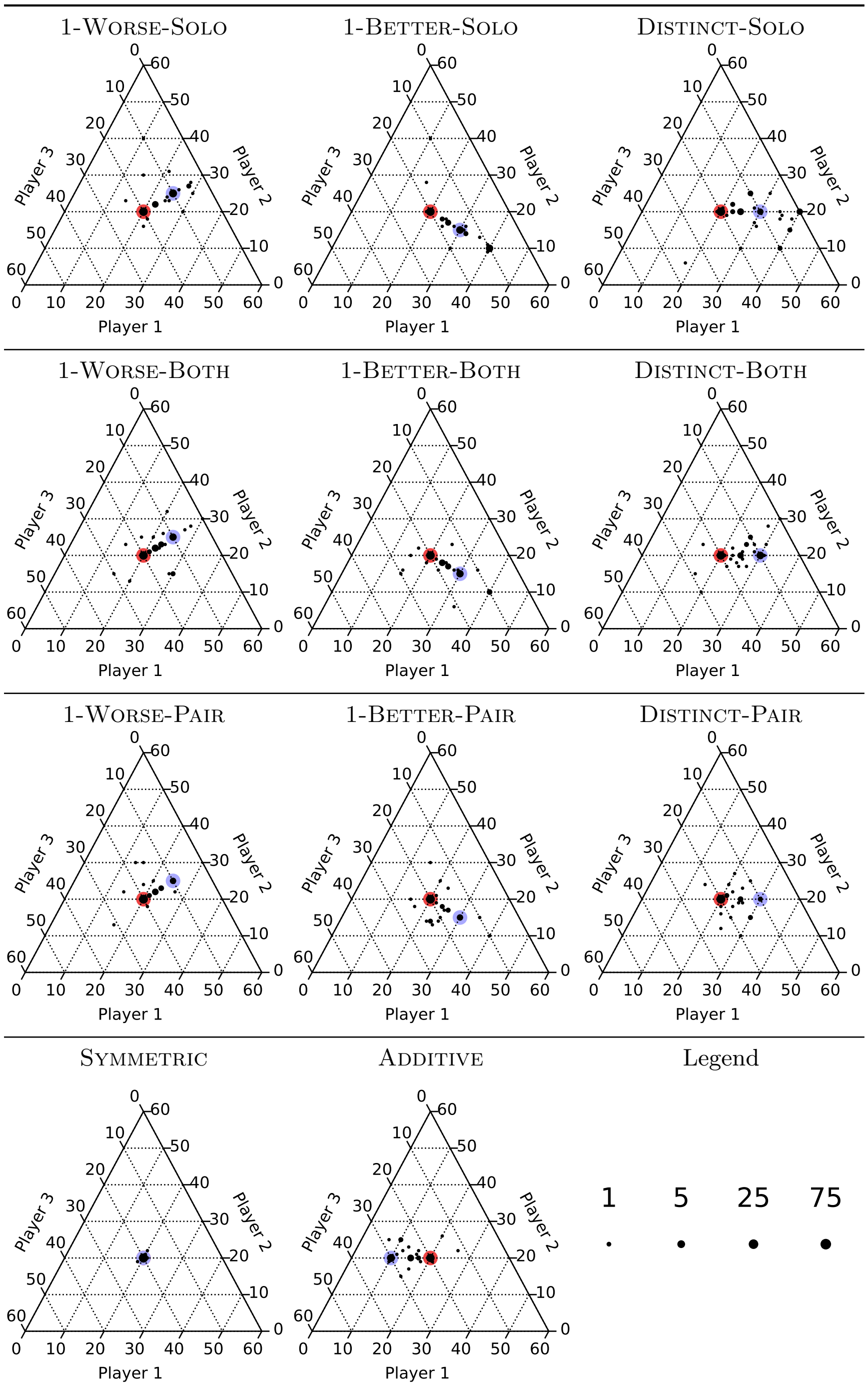

Fig. 2. The rewards that participants submitted for each game in Experiment 1. On each plot, $ED(f)$ is circled in dark red, and $Sh(f)$ is circled in light blue.

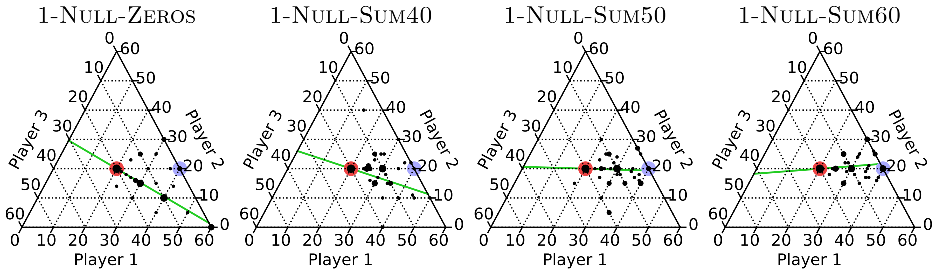

Fig. 3. The rewards that participants submitted for the 1-Null games in experiment 2. In each plot, $ED(f)$ is circled in dark red, and $Sh(f)$ is circled in light blue. Green lines indicate the direction of the main PCA component.

1-Better games, no participants chose the Shapley values. However, there is still a clear linear pattern to the rewards in most games. In the 1-Worse games, most of the rewards lie between the equal division and the value $[60, 0, 0]$; in the 1-Better games, they lie between the equal division and $[30, 30, 0]$. The 1-Null-Zeros game appears to be similar to the 1-Worse games, with many of the responses giving a disproportionately high amount of reward to player 1. Lastly, the other 1-Null games have more rewards close to the Shapley values.

We described these trends using principal component analysis (PCA). For each game, we computed the main principal component of the rewards. These components are plotted as green lines in Figures 3 and 4. Due to the high number of participants selecting equal splits in all games, we plotted these components as passing through the equal division. These components are highly consistent, with nearly identical directions in each 1-Worse game and in each 1-Better game. They also show the differences between the 1-Null games, where the components steadily shift from the extreme value in 1-Null-Zeros towards the set of egalitarian Shapley values in 1-Null-Sum50 and 1-Null-Sum60.

To make a formal comparison between the data and the Shapley values, we found bootstrapped 99% confidence intervals for the angles of each of these PCA components (see Appendix B for details). Only three of these confidence intervals – the three 1-Null-SumX games – contain $d^{Sh}(f)$. However, almost all of them contain the $d^1(f)$ vector; the only exception is 1-Better-Sum30, where it is $0.7°$ outside of the interval. This data strongly suggests that egalitarian Shapley values are not a good model for human-selected rewards in these games.

5 Discussion

Our results suggest that the single-player coalitions in cooperative games play an important role in people's reward divisions. In this section, we provide additional insights about these rewards, comparing our data against the Shapley value axioms and the rewards predicted by procedural values.

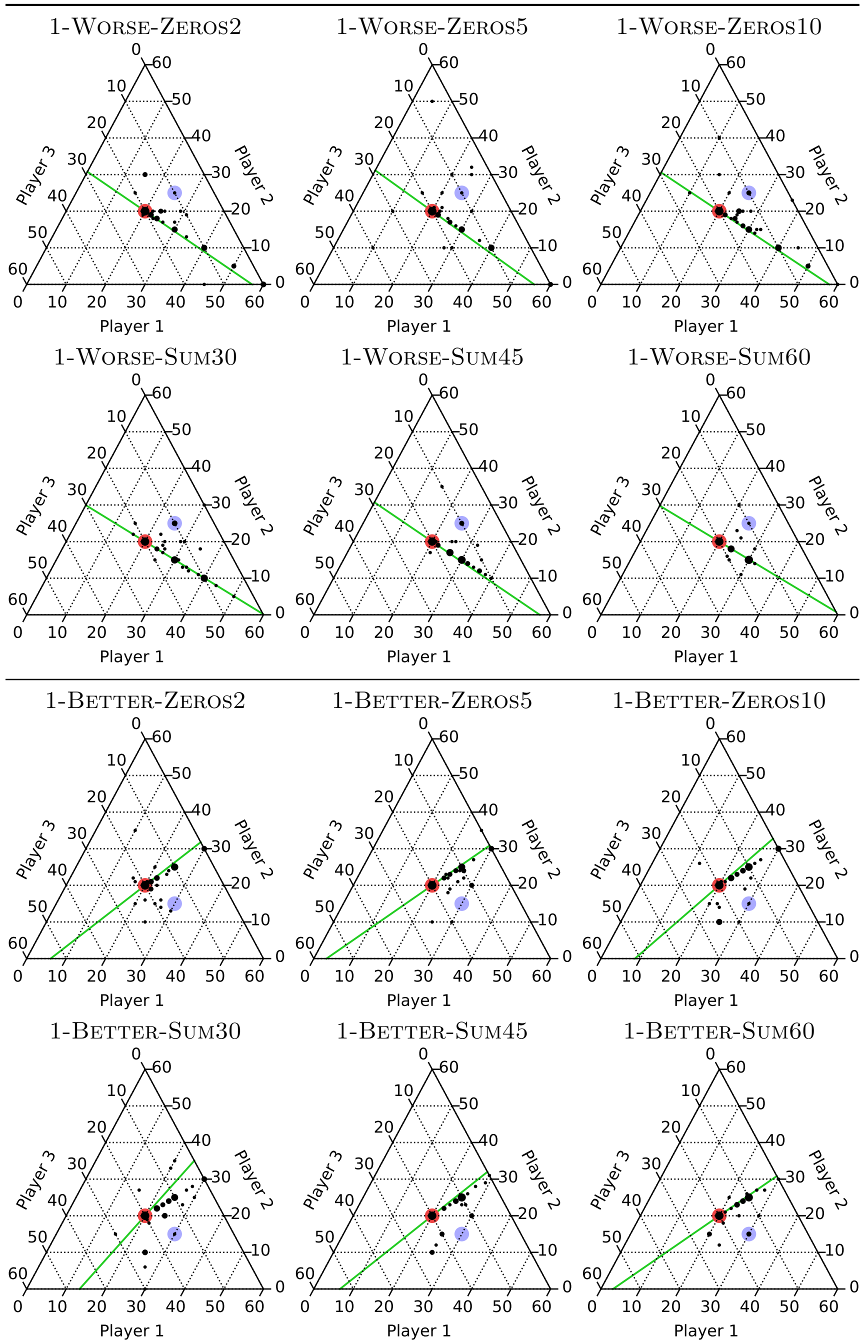

Fig. 4. The rewards that participants submitted for each of the 1-Worse and 1-Better games in Experiment 2. In each plot, $ED(f)$ is circled in dark red, and $Sh(f)$ is circled in light blue. Green lines indicate the direction of the main PCA component. In all 12 games, the PCA component is close to $d^1(f)$, but far from $d^{Sh}(f)$.

5.1 Shapley Value Axioms

In both of our experiments, we only allowed participants to submit efficient rewards, but we made no restrictions related to the other three Shapley value axioms. Did participants obey these axioms?

Symmetry: Six of the games in Experiment 1 have two symmetric players: players 1 and 2 in the 1-Worse games, and players 2 and 3 in the 1-Better games (Figure 2). We used paired Wilcoxon signed-rank tests to test whether these symmetric players received different rewards. We found no significant differences between these rewards in any of these six games (all $p > 0.1$). Thus, we cannot reject the hypothesis that participants obey the symmetry axiom.

Null Players: In all four of the 1-Null games in Experiment 2, player 3 is a null player. It is clear from Figure 3 that a majority of players give a positive reward to player 3, breaking the null player axiom. In the best case (1-Null-Sum60, only 14 of the 74 participants satisfied the null player axiom; in the other games, this proportion is even smaller. However, in each game, 30 to 40 participants gave a reward of 10 or less to the null player. While participants tend to recognize that null players contribute little to the group, they rarely go so far as to assign no reward to these null players.

Additivity: Several games in Experiment 2 are closely related. For instance, the 1-Worse-Sum30 and 1-Worse-Sum45 games only differ by the game

$$f(C) = \begin{cases} 5, & |C| = 1, \\ 0, & |C| \neq 1. \end{cases}$$

It is difficult to argue that any value other than $[0, 0, 0]$ is reasonable for f: $f(N) = 0$, and all three players are symmetric. Thus, to satisfy additivity, participants must select the same rewards for all three of the 1-Worse-SumX games and for all three of the 1-Better-SumX games.

We used 6 within-subjects Friedman tests to check for additivity violations. (For instance, one test checked whether participants assigned the same rewards to player 1 in the 1-Worse-Sum30, 1-Worse-Sum45, and 1-Worse-Sum60 games.) We found that the rewards for players 1 and 3 varied significantly in the 1-Worse games (both $p < 0.01$). We also found marginally significant results in the 1-Better games, with inconsistent rewards assigned to player 1 ($p = 0.08$) and player 3 ($p = 0.07$). These results imply that our data violates additivity.

5.2 Models for Human Rewards

Motivated by the concept of procedural values, we found that the values of the solo coalitions have the largest impact on human-selected rewards. Can procedural values accurately describe our participants' rewards?

For each participant, we searched for parameters (s_1, s_2) such that the predicted values differed from their rewards by no more than a threshold t for any player in every game. We tested all combinations of parameters $s_1 \in [-1, 2]$ and $s_2 \in [-2, 2]$ in steps of 0.01 with thresholds of $t = 2$ and $t = 5$.

We had little success fitting these procedural values to individual participants. In experiment 1, 25 participants submitted an equal division for every game, so they could be described by $s = (0, 0)$. However, few other participants could be described by any procedural value: we could only find good fits for 1 and 19 participants with $t = 2$ and $t = 5$, respectively. The remaining 31 participants could not be described by any procedural values. We found similar results in experiment 2. Here, 19 participants submitted equal divisions; otherwise, procedural values only described 2 and 8 participants at $t = 2$ and $t = 5$ respectively. We conclude that procedural values are generally not suitable for describing individual people's reward divisions.

Surprisingly, though, we found that procedural values describe the population averages for each game quite well. In experiment 1, we found a set of good s-values at $t = 2$; in experiment 2, we also found good fits at $t = 4$. One set of parameter values that fits both datasets is $s = (0.65, 0.25)$. However, neither set of parameters contains the egalitarian Shapley values or solidarity value.

There are several possible reasons why few participants can be described accurately by procedural values. One potential reason for this is that people may be performing a completely different type of calculation. People might be estimating the relationships between the players' power with quick, simple heuristics that are unrelated to the Shapley values – similar to Selten's equal division payoff bounds for bargaining [14]. Another possibility is that people are applying a non-linear utility function $u(x)$ to each of the game's values, then computing procedural values using the rewards $u(f(C))$ rather than $f(C)$. Finally, people might be concerned with stability, and their rewards could be affine combinations of equal payments and core-like solution concepts. This theory is difficult to test, as some of these games have multi-valued least-cores, but a focused experiment could investigate this idea.

6 Conclusion

In this paper, we studied how humans divide rewards when acting as impartial decision makers in cooperative games. Our results showed that the values of the single-player coalitions in these games, which have typically been fixed at zero in previous work, play an important role in people's reward division decisions. First, humans are more likely to select unequal reward divisions when the players in the game can earn different rewards alone. Second, the values of these single-player coalitions appear to take precedence over the two-player coalitions in many situations. We also used our data to show that humans respect the symmetry axiom, but not the null player or additivity axioms that are used to characterize the Shapley value, and their rewards are not described well by existing values in the literature. Our results serve as a starting point to research heuristics or utility functions that capture people's opinions in general cooperative games. Future work can also investigate how our results extend to other Shapley values, numbers of players, and representations of games.

References

1. Bolton, G.E., Chatterjee, K., McGinn, K.L.: How communication links influence coalition bargaining: A laboratory investigation. Management Science **49**(5), 583–598 (May 2003). https://doi.org/10.1287/mnsc.49.5.583.15148
2. Casajus, A., Huettner, F.: Null players, solidarity, and the egalitarian Shapley values. Journal of Mathematical Economics **49**(1), 58–61 (Jan 2013). https://doi.org/10.1016/J.JMATECO.2012.09.008
3. De Clippel, G., Rozen, K.: Fairness through the lens of cooperative game theory: An experimental approach (2013)
4. Joosten, R.: Dynamics, equilibria, and values. Ph.D. thesis, Maastricht University (1996)
5. Kahan, J.P., Rapoport, A.: When you don't need to join: The effects of guaranteed payoffs on bargaining in three-person cooperative games. Theory and Decision **8**(2), 97–127 (1977)
6. Kahan, J.P., Rapoport, A.: Theories of coalition formation. Psychology Press (1984)
7. Kalisch, G.K., Milnor, J.W., Nash, J.F., Nering, E.D.: Some experimental n-person games. In: Thrall, R.M., Coombs, C.H., Davis, R.L. (eds.) Decision Processes, pp. 301–327 (1954)
8. Malawski, M.: Procedural values for cooperative games. International Journal of Game Theory **42**(1), 305–324 (Feb 2013). https://doi.org/10.1007/s00182-012-0361-7
9. Maschler, M.: The bargaining set, kernel, and nucleolus. In: Aumann, R.J., Hart, S. (eds.) Handbook of game theory with economic applications, vol. 1, pp. 591–667. Elsevier (1992)
10. Nash, J.F., Nagel, R., Ockenfels, A., Selten, R.: The agencies method for coalition formation in experimental games. Proceedings of the National Academy of Sciences **109**(50), 20358–20363 (2012). https://doi.org/10.1073/pnas.1216361109
11. Nowak, A.S., Radzik, T.: A solidarity value for n-person transferable utility games. International Journal of Game Theory **23**(1), 43–48 (Mar 1994). https://doi.org/10.1007/BF01242845
12. Nowak, A.S., Radzik, T.: On convex combinations of two values. Applicationes Mathematicae **24**(1), 47–56 (1996)
13. Radzik, T., Driessen, T.: On a family of values for TU-games generalizing the Shapley value. Mathematical Social Sciences **65**(2), 105–111 (mar 2013). https://doi.org/10.1016/J.MATHSOCSCI.2012.10.002
14. Selten, R.: Equity and coalition bargaining in experimental three-person games. In: Roth, A.E. (ed.) Laboratory Experimentation in Economics: Six Points of View, pp. 42–98. Cambridge University Press (1987)
15. Shapley, L.S.: A value for n-person games. In: Kuhn, H., Tucker, A. (eds.) Contributions to the Theory of Games, pp. 307–317. Princeton Unversity Press, Princeton, 2nd edn. (1953)
16. Zick, Y., Gal, K., Bachrach, Y., Mash, M.: How to form winning coalitions in mixed human-computer settings. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. pp. 465–471. International Joint Conferences on Artificial Intelligence Organization (aug 2017). https://doi.org/10.24963/ijcai.2017/66

A Experiment Games

	Characteristic function								$Sh(f)$		
Condition	∅	1	2	3	12	13	23	123	1	2	3
1-Worse-Solo	0	40	40	10	60	60	60	60	25	25	10
1-Worse-Both	0	15	15	0	45	30	30	60			
1-Worse-Pair	0	0	0	0	45	15	15	60			
1-Better-Solo	0	40	10	10	60	60	60	60	30	15	15
1-Better-Both	0	15	0	0	45	45	30	60			
1-Better-Pair	0	0	0	0	45	45	15	60			
Distinct-Solo	0	40	20	0	60	60	60	60	30	20	10
Distinct-Both	0	20	10	0	60	50	40	60			
Distinct-Pair	0	0	0	0	60	40	20	60			
Symmetric	0	20	20	20	40	40	40	60	20	20	20
Additive	0	10	20	30	30	40	50	60	10	20	30

Table 1. The 11 games used in experiment 1 and their Shapley values.

	Characteristic function								$Sh(f)$		
Condition	∅	1	2	3	12	13	23	123	1	2	3
1-Worse-Zeros2	0	2	0	0	40	10	12	60	25	25	10
1-Worse-Zeros5	0	5	0	0	40	10	15	60			
1-Worse-Zeros10	0	10	0	0	40	10	20	60			
1-Worse-Sum30	0	20	5	5	60	30	45	60			
1-Worse-Sum45	0	25	10	10	60	30	45	60			
1-Worse-Sum60	0	30	15	15	60	30	45	60			
1-Better-Zeros2	0	2	2	0	38	40	10	60	30	15	15
1-Better-Zeros5	0	5	5	0	35	40	10	60			
1-Better-Zeros10	0	10	10	0	30	40	10	60			
1-Better-Sum30	0	15	15	0	45	60	30	60			
1-Better-Sum45	0	20	20	5	45	60	30	60			
1-Better-Sum60	0	25	25	10	45	60	30	60			
1-Null-Zeros	0	20	0	0	60	20	0	60	40	20	0
1-Null-Sum40	0	30	10	0	60	30	10	60			
1-Null-Sum50	0	35	15	0	60	35	15	60			
1-Null-Sum60	0	40	20	0	60	40	20	60			
Symmetric	0	20	20	20	40	40	40	60	20	20	20

Table 2. The 17 games used in experiment 2 and their Shapley values.

B PCA Confidence Intervals

We used PCA to find linear relationships between participants' rewards in our games. For our data, the main PCA component can be interpreted as a d vector: each datapoint x in our dataset lies on the plane $x_1 + x_2 + x_3 = 60$, so the main component must be parallel to this plane. We note that PCA is sensitive to outliers, but we apply it to our dataset because we manually removed most of these outliers.

We used bootstrapping to find confidence intervals for the angles of the PCA components. Specifically, to compute one bootstrapped estimate of the angle for a game, we sampled 74 points with replacement from our Experiment 2 dataset of 74 rewards. We calculated the main PCA component of this resampled data and found the angle of this component as it would be visualized on a ternary plot. We repeated this process 10000 times for each game to get a distribution of the PCA angles, and we took the middle 99% of these angles as the confidence interval. These intervals are shown in Figure 5.

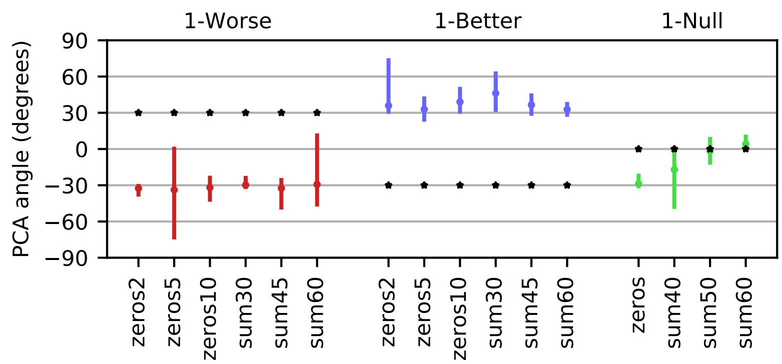

Fig. 5. Bootstrapped 99% confidence intervals for the angles of each of the PCA components. For each game, the angle of the vector $d^{Sh}(f)$ is indicated with a black point. All but one (1-Better-Sum30) of the confidence intervals contain $d^1(f)$; only the rightmost three intervals contain $d^{Sh}(f)$.