Algorithmic Mediation in Group Decisions: Fairness Perceptions of Algorithmically Mediated vs. Discussion-Based Social Division

Min Kyung Lee
Center for Machine Learning and Health, and
Machine Learning Department
Carnegie Mellon University
mklee@cs.cmu.edu

Su Baykal
Google Inc.
bsu394@gmail.com

ABSTRACT

How do individuals perceive algorithmic vs. group-made decisions? We investigated people's perceptions of mathematically-proven fair division algorithms making social division decisions. In our first qualitative study, about one third of the participants perceived algorithmic decisions as less than fair (30% for self, 36% for group), often because algorithmic assumptions about users did not account for multiple concepts of fairness or social behaviors, and the process of quantifying preferences through interfaces was prone to error. In our second experiment, algorithmic decisions were perceived to be less fair than discussion-based decisions, dependent on participants' interpersonal power and computer programming knowledge. Our work suggests that for algorithmic mediation to be fair, algorithms and their interfaces should account for social and altruistic behaviors that may be difficult to define in mathematical terms.

Author Keywords

Algorithms; decision-making; fairness; groups; fair division; collaboration

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION

Algorithms increasingly mediate groups and make social decisions that used to be made by humans. Algorithms allocate limited tasks to crowd-sourced workers and Uber drivers [22, 30]; they assign credit and rewards to virtual team members [12]; and they sort MOOC students and crowd-sourced workers into small work groups [48]. Emerging research on smart cities investigates ways to use algorithms to allocate different social resources to different neighborhoods [46]. Algorithms in these settings could aid efficient, data-driven decision-making for groups, yet we know relatively little about the impact of algorithmic mediation in this context. How might people feel about algorithmically "mediated" decisions for their groups? How might their perceptions of algorithmic decisions differ from their perceptions of decisions made through discussion? Would people feel that the algorithms' decisions were fair?

We examined people's perceptions of algorithmically-generated decisions for groups by conducting two studies using Spliddit [42], a website that applies fair division algorithms to social decisions. The site uses mathematically-proven fair division algorithms from the field of economics, which generally measure how much each individual values limited resources that a group needs to share and then compute division decisions for each individual.

To explore how people perceive algorithmic decisions, we first conducted a qualitative laboratory study using four division features on Spliddit for rent, tasks, credit, and goods. We then conducted a controlled experiment where a group of participants divided tasks (specifically, chores) that each had to complete in a kitchen laboratory, using either Spliddit or discussion as their method of making the division. The second study validates findings from the first study using actual tasks; it also provides a good baseline for estimating the performance of algorithmically-mediated group decisions and understanding how that differs from the performance of group discussions in division decisions.

The results suggest that even mathematically-proven fair division algorithms were thought to be less than fair one third of the time (30% for self, 36% for group). Algorithmic decisions were viewed as being unfair when the algorithm's assumptions of users did not account for multiple concepts of fairness and cognitive and social behaviors in groups, such as the presence of altruism and group dynamics. These factors can make perceptions of fairness differ from economic fairness. Decisions made through discussion were thought to be fairer. This effect depended on participants' interpersonal power and computer programming knowledge. The interviews suggest that participation in the

CSCW '17, February 25-March 01, 2017, Portland, OR, USA

ACM 978-1-4503-4335-0/17/03…$15.00
DOI: http://dx.doi.org/10.1145/2998181.2998230

process of discussion made participants responsible for the division outcome, allowing each group to decide what fairness meant to them, which may account for this effect.

Our work makes the following contributions to research on computer-supported collaborative work (CSCW): first, we offer a new understanding of how cognitive and social factors, intertwined with the assumptions of algorithms and their interfaces, influence perceptions of fairness in algorithmic group decisions; second, we offer an understanding of how algorithmic mediation differs from group discussion based on experimental comparison; finally, we consider the implications of our results for the design of algorithmic decision mediators that are not only efficient, but also fair from human perspectives.

RELATED WORK

Three threads of research motivate our work.

Technology for collaboration and group decision-making

A long and rich stream of research in CSCW, organizational behaviors, and management has investigated the practices and design of technological tools for collaboration and group decision-making. The literature highlights how technology can change surrounding social and organizational practices; at the same time, these sociotechnical factors collectively shape the adoption of the technology [3, 11, 35]. Researchers in CSCW have designed novel tools and interaction principles that help groups work better.

One subset of research most relevant to our paper concerns group decision support systems [2, 14, 23, 26]. These systems support group decisions by facilitating intergroup communication, highlighting potential group conflicts, or calculating the tradeoffs of different group decisions. Unlike this set of research, which focuses on systems that play supporting roles for groups making decisions, our inquiry focuses on algorithms that assume the roles of final decision makers, mediating group decisions by making division decisions for each individual.

Interaction with algorithmic technologies

Increasingly more attention has been given to understanding the social implications of algorithmic technologies. Recent studies have examined how people make sense of algorithmically-curated social media [16] and algorithmically-managed workplaces [30, 39]. A long stream of research on intelligent systems more broadly, such as recommender systems [9], automated systems [28], and robots [29, 34], is also relevant. However, to our knowledge, few studies have examined the contexts in which algorithms mediate group decision-making and determine final outcomes.

Fairness and fair division algorithms

Fairness has been a central interest of scholars for centuries and there exist multiple ontological, psychological, and mathematical perspectives on its definition. In the Cambridge Dictionary, fairness is defined as "the quality of treating people equally or in a way that is right or reasonable." When thinking about the meaning of "treating people equally," an important distinction to make is that between equality and equity. Numerical equality is defined by Aristotle as "treat[ing] all persons as indistinguishable, thus treating them identically" [21]. This definition is similar to modern definitions of equality, which assume that everyone is at the same level and therefore deserves to receive the same distribution of a total. Equity – called "proportional equality" by Aristotle, is "not concerned primarily with what the final distribution of some good is," but rather with "how the distribution respects the nature of the goods and certain features of the people between whom they are being distributed," – also a key feature of distributive justice [24]. It is therefore important to consider whether fairness is based on the equal distribution of resources, regardless of the people those resources are distributed to, or whether a distribution is only fair if it takes individual differences into account.

More recently, fairness has been defined mathematically [4, 33]. Many economists have made algorithms that supposedly guarantee fairness in social division problems. Many of these fair division algorithms assume equity, maximizing individuals' ultimate responsibilities equally given varying levels of each individual's baseline. Many of these algorithms have been validated theoretically and mathematically, but very few have been evaluated from social perspectives in the real world, with the notable exception of an anecdotal evaluation of a residency matching program [17].

STUDY 1: INDIVIDUALS' EXPERIENCE OF ALGORITHMICALLY MEDIATED GROUP DIVISION

We first empirically examine how people would use a system that algorithmically makes division decisions for groups. Our research questions ask how people use an algorithmic system that mediates group division and what factors influence their perceptions of the division outcomes. As we examine a system that applies fair division algorithms, perceptions of fairness are one of our core interests in answering these questions.

Methods

We conducted a qualitative laboratory study in which groups of 2-5 participants divided rents, house chores, snacks, or credit for a game outcome using the Spliddit website. Each group of participants did one division task in one experiment session. Each participant was interviewed independently after the task.

Spliddit website

Spliddit [42] is a non-profit, public website that uses algorithms to find fair solutions to everyday division problems using methods from research on fairness division in economics, mathematics, and computer science. The website was developed by Ariel Procaccia, a computer scientist, and his students at Carnegie Mellon University, with the goal of providing easy access to fair division methods; it currently offers division of rent, fare, credit, goods, and tasks. The website had attracted 40,000 users by April 2015, a mere five months after its launch, and reviewed positively in major press [5, 6, 44]. On Spliddit, a user can determine the options that the group needs to make the division. All users are then able to input their preferences for each option as individuals. Once everyone finishes inputting their preferences, they are presented with the results for the entire group. In our study, we focused on division of rent, fare, credit, and goods, as they require individuals' valuation and preferences as inputs, whereas Spliddit's taxi fare feature uses only the distance of travel as division criteria. In the following sections, we explain the different algorithms behind each of the four division features we studied. In all features, each individual's inputs are not shared with the other group members.

Division algorithms and input interfaces on Spliddit

Rent division. Spliddit's rent feature aims to help roommates fairly share rent when sharing an apartment or house by taking into account the maximum amount each individual would contribute to the rent for each room. The algorithm [18] claims to guarantee envy-free rent splitting, as each roommate is assigned to a room that he/she should feel is the best deal for him/her. Efficiency is also guaranteed, as the algorithm assigns rooms in a way that makes it impossible for one roommate to find a more beneficial room without making another roommate worse off. These properties are all guaranteed on the assumption that each roommate wishes to maximize the difference between how much he/she believes the room is worth and how much he/she has to pay for rent.

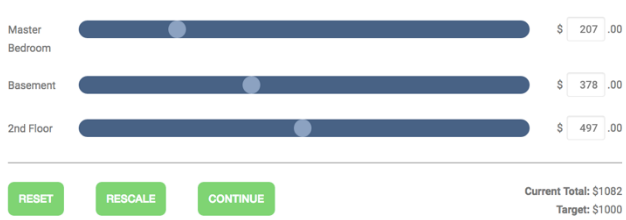

Figure 1. Spliddit rent input interface ©Spliddit

Users enter their preferences by choosing a point on a slider to indicate the maximum they are willing to pay for each room (Figure 1). The sum of the input should add up to the total rent for the whole house. The Spliddit algorithm then assigns each user both a room and a rent price.

Task division. Spliddit's task feature aims to fairly divide any set of tasks (i.e. household chores) among a group by gathering each participant's preferences for each task relative to the others. The algorithm [7, 38] guarantees equity, with each user believing their workload is identical. The efficiency property is also guaranteed in that the algorithm assigns tasks in a way that would make it impossible for one participant to find another more beneficial assignment without making another participant worse off.

Spliddit gets users' preferences through an interface that compares each one of the tasks to the others, one by one (Figure 2). Users are prompted to select the task they prefer between the "baseline" task and the comparison task. They then enter a multiplier indicating how many times they would be willing to do the selected task instead of doing the task they did not select.

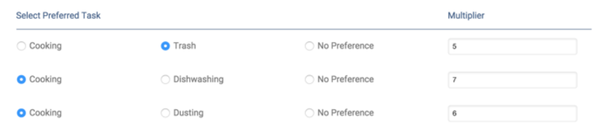

Figure 2. Spliddit tasks input interface ©Spliddit

Goods division. Spliddit's goods feature aims to fairly divide any kind of goods among two or more individuals [8, 27, 40]. In divisions between two people, the equitability property is met as long as both participants believe that their sets of goods have the same value. Envy-freeness is guaranteed as long as an individual's set of goods is at least as valuable as the other set and neither participant is willing to swap goods. The efficiency property is guaranteed in that the algorithm assigns goods in a way that would make it impossible for one participant to be assigned another more beneficial set of goods without making the other participant worse off. In divisions between three or more people, the envy-freeness property is guaranteed along with the maximin share fairness property. A person's maximin share is the amount that he/she could have given that he/she were allowed to divide the goods into sets, but the other participants were allowed to choose their sets before him/her. The algorithm guarantees that each person will receive at least ⅔ of his/her maximin share, with a greater likelihood that each individual will receive their total maximin share.

Figure 3. Spliddit goods input interface ©Spliddit

Users enter their preferences by choosing a point on a slider to indicate how much they value each item or "good" (Figure 3). Their input must add up to 1000.

Credit division. Spliddit's credit feature aims to fairly determine the contribution of each individual to a group project [10]. The algorithm guarantees impartiality, which means that a participant cannot affect his/her own share of credit – it can only be determined by the other participants' reports. The algorithm also guarantees consensus, which means that if all participants agree on each other's relative contributions, then credit is assigned according to a consensual division that is consistent with the inputs of all participants.

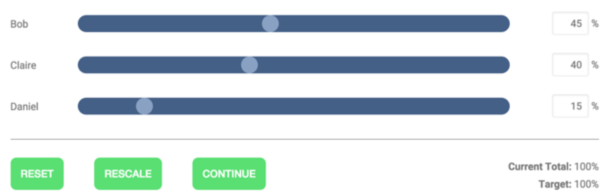

Figure 4. Spliddit credit input interface ©Spliddit

Users enter their preferences by choosing a point on a slider to indicate how much credit they think each user should get (Figure 4). They do not rate themselves. Their input must add up to 100%.

Task descriptions

In our study, we asked groups of two to five participants to make division decisions for rents, weekly house chores, snacks, or credit for a game outcome. The tasks of dividing rents and house chores were done in the setting of a hypothetical shared apartment, whereas the tasks of dividing snacks and credit were based on actual tasks in the laboratory.

Dividing rents. In the rent splitting condition, participants were asked to imagine that they were moving to a city and sharing an apartment with their group members. Participants used the Spliddit website to divide room assignments and rent amongst themselves. Experimenters selected apartments from Craigslist depending on total rent and variability between rooms, closets, and bathroom accessibility. Rent for apartments ranged from $1009 to $2599. Participants were then given a handout with a bird's-eye view drawing of the apartment, total price, and size. Additionally, participants were also given short descriptions of each room, including each room's size relative to the other rooms, the presence of a closet, and whether the bathroom needed to be shared with others. Individual inputs were then submitted in the Spliddit interface (Figure 1). The total values inputted for each room had to add up to the total valuation of the apartment.

Dividing house chores. In the tasks condition, participants were asked to imagine that they had recently moved to a city and were sharing an apartment with the other group members. They were instructed to split up house chores, which included cooking seven times a week, dishwashing seven times a week, emptying the trash once a week, and dusting once a week, between the members of their groups. Participants inputted their preferences for each task (Figure 2).

Dividing snacks. In the goods condition, participants divided snacks amongst each other. The snacks that were split included candy, gourmet chocolate, chips, and popcorn. Participants inputted their preferences by assigning numerical values to each snack (Figure 3). The total sum of the values was required to add up to 1,000.

Dividing credit for a game outcome. In the credit splitting condition, participants worked in groups of four and were instructed to guess an animal by asking the experimenter a maximum of twenty yes-or-no questions. They were given eight minutes to guess the animal as a group. The participants were then asked to assign credit to one another based on how much the success of the game could be attributed to each person's contributions. The group was rewarded $10 for their efforts, with the money being split according to the amount of credit assigned per person. They assigned credit percentages for the other three people in the group, not including themselves (Figure 4), with the sum of the percentages adding to 100%.

Discussion or no discussion before using Spliddit. For about half of the participants, we allowed them to discuss their preferences before they input their preferences into the Spliddit website. This was done to investigate perceptions of Spliddit's outcomes in the real world, where people may or may not choose to discuss their preferences before using the website.

Participants

We conducted the study at Carnegie Mellon University in November and December 2015. Participants were recruited through a participant recruitment website managed by the university and through flyers posted on campus. We ran 63 participants in 23 sessions. The data for 3 sessions was excluded due to a website glitch, no-show participants, or errors in website input, which left us 55 participants (M age=28.7 years (SD=12.9), 55% female) in 20 sessions for analysis. The rent-splitting task had 5 sessions (N=12), task-splitting had 5 sessions (N=11), goods-splitting had 8 sessions (N=24) and credit-splitting had 2 sessions (N=8). In half of the sessions, participants had a short discussion with the other participants about their preferences before entering them on the Spliddit website.

Participants had diverse ethnicities: there were 26 Caucasians, 26 Asians (or Pacific Islanders), and 3 African Americans. Participants recorded an average education level of 4.2 ("associate degree"=4, "bachelor degree"=5). Participants reported having basic concepts of programming on average (M=2.4 (SD=1.1)) and basic concepts of algorithms (M=2.0 (SD=1.1)).

Procedures

A study session took between 30 minutes and one hour, depending on the size of the group, and each participant was compensated $5. Participants worked in groups ranging from two to five people. Each session could accommodate up to five participants, and the group size was determined based on the total amount of participants that signed up and showed up to the session. Each session was assigned to one of four conditions for using the Spliddit website – rent, household chores, goods, or credit splitting. All participants signed the consent form prior to participating in the study. In some of the sessions, participants were asked to discuss their preferences for their assigned condition, while in others, no discussion was conducted prior to using Spliddit. In both cases, participants were given a handout with an explanation of how to use the Spliddit interface. After reviewing the handout, participants then added their inputs to Spliddit individually, without consulting other group members. Once all inputs were finalized, the researcher took photos of each participant's input for experiment records. Participants were then instructed to submit their inputs on the website and their results were displayed on the Spliddit website. After receiving results, participants were instructed to fill out a survey detailing their experience in using Spliddit. Finally, the experimenter conducted an interview with each individual in a separate room.

Measures

The input of each participant and the results produced by the Spliddit website were documented. The experimenter took pictures of each participant's input before they submitted it and received an email from Spliddit with group results.

Our survey measured perceptions of fairness related to the decision that was arrived at by Spliddit for themselves, other participants in the group activity, and the group as a whole. The survey also checked if participants knew each other before the study to control for possible effects of social proximity. Only a few participants knew each other, and as this did not influence our results, we excluded it in the analysis.

Fairness of self and group division outcomes. To measure perceived fairness of the participant's own results and others' results, we asked participants to indicate how much they agreed or disagreed, on a 7-Likert scale, with the statement, "My task assignment is fair," and, referring to each other group member, the statement, "This participant's task assignment is fair" [31]. We also measured their perception of the overall fairness of the results for the whole group by asking them to indicate how much they agreed with the statement, "The overall result for the group was fair."

Individual difference and demographic information. We asked about knowledge in computer programming and knowledge in computational algorithms as well as demographic information such as education level, age, gender, and ethnicity. To measure programming knowledge, we used the following 4-point scale: "No knowledge at all," "A little knowledge–I know basic concepts in programming," "Some knowledge–I have coded a few programs before," "A lot of knowledge–I code programs frequently." To measure knowledge in computational algorithms, we used the following 4-point scale: "No knowledge," "A little knowledge–I know basic concepts in algorithms," "Some knowledge–I have used algorithms before," "A lot of knowledge–I apply algorithms frequently to my work or I create algorithms."

Interview. We conducted 10-15 minute semi-structured interviews with each participant. We started with questions about the participants' perceptions of their results, asking if they were satisfied with their results, if they thought their results were fair and why, how they chose their input, and how they thought the Spliddit website had given them their results. Questions then probed their perceptions of other group members' results, if they thought members were satisfied with results, if they thought the results for the other members were fair, and how they thought group members had gotten those results. Lastly, we asked questions that looked more specifically into participants' perceptions of Spliddit and if they would use the website again.

Analysis

All interviews were recorded and transcribed. Three researchers qualitatively analyzed [36, 43] the interview transcripts. We first open-coded the transcript at a sentence or paragraph level to note factors that influenced participants' perceptions of the fairness of decision outcomes, which resulted in 67 concepts. We then synthesized the concepts into themes, which resulted in 4 high-level categories. The final coding scheme was reliable (Kappa=.71), and we used it to recode interview transcripts on Dedoose [13] to get proportions of different themes.[1]

Results

Overall, participants thought that algorithmic divisions for themselves were somewhat fair on average (M=5.04 (SE=.23)); they also thought that the overall division outcomes for the group were somewhat fair on average (M=4.95 (SE=.2)). Further analysis shows that 30% of the participants rated their own outcome 4 or less (out of the 7 point Likert scale) and 36% of the participants rated the outcome for the group to be 4 or less.

Multiple concepts of fairness

When the results reflected their input (in proportion to what they wanted), many participants said their results were fair (because the results reflected what they wanted) and vice versa. Some participants (34.55% of the overall) rated the fairness of their own outcome highly when their division outcomes mirrored their input. For example, participant 6A explained why they thought their results were fair with the following words: *"I think what I got, for the most part, reflected my preferences. [...]It's pretty close to the three numbers I put into it. (6A)"*

For others, their fairness perception depended on their general ideas about what fairness means in group divisions. For many participants (65.55% of all participants from 25

[1] We report the percentage in order to note relative frequency of different opinions and behaviors in our study. As a qualitative study with a small sample, we note that this should not be taken as an exact weight of whether one opinion is more significant or representative.

sessions), even distribution of individual division outcomes among group members was an essential factor for fairness. Even when their own or others' preferences were satisfied, many said that the objective values of options (such as monetary values or quantity of tasks) should be equally distributed. An exchange between participant 12C and the interviewer illustrates this: "*Interviewer: So I mean are you satisfied with your results? Interviewee: I mean it's nice that I got more I guess, but it's also not fair.*"

Some participants (9 participants, or 16.36% of all participants, from 7 sessions, or 35% of all sessions) considered preferences and even distribution as being equally important to fairness and felt that if one were lacking, the presence of the other could make up for it. As participant 12C stated, *"I know they both said they don't like dusting. So Kelsey has the dusting task. But the thing is she only has to do cooking two times a week with nothing else. I think that's fair for her I guess."*

Altruistic behaviors and social norms around division options

Some participants were willing to give up unevenly distributed results to help other participants. They were willing to make compromises and forgo some of their own preferences. Some mentioned that they value the happiness of the group; others mentioned being willing to accommodate people who had stronger feelings about certain options; others wanted to respect older adults, others' financial constraints, or others' reasons for their preferences. For example, one participant (17A) mentioned that he would want to know the reasons why different people want certain options, so that he could compromise and alter his share depending on the reasons. However, the Spliddit website does not accommodate these behaviors.

Heuristics and biases in quantifying subjective preferences.

Another factor that led to unfair outcomes from Spliddit was the presence of heuristics and biases, and social behaviors that influence how people quantify subjective preferences, which made the input to the algorithms error-prone. To translate their subjective preferences into numbers, all participants used various heuristics to anchor and adjust their input [25]. Some participants first ranked different options and assigned proportional numbers; others divided the total values "evenly" by the number of options and adjusted the results based off projected values or preferences. In this process, some participants translated strong preferences with rather extreme numbers, (i.e., I would rather cook 10 times than dust once versus I would rather cook two times than dust once), which skewed the overall group division outcomes.

While the fair division algorithms assume that individuals' input reflects individuals' "true" preferences, participants used the input interfaces in different ways. Some participants attempted to accurately express their own preferences, but other participants strategized their input in order to increase their chance of getting their most desirable division outcomes, to help other participants get what they wanted, or to avoid competition. The strategized input did not always return the outcomes that they desired, as the outcome depended on other participants' input as well, and it was difficult to predict what each group member would actually input into the system.

The designs of some of the input interfaces were conducive to potential biases and did not always embody accurate assumptions about users. For example, the task-splitting interface used anchoring and adjustment methods and randomly chose a baseline task for comparison. The fact that all task-splitting options were compared to one baseline task could skew participant preferences, especially when participants had strong preferences toward the baseline task. For instance, participant 12B said: *"I really don't like dishwashing at all, so I pretty much did not select that one any time it came up. I don't mind tidying the living room if it's just once, so I put a three there.*"

Participant 19B pointed out: "*I mean the process can be unfair, but the calculation is – I think it's fair. Interviewer: Oh okay. What do you mean by process? Interviewee: The process like being a multiplier thing, and you have to rate everything against it [to be fair]. […] [A]lways I'm comparing dishwashing against something."*

In addition, the allocation of the same amount of total value units to all participants in goods-splitting and credit-splitting features assumed that each user cared an equal amount, even though some participants indicated that they lacked any strong preferences or were willing to reduce their overall input if someone had a really good reason or desire to have something.

Decisions mediated by algorithms vs. discussion

Without any prompt from the interviewers, most participants compared decision-making through Spliddit to discussion-based decision-making. An advantage of discussion was the transparency that it provided participants with. Knowing other participants' preferences helped participants understand if a result was fair or not, both for participants' own results and those of other participants. Knowing others' preferences through discussion and the ability to adjust results was also stated as an important factor in knowing how to make compromises for participants to make results fairer. 7 participants (12.72% of all participants) from 7 sessions (35% of all sessions) stated that they thought discussion would have allowed for more compromise. As participant 10E stated, *"We do our best to make people happy, and giving them what they want for the most part. And it's sometimes just – it's not an intentional thing but that's how it goes. And had it been human interaction instead of computer interaction we probably would have got candy and a box of Twizzlers. But with the computer there's no emotions in it. Just you put in whatever input you put in, and it just bam bam bam bam, doles it out and that's just how it goes."*

While discussion was seen as allowing compromises to make results more fair, Spliddit was seen as being more fair for its objectiveness and equal treatment of all participants. The fact that it was an algorithm was often seen as enough reason for it to be fair. 9 participants (16.36% of all participants) from 8 sessions (40% of all sessions) stated that they thought Spliddit was more fair because of its objectivity. As participant 6A explained, *"even if you're trying to be fair ultimately you're going to have your best interest in mind. But the computer doesn't."*

Spliddit was also seen as a mediator between individuals that may be in dispute or may feel uncomfortable communicating. It let participants honestly express their preferences without feeling embarrassed or uncomfortable, and this was seen to make the system more accurate as well. 14 participants (25.45% of all participants) from 14 sessions (70% of all sessions) felt that Spliddit was a mediator for negative or uncomfortable social situations. As participant 22B expressed, *"[N]obody would wanna say someone did less work than another person so I guess it would be more accurate to do it on a computer."*

Discussion

The results suggest that cognitive and social factors, intertwined with the assumptions algorithms make and their interfaces, influence perceptions of fairness in algorithmic group decisions. In the study we observed several points where algorithms' assumptions and the processes and interfaces that allow algorithms to interact with people did not accommodate multiple concepts of fairness, altruistic behaviors and norms, or the social psychology of users.

First, fair division algorithms are based on equity, or the "proportional equality" concept of fairness, which emphasizes maximizing each individual's preferences and needs along with the overall group's welfare. However, not all the participants expected proportional equality. Many participants expected numeric equality, and some emphasized the need for self-sacrifice and compromise. Some participants emphasized the process, wanting the ability to make sure that nobody is unsatisfied to level out the satisfaction and perception of fairness among all the group members.

Fair division algorithms make several assumptions: users will be rational actors seeking to maximize self-benefit (the returning utility); users will have the same intensity of preferences; users' inputs will reflect their true preferences; and as long as their own preferences are satisfied, they won't "envy" others' results. Not all the participants fit this description: some participants were biased in the way they quantified their preferences; some participants argued that they had weaker or stronger preferences than others; and some chose to strategize their input, which became salient when the resources were not divided into equal quantities.

The design of Spliddit imposed a rigid operationalization of what fairness means; it did not allow people to discover what fairness meant to each individual in the group and come to a consensus about what a "fair" decision on the task at hand might look like. It took an individual-centric perspective, asking people to input their preferences separately, and did not provide any social transparency into how satisfied other people in the groups were. This lack of social transparency seemed to decrease participants' fairness perceptions about the group outcome. It seems apparent that individual preferences are not based only on the individual; various social factors come into play. Because Spliddit did not realize those factors and incorporate them into the outcome, some users felt that the algorithm had not produced a fair outcome.

STUDY 2: COMPARING ALGORITHMICALLY-MEDIATED VS. DISCUSSION-BASED DIVISION

The first study shows how people's perceptions of the fairness of algorithmically-mediated division decisions depend on the tension between mathematical definitions of fairness and the harder-to-define social sense of fairness. In Study 2, we conducted a between-subjects experiment to compare how people judge algorithmically-mediated outcomes compared to group discussion-based division decisions. Discussion is a standard way of making group decisions; comparing algorithmic mediation to discussion serves as a baseline for gauging the performance of algorithmic division outcomes. The division of rents and tasks (household chores) in Study 1 was evaluated in a hypothetical setting, and, in Study 2, participants actually completed the tasks they were assigned by the division decision.

Research questions

We hypothesized that different forms of mediation would evoke different levels of perceived control over and trust in the decision-making process and resulting outcomes, as suggested in the qualitative findings from Study 1. Social justice and fairness literature suggest greater perceived control over and trust in the decision-making process increase people's fairness perceptions of outcomes [31]. Discussion is a social process which people with high interpersonal power trust and feel they can control, which could increase their fairness perceptions. On the other hand, algorithmic mediation uses a technological tool that people with relevant technological knowledge can better understand than those without the knowledge. This greater understanding could increase their trust in and perceived control over the process, which could increase their fairness perceptions.

H1. Participants with greater interpersonal power will perceive division outcomes derived through discussion as fairer than those mediated by algorithms.

H2. Participants with greater computer programming knowledge will perceive division outcomes mediated through algorithms as fairer than those derived through discussion.

Methods

We conducted a between-subjects experiment in which people divided chores among themselves to compare discussion-based versus algorithmically-mediated decisions.

Conditions

Algorithmic mediation condition. We used the same procedure that we used in Study 1 for the algorithmic mediation condition. Participants were given a handout with an explanation of how to use the Spliddit interface and input their preferences. Participants were told that their inputs would not be shared with the other participants. After reviewing the handout, participants put their inputs into Spliddit individually, without consulting other group members. Once all inputs were finalized, the researcher took photos of each participant's input for experiment records. Participants were then instructed to submit their inputs on the website, and their results were displayed on the Spliddit website.

Discussion condition. In the discussion condition, participants were given a list of the tasks they needed to divide (depending on their group size) and asked to discuss amongst themselves how to divide the tasks. The group discussion was audio-recorded and participants were given a time limit of ten minutes for finalizing their task assignments in order to keep the study running within the time limit. When the group indicated to the researcher that they had decided on task assignments, they were asked to report their tasks in order for the researchers to keep track of what each individual would be responsible for completing.

Chore division

We used house chore task division. Participants were told to prepare for a house party and divide the tasks that they needed to complete in a kitchen laboratory. To ensure that some tasks would generally be more desirable than others, we conducted a pilot survey where twelve participants ranked seventeen tasks varied in difficulty and predicted desirability. We chose the tasks that were most consistently low-ranked and high-ranked.

For groups of two to three individuals, five tasks were used: washing 11 dishes; making tea and coffee twice; mopping the floor; sorting trash into categories of aluminum, plastic, paper, and landfill; and sorting twenty academic papers that had been scattered on a table four times, for a total of 80 papers that needed to be sorted. Only the tasks of making tea and coffee and sorting academic papers could be further divided among participants–for example, one participant could make one pot of tea and coffee and the other could make another pot of tea and coffee for a total of two pots of tea and two pots of coffee. The academic papers could be split into four groups of twenty papers–for instance, two people could sort forty papers each for a total of eighty papers.

For groups of four to five individuals, three extra tasks were added to the five described above: Making freshly squeezed lemonade; sorting candies into different jars; and filling up five ice cube trays. Like the tasks of making tea and coffee and sorting papers, filling up the ice cube trays could also be divided between participants; for example, in a group of five, the task could be split so that each person fills up one ice cube tray.

Participants

We conducted the study at Carnegie Mellon University in March and April 2016. Participants were recruited through a participant recruitment website managed by the university and through flyers posted on campus. We ran a total of 103 participants and 33 sessions with 50 participants and 16 sessions in the Algorithmic Mediation condition and 53 participants and 17 sessions in the Discussion condition (M=24.06 (SD=6.44), 57% female). The sessions were conducted with either a small group (2-3 participants) or a large group (4-5 participants). Both the Algorithmic Mediation condition and the Discussion condition had 10 small groups. They had 6 and 7 large groups, respectively.

Participants had diverse ethnicities: there were 64 Asians (or Pacific Islanders), 29 Caucasians, 6 African Americans, 1 Latino, and 1 Caucasian and Asian (1 participant selected "other" and 1 participant preferred not to answer). Participants recorded an average education level of 4.1 ("associate degree"=4, "bachelor degree"=5). The participants recorded having a mean of 2.6 (SD=1.1) in general programming knowledge and a mean of 2.1 (SD=1.1) in algorithm knowledge.

Procedures

Each session took between 1 hour and 70 minutes, depending on the size of the group, and each participant was compensated $10. Participants were asked to imagine that they all shared a house and were returning home from a vacation. Upon their return, they see that the kitchen and dining area is a mess and they need to make sure that it is clean in order to host a party for some friends. The tasks to complete have already been established, but the group needs to decide how to split up the tasks. The researcher gave a brief walk-through of all the tasks the participants would need to divide, accompanied by written directions for each task. Participants were then asked to split up the tasks either by inputting their preferences into Spliddit or by discussing and assigning tasks by themselves. All participants signed the consent form prior to participating in the study.

In all conditions, participants completed a total of three surveys and an audio-recorded interview. The first survey was taken after the group had finalized task assignments to determine their perceptions of fairness and satisfaction. The second survey was taken immediately after completing their tasks to determine if their perceptions of those attributes had changed. The third survey was taken after a brief interview that asked about individual differences and demographic information.

Measures

The input of each participant and the results produced by the Spliddit website were documented. The experimenter took pictures of each participant's input before they submitted them and received an email from Spliddit with group results.

Fairness of self and group division outcomes. We used the fairness measures that we used in Study 1. These questions were asked twice, once right after task assignment and once after actual task completion. Fairness perceptions did not differ between the two surveys, so we used the first survey in our analysis.

Individual differences and demographic information. The final survey looked into individual differences between participants and more specifically, interpersonal power. We used the interpersonal power scale and included two items ("I tend to lead a group discussion" and "I am less influential in group settings than others." (Cronbach's α=0.75) [19, 41]. We used the same questions in Study 1 to measure knowledge in computer programming and knowledge of computational algorithms, as well as demographic information.

Interview. We used the same set of questions we used in Study 1 in the Algorithmic Mediation condition. In the Discussion condition, we started with questions about the participants' perceptions of their task assignment, asking if they were satisfied with their assignment, if they thought their assignment was fair and why, and what they thought the main influencer in their task assignment was. Questions then probed the group discussion, looking into how the group came to a final decision, if they felt that everyone equally contributed to the discussion, and if they would change anything about the discussion. Questions then probed their perceptions of other group members' task assignments, if they thought members were satisfied with their tasks, and if they thought their assignments were fair.

Analysis

We analyzed the survey using a multilevel regression model to test the main effect of the condition and the interaction effect of the condition and individual differences in interpersonal power and computer programming knowledge [20]. We nested individual response into groups, and nested groups into conditions to control for groups [2].

To analyze the interview data, we took the qualitative approach used to analyze the interview findings from Study 1. The results from the analysis of the interviews in the algorithmic mediation condition were similar to those of Study 1. Thus we focus on reporting the results from the discussion condition.

Results

The results suggest that participants felt that divisions derived through discussion were perceived to be fairer, but that this impact depended on individuals' interpersonal power and knowledge of computer programming. The interviews helped us understand what might be contributing to this result.

Fairness perceptions of algorithmically mediated vs. group discussion-based outcomes

Overall, there was a main effect of decision-making medium. Participants thought that division decisions made through discussion were fairer than those mediated by algorithms (Figure 5a). Participants rated their own outcomes as more fair when they arrived at them through discussion (M=6 (SE=.2)) as compared to the algorithmic mediator (M=4.76 (SE=.2), $F(1,31.5)=18.6$, $p<.001$). Similarly, participants thought that the fairness of the overall group division was greater when devised through discussion (M=6 (SE=.22)) rather than through the algorithmic mediator (M=4 (SE=4.31), $F(1, 31.48)=29.6$, $p<.0001$).

Interpersonal power and overall group fairness

There was an interaction effect of interpersonal power on participants' perceptions of division outcomes' overall fairness for the group (Figure 5b) ($F(1, 91.82)=3.91$ $p=.05$), supporting Hypothesis 1. Participants with low interpersonal power had fairly similar judgments of fairness for the algorithmically mediated decision and the discussion-based decision. However, participants with high interpersonal power judged the decisions differently, seeing the discussion-based decision to be noticeably more fair than the algorithmically mediated decision. There was no interaction effect on participants' fairness perceptions of

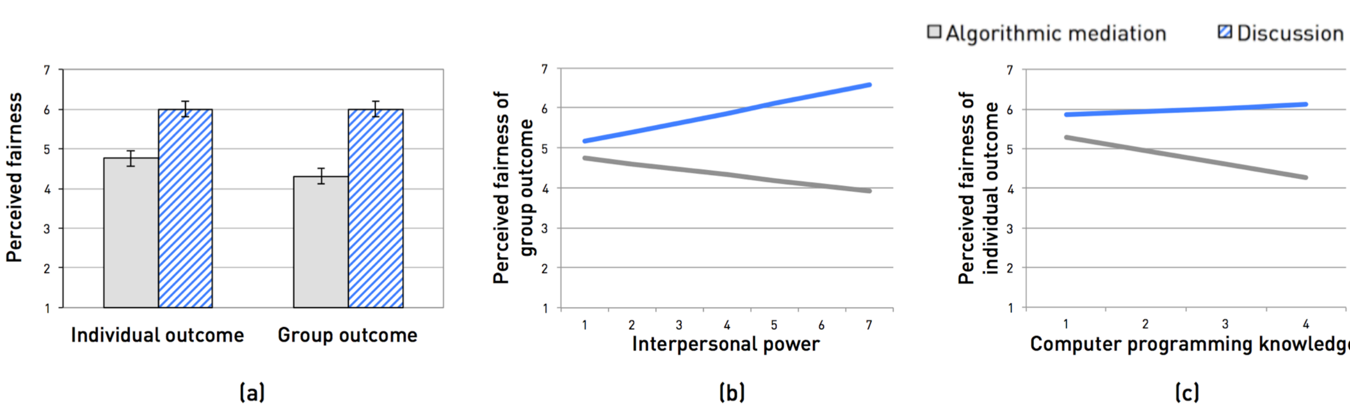

Figure 5. Fairness perceptions of participants' individual and group outcomes in the algorithmic mediation and discussion conditions

their own outcomes.

Knowledge of computer programming and own outcome fairness

Hypothesis 2 was not supported, and there was a marginally significant effect that shows the opposite effect of the prediction. There was a marginal interaction effect of knowledge in computer programming on fairness perception of individuals' own outcomes ($F(1, 95.5)=3.42$, $p=.07$) (Figure 5c). Participants with low computer programming knowledge saw the two decisions as fairly close in fairness, but as the level of computer programming knowledge grew, the algorithmically mediated decision was seen as increasingly less fair. There was no interaction effect of participants' programming knowledge on their overall group fairness perceptions.

The interview results provide some insight into why participants generally felt that decision outcomes derived through discussion were fairer.

Influence of choice and participation in the process on fairness

3 participants (5.66% of discussion participants) from 3 sessions (17.65% of discussion sessions) felt that their division outcomes were fair precisely because they had chosen and/or agreed to them. Even if their tasks ended up taking more time or were more difficult than those of others, participants would blame this on their own decisions. As participant 5A stated: *"I think it was fair, because I volunteered for it…. I think it's not anyone's fault, I would say, like how it turns out. I would say it's just like, oh, yeah, I kind of got the short end of the stick. No one knew that."*

This perspective also applied to other participants' results. 14 participants (26.42% of discussion participants) from 13 sessions (76.47% of discussion sessions) assumed that other participants perceived their own results as fair because they had accepted the tasks they were doing during their discussion. Referring to the other group members, participant 2A stated that *"they volunteered so they obviously didn't have a problem with those tasks so it seemed pretty fair all around."*

Social transparency through discussion

Discussion gave participants the opportunity to understand other group members' preferences as well as their responses to division outcomes. As previously stated, seeing that a participant had agreed to a task made them believe that the participant perceived the task assignment as fair and was satisfied with the task. Because discussion gave a clear understanding of others' preferences, participants were also able to make adjustments and compromises for other group members to increase overall fairness, even if this led to a less even distribution and went against the participant's own preferences. 14 participants (26.42% of discussion participants) from 12 sessions (70.59% of discussion sessions) made compromises for the rest of the group. As participant 10D stated, *"I looked at the dishes. It wasn't much. It could go quickly. It obviously wasn't a task that people are going to volunteer for anyway. So, I was like I'll just do it."*

Different fair division strategies emerged in each group

The process of arriving at a fair decision varied from group to group. In some groups (7 sessions, or 41.18% of discussion sessions), a few group members (13 participants, or 24.53% of discussion participants) would volunteer to do a few tasks and the rest of the remaining tasks were split between the remaining members. Participant 34C was one of the members who took one of the remaining options once others had volunteered. *"I was actually pretty quiet at first and just kinda let them pick things, and then, once I saw a couple things were gone, I picked washing the dishes 'cause it was still there, and it was better than sorting the trash. I guess. So I figured those would get left to be last."*

In other groups, one person would take the lead and distribute tasks. This was the case for 3 participants (5.66% of discussion participants) from 2 sessions (11.76% of discussion sessions).

These actions may have been caused by differences in interpersonal power. Those with higher interpersonal power may volunteer for tasks first or be more likely to take the lead during discussion. Those with lower interpersonal power were generally comfortable with someone else taking the lead, allowing them to stay passive during the discussion. Participant 28C was one of the participants who took the lead: "*I think just because I took lead in the beginning and I sort of made the point that these tasks take relatively the same amount of time – Because I made that I sort of asserted my credibility I guess. I was able to get my pick of tasks and I just chose what I thought I wanted."*

Across groups, even distribution and the minimization of time taken to do tasks was seen as essential to make the distribution as fair as possible, though groups had different strategies of achieving this. Some groups divided bigger tasks among group members (12 participants, or 22.65% of discussion participants, from 9 sessions, or 52.93% of discussion sessions). As participant 34D stated, *"I think that splitting up the papers was good, just 'cause that's like really boring and a big task."*

Other groups had each member take up one long and one short task (2 participants, or 3.77% of discussion participants, from 1 session, or 5.88% of discussion sessions). Participant 22C was in a group who used this strategy: *"The way that we split up tasks was we decided that four of the tasks would take a longer time and four of them would take a shorter time. Everyone picked one long and one short task."*

Different groups took different factors into account to find a task distribution that seemed fair to them. Some groups focused on their preferences for certain tasks while assessing the fairness of the task they had gotten. As participant 28D stated, *"I think at that point I guess it just* *became preferential. So whatever you thought was something you liked to do you chose that."*

Many groups focused on their experience and skill while assigning a task to minimize time spent doing the task and to increase fairness (16 participants, or 30.19% of discussion participants, from 12 sessions, or 70.59% of discussion sessions). As 10A stated when explaining his task choice, *"I guess lemonade jumped out at me because I've juiced a lot of lemons in the past. I've worked in kitchens. I feel like I can juice lemons pretty fast."*

Discussion

The results of Study 2 suggest that the issues surrounding algorithmic decision-making seen in Study 1, in which Spliddit was used to divide hypothetical tasks, are observed in and validated with real tasks. In addition, the results suggest that overall, participants thought that decisions made through discussion were fairer than those mediated by algorithms. The interview results largely center on the importance of the level of participation in the way participants perceive the fairness of their outcomes, which is related to previous literature on social justice. The autonomy that group members have in discussion allows each group to decide what rules, factors, or principles they want to incorporate into their decisions. While the process might be influenced by people's personalities or interpersonal power and the social dynamics of the group, the discussion makes the decision process transparent and gives people an opportunity to intervene or voice objections during the process. This makes people more accountable for the division outcome, and influences their fairness perceptions [15].

Participants perceived the discussion-made outcomes as fairer when they had high interpersonal power, as predicted in Hypothesis 1, but only for the group outcomes. Interestingly, participants' perceptions of the fairness of their individual outcomes were not influenced by their interpersonal power. This suggests that participants with higher interpersonal power might have sought to lead the discussion or volunteered to take tasks, often ones that they did not desire, thus playing a greater role in setting division criteria and rules for the group, but they were not necessarily trying to maximize their own benefit.

Computer knowledge had a marginal impact on individual outcome fairness in an opposite direction than that predicted in Hypothesis 2. We believe that participants with greater computer knowledge might have felt that they could control algorithmic decision outcomes through input adjustment, and were disappointed when the algorithms did not act in the ways they expected. This experience might have decreased their perceived control over the process, in turn decreasing their perceptions of fairness. It is also possible that participants with more computer knowledge might have known more about the limitations of algorithmic systems, which lowered their perceived control over and trust in algorithmic mediation. Further studies are required to unpack the mechanism.

LIMITATIONS & FUTURE WORK

Like any study, this paper has many limitations. We used a limited set of tasks in a group of strangers with a few types of fair division algorithms in our laboratory study. The findings from the study should be validated with different types of tasks, social contexts, and algorithms, eventually through a longitudinal study in the field. We used group discussion as a comparison for algorithmic mediation; other social decision-making processes, such as using facilitators, and variations in social power should be tested. We also examined the division problems in a relatively small group. Further research is needed to examine human perceptions of algorithmic decisions in larger groups, and to unpack the mechanisms that underlie the impact of psychological and social characteristics and technological knowledge of users on their perceptions of algorithmic decisions. We used a measure of interpersonal power in the study; future research could examine the role of other social constructs, such as team scales.

IMPLICATIONS

Algorithms are increasingly being introduced and incorporated as tools for governance in many different sectors of society [37]. We draw from the results of Study 1 and Study 2 to reflect on the potential unintended consequences of introducing algorithmic mediation into group decision-making. We then revisit the results, with the goal of informing algorithmic mediation that is not only efficient, but also fair from social perspectives.

Tensions between social and algorithmic decision-making

A rich stream of sociology and CSCW research has pointed out the tensions between the seemingly irrational, nuanced aspects of social behaviors and the simple, rational human behavior models commonly used by technological properties [1, 45]. Our research adds to this literature by showing that a similar tension emerges in people's social interactions with algorithmic technologies. Our studies suggest that the assumptions that algorithms hold about users – such as a desire to maximize self-interest – do not easily lend room to altruistic behaviors such as gifting, compromise, and sacrifice, which are critical elements of people's motivations and natures that help society function [32]. If economic-fairness division algorithms are embedded into the ways organizations and cities run, they may inadvertently promote interactions and decision-making that follow economic or mathematical models, and diminish the positive effects of altruism and other human behaviors that are not accurately represented in such models. This finding calls for more research on social and human perspectives on algorithmic technologies, and greater collaboration between fields such as artificial intelligence and human-computer interaction.

Materiality of algorithms

Our work suggests that we need to pay more attention to the materiality of algorithms. While biases of algorithms themselves have been the subject of much recent research, interfaces that embody and enact algorithms in situ have received relatively little attention. Our findings suggest that even algorithms mathematically proven to be "fair" may not achieve "fair" social division from human perspectives. Interfaces take input for algorithms, communicate their output, and direct how they become embedded in human practices. This process can sometimes negate the efficacy of algorithms, as in the case of Spliddit. Spliddit input interfaces did not help people better quantify their preferences in ways that fit with the algorithms' assumptions. The axioms and principles of the algorithms in Spliddit were explained on the website, but it was still difficult for participants to make sense of the algorithms' decisions, especially when the outcomes differed from their expectations. This work adds to research on the importance of algorithmic transparency and accountability [16, 47], and suggests that more research needs to be done to understand how to design algorithms with a level of transparency that is actually understandable to and useful for people.

Creating synergy between human and algorithmic decision-making

Our comparison between the algorithmic and discussion mediations raises a central question in designing algorithmic technologies: how can we combine the best parts of human and algorithmic decision-making? Spliddit uses mathematically proven definitions of fairness, makes decisions efficiently, and removes social influence that might bias the results. However, it also utilizes only one definition of fairness, requires groups of participants to fit into that definition, and fails to give participants any control over the decision-making process other than their input.

For services like Spliddit, participants can evaluate whether their assignments were fair, as their preferences were a key factor in the final decision and they were among the decision-makers. On the other hand, if algorithms are used in a different power structure – as when algorithms take on a managerial role in allocating incentives or budgets based on individual or project performances – it is much more difficult to refute algorithmic decisions, even when they feel unfair. In discussion groups, participants could organically decide what they believed was fair for their local context. They felt like they could influence the process by voicing their opinions, which resulted in greater perceptions of fairness of the discussed decision as compared to the algorithmically-made one. Yet discussion cannot be scalable for a very large group of people. How can we help people feel in control, influence decision-making principles, and negotiate the results in algorithmic mediation? One way could be to allow people to use algorithmic decisions as a basis for discussion; yet the process would need to be done very carefully to prevent any biased anchoring points. Further research is needed to find ways for individuals and groups to flexibly determine what factors algorithms need to account for and to negotiate the results.

CONCLUSION

In this paper, we explored fairness perceptions of divisions mediated by algorithms or discussion through two qualitative and controlled laboratory studies. The results suggest that even mathematically-proven fair division algorithms were thought to be less than fair one third of the time (30% for self, 36% for group). Algorithmic decisions were viewed as being unfair when the algorithm's assumptions of users did not account for multiple concepts of fairness and cognitive and social behaviors in groups, such as the presence of altruism and group dynamics, and when people's preferences and input through interfaces reflected biases and errors. These factors can make perceptions of fairness differ from economic fairness. Decisions made through discussion were thought to be fairer. This effect depended on participants' interpersonal power and computer programming knowledge. The interviews suggest that participation in the process of discussion made participants responsible for the division outcome, allowing each group to decide what fairness meant to them, which may account for this effect. The work suggests that algorithmic mediation in group decisions should account for social and altruistic behaviors that may be difficult to define in traditional mathematical or economic terms.

ACKNOWLEDGEMENT

We thank Daya Lee and Emma Shi who helped us collect and analyze data and participants who provided us with valuable insights.

REFERENCES

1. Mark S. Ackerman. 2000. "The intellectual challenge of CSCW: the gap between social requirements and technical feasibility." *Human-computer interaction 15, no. 2*, 179-203.
2. Robert Anson, Robert Bostrom, and Wynne Bayard. 1995. "An Experiment Assessing Group Support System and Facilitator Effects on Meeting Outcomes." *Management Science* 41, no. 2: 189-208.
3. Stephen R. Barley. 1986. "Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments." *Administrative Science Quarterly* 31, no. 1:78-108.
4. Steven J. Brams, Michael A. Jones, and Christian Klamler. 2006. "Better ways to cut a cake." *Notices of the AMS 53, no. 11*: 1314-1321.
5. Molly Brown. 2015. *Don't fight over the check: Spliddit helps you split expenses*. Geekwire. www.Geekwire.com.

6. Sydney Brownstone. 2014. “Finally, Computer Scientists Come Up with a Way to Split Rent that is Undeniably Fair.” Co.Exist.
7. Eric Budish, Yeon-Koo Che, Fuhito Kojima, and Paul Milgrom. 2013. "Designing random allocation mechanisms: Theory and applications." *The American Economic Review* 103, no. 2: 585-623.
8. Ioannis Caragiannis, David Kurokawa, Herve Moulin, Ariel D. Procaccia, Nisarg Shah, and Junxing Wang. 2016. “The Unreasonable Fairness of Maximum Nash Welfare.” In *Proceedings of the 2016 ACM Conference on Economics and Computation*, 305-322.
9. Enrico. Costanza, Joel E. Fischer, James A. Colley, Tom Rodden, Sarvapali D. Ramchurn, and Nicholas R. Jennings. 2014. "Doing the laundry with agents: a field trial of a future smart energy system in the home." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 813-822. ACM.
10. Geoffroy De Clippel, Herve Moulin, and Nicolaus Tideman. 2008. "Impartial division of a dollar." *Journal of Economic Theory* 139, no. 1: 176-191.
11. Paul Dourish and Victoria Bellotti. 1992. "Awareness and coordination in shared workspaces." In *Proceedings of the 1992 ACM conference on Computer-supported cooperative work*, pp. 107-114. ACM.
12. “Crowd Research.” 2016. Stanford HCI Group.
13. Dedoose http://www.dedoose.com
14. Gerardine DeSanctis and Brent Gallupe. “Group decision support systems: a new frontier.” *ACM SIGMIS Database*, 16(2): pp. 3-10.
15. Christopher P. Earley and E. Allan Lind. 1987. “Procedural Justice and Participation in Task Selection: The Role of Control in Mediating Justice Judgments.” *Journal of Personality and Social Psychology* 52, no.. 6: 1148.
16. Motahhare Eslami, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. 2015. "I always assumed that I wasn't really that close to [her]: Reasoning about Invisible Algorithms in News Feeds." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pp. 153-162. ACM.
17. Carl Erik Fisher. 2009. “Manipulation and the Match”. *JAMA*, 302(12):1266-1267.
18. Ya’Akov Gal, Moshe Mash, Ariel D. Procaccia, and Yair Zick. 2016. “Which Is the Fairest (Rent Division) of Them All?” EC-16: Proc. 17th ACM Conference on Economics and Computation.
19. John P. Garrison, and Larry E. Pate. 1977. "Toward development and measurement of the interpersonal power construct." *The Journal of Psychology* 97, no. 1: 95-106.
20. Andrew Gelman and Jennifer Hill. 2006. *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
21. Stefan Gosepath, (Spring 2011 Edition), "Equality", *The Stanford Encyclopedia of Philosophy* Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/equality/>.
22. Umair Ul Hassan, Sean O’Riain, and Edward Curry. 2013. "Effects of expertise assessment on the quality of task routing in human computation." *In Proceedings of the 2nd International Workshop on Social Media for Crowdsourcing and Human Computation*, Paris, France.
23. Joshua Introne. 2009. “Supporting group decisions by mediating deliberation to improve information pooling.” In *Proceedings of the ACM 2009 International Conference on Supporting Group Work:* pp. 189-198. ACM.
24. Harry Jones. 2009. *Equity in Development: Why it is important and how to achieve it*. London: Overseas Development Institute.
25. Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.
26. Nikos Karacapilidis and Dimitris Papadias. 1996 “A group decision and negotiation support system for argumentation based reasoning.” *Learning and Reasoning with Complex Representations*. Springer Berlin Heidelberg: pp. 188-205.
27. Flip Klijn. "An algorithm for envy-free allocations in an economy with indivisible objects and money." *Social Choice and Welfare* 17, no. 2 (2000): 201-215.
28. John D. Lee and Katrina A. See. 2004. “Trust in Automation: Designing for Appropriate Reliance.” In *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, no. 1: 50-80.
29. Min Kyung Lee, Sara Kiesler, Jodi Forlizzi, and Paul Rybski. 2012. "Ripple effects of an embedded social agent: a field study of a social robot in the workplace." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 695-704. ACM.
30. Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. “Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers.” In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems,* pp. 1603-1612. ACM.

31. Allan E. Lind, Ruth Kanfer, and P. Christopher Earley. 1990. "Voice, control, and procedural justice: Instrumental and noninstrumental concerns in fairness judgments." *Journal of Personality and Social psychology 59*, no. 5, 952-959.
32. Marcel Mauss. 2011. *The Gift: Forms and Functions of Exchange in Archaic Societies*. Martino Fine Books.
33. Hervé Moulin. *Handbook of Computational Social Choice*. 2016. Edited by Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. Cambridge University Press.
34. Bilge Mutlu and Jodi Forlizzi. 2008. "Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction." In *Human-Robot Interaction (HRI), 2008 3rd ACM/IEEE International Conference on*, pp. 287-294. IEEE.
35. Wanda J. Orlikowski. 1992. "The Duality of Technology: Rethinking the Concept of Technology in Organizations." *Organization Science* 3.3: 398-427.
36. Michael Quinn Patton. 1990. *Qualitative evaluation and research methods*. SAGE Publications, inc.
37. Frank Pasquale. 2015. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.
38. Elisha A. Pazner, and David Schmeidler. 1978. "Egalitarian equivalent allocations: A new concept of economic equity." *The Quarterly Journal of Economics*: 671-687.
39. Gary Pritchard, John Vines, Pam Briggs, Lisa Thomas, and Patrick Olivier. 2014. "Digitally driven: how location based services impact the work practices of London bus drivers." In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pp. 3617-3626.
40. Ariel D. Procaccia, and Junxing Wang. 2014. "Fair enough: Guaranteeing approximate maximin shares." In *Proceedings of the fifteenth ACM conference on Economics and computation*, pp. 675-692. ACM.
41. Spencer A. Rathus 1973."A 30-item schedule for assessing assertive behavior." *Behavior therapy 4,* no.3: 398-406.
42. Spliddit. http://www.spliddit.org
43. Anselm Strauss and Juliet Corbin. 1990. Basics of qualitative research. Newbury Park, CA: Sage.
44. Albert Sun. 2014. "To Divide the Rent, Start with a Triangle." The New York Times: pp. D2.
45. Lucy Suchman. 2007. *Human-machine reconfigurations: Plans and situated actions*. Cambridge University Press.
46. Hong Ye. 2015. "Research on Emergency Resource Scheduling in Smart City based on HPSO Algorithm." *International Journal of Smart Home* 9, No. 3: 1-12.
47. Jack Whalen. 1995. "Expert systems versus systems for experts: Computer-aided dispatch as a support system in real-world environments." *Cambridge Series on Human Computer Interaction. 161-183.*
48. Zhilin Zheng, Tim Vogelsang, and Niels Pinkwart. 2014. "The Impact of Small Learning Group Composition on Student Engagement and Success in a MOOC." In *Proceedings of Educational Data Mining* 7.