COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

JOINT DECISION-MAKING AND PUNISHMENT IN THE REPEATED ULTIMATUM GAME

MASTER THESIS

2024
Tiam Ghorab, BA

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

JOINT DECISION-MAKING AND PUNISHMENT IN THE REPEATED ULTIMATUM GAME

MASTER THESIS

Study Programme: Cognitive Science
Field of Study: Computer Science
Department: Department of Applied Informatics
Supervisor: Mgr. Xenia Daniela Poslon, PhD.
Consultants: Prof. Dr. Natalie Sebanz & Luke McEllin, PhD.

Bratislava, 2024
Tiam Ghorab, BA

66707258

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Tiam Ghorab
Study programme: Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study: Computer Science
Type of Thesis: Diploma Thesis
Language of Thesis: English
Secondary language: Slovak

Title: Joint decision-making and punishment in the repeated Ultimatum Game

Annotation: The role of decision aggregation has been studied widely for discrete tasks, however the details driving collective decision-making when navigating through social bargaining situations against uncooperative opponents requires more investigation. While there is existing knowledge on group deliberation increasing individual decision accuracy, the direct comparison between individual and joint performance against uncooperative parties remains unexplored. This thesis aims to explore potential differences between individual and joint performance of dyads in uncooperative social bargaining situations.

Aim: The thesis aims to investigate the difference between individual and joint performance of dyads in social bargaining situations.

Literature:
1. Dezecache, G. et al. (2022). Democratic forecast: Small groups predict the future better than individuals and crowds. Journal of experimental psychology. Applied, 28(3), 525-537. https://doi.org/10.1037/xap0000424
2. Kugler, T., Kausel, E., Kocher, M. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. Wiley Interdisciplinary Reviews: Cognitive Science, 3(4), 471-482. https://doi.org/10.1002/wcs.1184
3. Mercier H; Claidière N. (2022). Does discussion make crowds any wiser? Cognition, 222: 104912. https://doi.org/10.1016/j.cognition.2021.104912

Supervisor: Mgr. Xenia Daniela Poslon, PhD.
Consultant: Natalie Sebanz, PhD.
Department: FMFI.KAI - Department of Applied Informatics
Head of department: doc. RNDr. Tatiana Jajcayová, PhD.

Assigned: 17.03.2024

Approved: 18.03.2024 prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

....
Student Supervisor

66707258

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Tiam Ghorab
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Joint decision-making and punishment in the repeated Ultimatum Game
Spoločné rozhodovanie a trestanie v opakovanej Ultimatum Game

Anotácia: Úloha agregácie rozhodnutí bola podrobne preskúmaná pri diskrétnych úlohách, avšak podrobnosti, ktoré riadia kolektívne rozhodovanie pri navigácii v situáciách sociálneho vyjednávania proti nespolupracujúcim oponentom, si vyžadujú podrobnejší výskum. Hoci existujú poznatky o skupinovom zvažovaní zvyšujúcom presnosť individuálneho rozhodovania, priame porovnanie individuálneho a spoločného výkonu proti nespolupracujúcim stranám zostáva nepreskúmané. Cieľom tejto práce je preskúmať potenciálne rozdiely medzi individuálnou a spoločnou výkonnosťou dvojíc v situáciách sociálneho vyjednávania s nekooperujúcimi protihráčmi.

Cieľ: Cieľom práce je preskúmať rozdiel medzi individuálnym a spoločným výkonom dyád v situáciách sociálneho vyjednávania.

Literatúra:
1. Dezecache, G. et al. (2022). Democratic forecast: Small groups predict the future better than individuals and crowds. Journal of experimental psychology. Applied, 28(3), 525-537. https://doi.org/10.1037/xap0000424
2. Kugler, T., Kausel, E., Kocher, M. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. Wiley Interdisciplinary Reviews: Cognitive Science, 3(4), 471-482. https://doi.org/10.1002/wcs.1184
3. Mercier H; Claidière N. (2022). Does discussion make crowds any wiser? Cognition, 222: 104912. https://doi.org/10.1016/j.cognition.2021.104912

Vedúci: Mgr. Xenia Daniela Poslon, PhD.
Konzultant: Natalie Sebanz, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.

Dátum zadania: 17.03.2024

Dátum schválenia: 18.03.2024

prof. Ing. Igor Farkaš, Dr.
garant študijného programu

....
študent

....
vedúci práce

Acknowledgments

This thesis was supported by the Social Mind and Body Lab at the Central European University in Vienna. I would like to thank my supervisor Xenia Daniela Poslon for her openness to working together on this thesis and also for her support and encouragement as my supervisor on other projects throughout my studies.

This research would not have been possible without Luke McEllin and Natalie Sebanz, whom I would like to thank for making this project possible in the first place and furthermore for offering me an internship that allowed me to gain invaluable hands-on experience on the whole process of conducting a scientific study. My time with the SOMBY Lab fundamentally contributed to my understanding of doing science. I am grateful for all the opportunities to learn and for my supervisors' constant encouragements and patience, allowing me to follow my, sometimes inefficient, ideas and for commitment to this project. My big thanks go to Luke who often went with me through hundreds of lines of code to find small improvements to the experiment. Furthermore, I would like to thank Candasch Acar and Fanni Takatsy for their help with setting up the experiment, and for putting up with unexpected issues during data collection as well as for their emotional support throughout the whole process of testing. Finally, I want to thank my family and friends for the hours of testing the experiment, last-minute proof-reads and for pushing me through all kinds of struggles during my studies. Working on this project helped defy many uncertainties about being able to do science. Thank you for making this possible.

Abstract

Collective decision-making is a fundamental cornerstone of human cooperation. While there is a large body of research on the implications of joint deliberation for cooperation in economic games, the details of how dynamics change when subjects collectively make decisions in these scenarios are yet to be fully understood. Due to its simple and yet dynamic structure, the Ultimatum Game has been studied in various forms, showing that groups perform better in ultimatum bargaining than individuals. However, this has not yet been shown for situations where subjects' influences on their social environment are inhibited by an inflexible and uncooperative opponent. The present thesis aims to investigate the role of joint decision-making and punishment in ultimatum bargaining by further understanding collective strategies for cooperating against uncooperative opponents.

A within-subject experiment (N= 40) was conducted to investigate differences between individual and joint performances in the repeated Ultimatum Game. Subjects were assigned the role of the responders and played against an algorithm who they were told was a human proposer. Individuals were assigned to three condition blocks which consisted of 20 trials each. In the first block each participant played alone as a pre-baseline. In the second block, participants jointly decided upon offers together while communicating via a chat. The third block served as a post-baseline where participants played alone again.

In contrast to our hypothesis, the differences between overall individual and joint performance were not significant and negative as subjects performed gradually worse. No significant difference could be found between performance in the pre-baseline and joint performance or between joint and individual performance in the post-baseline. Albeit yielding small effect sizes, we have found significant differences between the performance in the pre- and post-baseline. Moreover, we find that the size of offers differed significantly from another across all condition blocks, with mean offers in the joint condition block being the highest. We interpret the latter results as higher acceptance thresholds in the joint condition as participants forced the algorithm into yielding higher offers. These results indicate that participants negotiated higher offers when deciding together and carried over their newly adopted strategies when playing alone again, however, they did not collect more rewards and thus did not perform better when deciding together compared to when deciding individually. Possible explanations and implications for future research are discussed.

Keywords: repeated ultimatum game; joint decision-making; punishment; cooperation

Abstrakt

Kolektívne rozhodovanie je základným kameňom ľudskej spolupráce. Hoci existuje veľké množstvo výskumov o dôsledkoch spoločného rozhodovania na spoluprácu v ekonomických hrách, podrobnosti o tom, ako sa mení dynamika, keď sa subjekty kolektívne rozhodujú v týchto scenároch, ešte nie sú úplne pochopené. Ultimátna hra sa vďaka svojej jednoduchej a zároveň dynamickej štruktúre skúmala v rôznych podobách, pričom sa ukázalo, že skupiny dosahujú pri ultimátnom vyjednávaní lepšie výsledky ako jednotlivci. Zatiaľ sa to však nepreukázalo v situáciách, keď je vplyv subjektov na ich sociálne prostredie brzdený nepružným a nespolupracujúcim súperom. Cieľom tejto práce je preskúmať úlohu spoločného rozhodovania a trestania pri ultimátnom vyjednávaní prostredníctvom hlbšieho pochopenia kolektívnych stratégií spolupráce proti nespolupracujúcim oponentom. Vnútrosubjektový experiment (N= 40) sa uskutočnil s cieľom preskúmať rozdiely medzi individuálnymi a spoločnými výkonmi v opakovanej hre Ultimátum. Subjektom bola pridelená úloha respondentov a hrali proti algoritmu, o ktorom im bolo povedané, že je ľudský navrhovateľ. Jednotlivcom boli pridelené tri bloky podmienok, ktoré pozostávali z 20 pokusov. V prvom bloku hral každý účastník sám ako predbeţnú situáciu. V druhom bloku účastníci spoločne rozhodovali o ponukách, pričom komunikovali prostredníctvom chatu. Tretí blok slúžil ako post-základný blok, v ktorom účastníci opäť hrali sami. Na rozdiel od našej hypotézy rozdiely medzi celkovým individuálnym a spoločným výkonom neboli významné a boli negatívne, keďže subjekty dosahovali postupne horšie výsledky. Nepodarilo sa zistiť žiadny významný rozdiel medzi výkonom v predvýskume a spoločným výkonom ani medzi spoločným a individuálnym výkonom v postvýskume. Aj keď priniesli malé veľkosti efektov, zistili sme významné rozdiely medzi výkonom v predvýkone a po výkone. Okrem toho sme zistili, že veľkosť ponúk sa významne líšila od inej vo všetkých blokoch podmienok, pričom priemerné ponuky v bloku spoločných podmienok boli najvyššie. Posledné výsledky interpretujeme ako vyššie prahové hodnoty akceptácie v spoločnej podmienke, keďže účastníci nútili algoritmus, aby priniesol vyššie ponuky. Tieto výsledky naznačujú, že účastníci pri spoločnom rozhodovaní vyjednali vyššie ponuky a svoje novo prijaté stratégie preniesli aj pri ďalšej hre osamote, avšak nezískali viac odmien, a teda nedosiahli lepšie výsledky pri spoločnom rozhodovaní v porovnaní s rozhodovaním osamote. V závere práce diskutujeme o možných vysvetleniach a dôsledkoch pre budúci výskum.

Kľúčové slová: opakovaná ultimátna hra; spoločné rozhodovanie; trest; spolupráca

Contents

List of Figures

List of Tables

1. Introduction

In the discipline of economics itself, the term behavioral economics is, to no surprise, the subject of a long history of tradition. After the Second World War, the convention of methodologically positivistic economics manifested itself even stronger as the field's main paradigm. Along with influences from behaviorism, cognitive psychology, experimental work as well as the cognitive revolution, light was also shed on factors such as task difficulty, the influence of beliefs and perceived information when analyzing the behavior of economic agents (Thaler, 2016). With it, as is argued, the already existing study of Game Theory gained importance in the field of social cognition as the study of outcomes of interacting economic agents. While being a specialized field in economics, with the rise of behavioral economics, Game Theory has also found its way to become a standardized framework in other disciplines such as psychology, biology, anthropology and computer science. Thus, Game Theory has become not only a useful methodological paradigm but also a connecting link that promotes interdisciplinary scientific exchange. Following the economists' tradition, Game Theory builds upon the paradigmatic tenet of methodological individualism, where the individual and their behavior are the main interest of analysis. Furthermore, it expands this by standardizing the study of interactions between individuals by providing a discrete set of options (Samuelson, 2016).

In game theoretical research, the main question lies in strategic behavior and optimal decision-making of agents involved in the game. Due to the emphasis on traceable structures of social interactions that game theoretical perspectives put on economic games, the paradigm of optimization has long been a traditional tenet of experimental economics. While the paradigm observes the interactions of agents with the assumption of rational strategic thinking exhibited by the players, experimental evidence suggests that the optimizing paradigm may not apply to empirical observations with human participants. This is where the intersection between psychological research and experimental economics tries to converge on models that explain these so-called *irrational* behaviors observed in humans interacting in economic games (Gale et al., 1995).

In social cognition, economic games both have a long tradition and have also become an even more widely used methodological approach for studying social behavior in recent years (Thielmann et al., 2021). Since the increasing emergence of the field of behavioral economics, the importance of the paradigm has become more relevant for psychologists and economists. Next to studying the social behavior of money, behavioral economists and economic psychologists argue that these scenarios not only directly represent behavior in regard to the division of resources, but also – depending on the scenario – indirectly represent the division of social power, hierarchies and dynamics between players. In more recent applications, economic games have also been used for studying the influence of affordances and dynamics that affect social choices or preferences (Balliet & Mulder, 2011).

These developments have led to the emergence of research on representative social dynamics arising from economic games. Aside from studying economic decision-making, studies tackle questions of how game dynamics and situational affordances shape beliefs about opponents, or vice-versa, the influence of beliefs and emotions on strategic decision-making. Due to their popularity, the literature on economic games is expanding vastly.

One of the topics that gained relevance in the research on social cognition are phenomena of collective decision-making and group effects in game theoretical scenarios. Group decision-making as well as in-group and out-group dynamics have been among the core interests of social psychologists ever since the formation of the field (Hertwig, 2012). Furthermore, knowledge of group effects opened the gate for further investigations into the factors facilitating these phenomena. Interestingly, research on interacting minds in economic games has helped to understand game theoretical predictions themselves better. While the paradigm of rational agents is a central tenet of game theory, empirical evidence of behavior and economic games defies these predictions. However, it has been found that the behavior of groups in economic games may come closer to game theoretical optima than individuals do (Robert & Carnevale, 1997; Bornstein & Yaniv, 1998).

These findings make the link between research on interacting agents and behavioral economics especially interesting, as they not only hold implications for human social psychology, but for social cognition in the broad sense, also touching fields such as cognitive biology and the information sciences. Economic games allow posing questions on the foundations of human cooperation by studying the mechanics of distributive problems (Hilbig et al., 2018). Mainly consisting of options to cooperate and defect, it is claimed that economic games create frameworks to understand underlying mechanisms of cooperation, such as pro- and antisociality in altruism, punishment and other norm-regulatory behaviors. Among games representing real-life social dilemmas that occur often, such as the Prisoner’s Dilemma or the Public Goods Game, the Ultimatum Game is one of the widely studied scenarios in behavioral economics as well as collective decision-making (Fehr & Gächter, 2000; Rand & Nowak, 2011). When positioning subjects in a situation of unequal action opportunities, the adaptation to asymmetric power dynamics can be observed. Additionally, preferences on the trade-off of fairness and rationality of subjects can be measured. However, more importantly, the scenario of ultimatum bargaining allows investigations into the causes for and the implications of these preferences and behaviors (Fehr & Schmidt, 1999).

The relevance of bargaining situations in everyday interactions goes beyond negotiations on the distribution of monetary resources. Representative positions of power, hierarchies, the establishment and maintenance of social norms as well as representative chances of survival are suggested to be implicitly negotiated in ultimatum bargaining (Arkes et al., 2017). As in other games, the impact of collective decision-making has also been studied in the Ultimatum Game. However, the mechanisms behind the emergence of cooperation and interaction between interacting decision-makers are yet to be adequately investigated. What drives agreement in joint decisions and what kind of dynamics facilitate the success of group over individual performance? What are the situational affordances of ultimatum bargaining that facilitate the processes discovered in previous research?

The present thesis aims to better understand interactions between responders in the Ultimatum Game by examining the dynamics behind joint decision-making in asymmetric distribution problems, and the role the task environment plays for on individual as well as joint performance. A literature review of present research on the Ultimatum Game as well as the role of collective decision-making and joint action will be provided, followed by an experiment that is conducted to test the difference between individual and joint performance of cooperating responders in an inflexible task environment against an unfair opponent. The following chapter reviews, previous literature and discusses possible connections between related theories explaining joint decision-making in the Ultimatum Game. Finally, based on previous findings and gaps, a research question and hypotheses are developed that represent empirical investigations made in this thesis. Lastly, results are presented and interpreted before the limitations and implications of the experiment for future research are discussed.

2. Theoretical Framework and Previous Research

2.1. Shared Representations in Joint decision-making

Mental representations serve as a facilitatory basis for the coordination of joint actions towards a shared goal. When performing planned coordination tasks, agents either plan their own actions in relation to the shared goal or their partner's actions. Tasks that cannot be performed alone often also require representations of partners next to representations of the task goal itself. These shared representations of tasks allow for predictions and the coordination of actions before and during task performance. Sebanz et al., (2006) postulate that "[...] the ability to form shared representations of tasks is a cornerstone of social cognition. It allows individuals to extend the temporal horizon of their action planning, acting in anticipation of others' actions rather than simply responding" (p. 73)

Evidence suggests that agents represent each other's tasks even if their tasks are independent from each other. However, depending on the task goals, properties, and affordances, factors such as the spatial proximity of agents may be necessary for shared representations (Knoblich et al., 2011). Nevertheless, participants' performance may even be influenced by solely knowing about the other participant's task without the opportunity to observe their partner's actions (Sebanz et al., 2005). Dyads (in the sense of experimental settings; two individuals linked by a task) may even share task representations if those negatively affect individual performances. In the majority of experimental cases, however, shared representations enable agents to predict and integrate their co-actor's actions for their own action-plans through monitoring their own and their partners' mistakes. Joint adaptation to the task environment, the shared goal, or the mental states of a partner further has been shown to increase joint action performances (Vesper et al., 2017).

Another mechanism that is crucial for successful joint actions is the emergence of shared attention between co-actors, that includes focusing one's attention on the object or event which is being attended to by one's co-actor. This framework suggests this shared attention and interaction of individual attentional focal points to create a new joint focus of attention, which is relevant for the dyad and their task representations as well as their ability to perform the task itself. Joint attention serves as a ground for initiating coordinating actions towards a shared goal but also for coordinating actions while cooperation towards a shared goal is already in process. Joint attention was found to be the base for joint task performance among dyads next to allowing the representation and predictions of the partner's actions (Sebanz et al., 2006). Michael et al. (2016) propose that agents may as well solely coordinate due to an external cue that may serve as the basis for coordination but not as a basis to infer the partner's willingness to adapt to the situation or their strategy. Agents may coordinate to reach a particular goal or coordinate as an end in itself when coordination may be required to facilitate a situation, e.g., when coordinating time and date for meetings or making group decisions.

One common explanation for an agent's understanding of decision-making is that similar to other interactions with the environment, individuals form models of themselves and their current situation when making decisions and performing subsequent actions to make sense of the processes they encounter. For this, the internal model has to relate the intended or predicted outcomes with the perceivable outcomes of the action. An individual's sense of agency builds on the strength of this relation and, therefore, emerges if predicted and actual outcomes match (Pesquita et al., 2018). In the sense of joint actions, these processes become relevant for coordination with a co-action partner. For said joint actions, next to one's own actions and outcomes, the intentions and actions of one's partner must be predicted and evaluated as well. Moreover, there are accounts proposing that from these joint efforts, an additional level emerges which requires the integration of the joint goal and the partner's action plan. Therefore, the experience of joint agency and thus the representation of the shared goal and co-actors actions is crucial for the coordinated success of joint actions (De Vicariis et al., 2022). This experience of joint task representations can also be crucial to more abstract tasks such as joint decision-making. Ginkel & Knippenberg (2008) have shown that the performance of groups making decisions together can rely on their members' being aware about sharing the same task representations. Specifically, sharing information through exchange of opinions, knowledge and perspectives and members having the experience of having the same integrated task representations as the rest of the group, enhanced group decision performance compared to groups without these experiences. While tasks in this study were of uncompetitive nature, the role of shared representations potentially evoking emotional arousal is of particular interest for the present work.

Tenenbaum and Land (2009) assume cognitive processes as a row of internal decision-making and prediction processes of how to interpret stimuli, cognitively appraise them, and, to integrate them into the memory system, and which motor action plans to use. One may interpret this approach as similar to predictive processing accounts (Friston et al., 2009; Diaconescu et al., 2017; Wiese, 2017; Isomura, 2022). Furthermore, they frame mental representations as underlying mechanisms for these decision-making processes, ultimately claiming that cognitive processes are based on mental representations and schemata. This also goes for the integration of and the adaptation to environmental change. In contrast, in cases of emotional pressure and arousal, the authors claim, some neural pathways and therefore schema-connections do not operate on the same level as they would under normal conditions as more strongly wired heuristics come into play which suppress processes for slow, deliberate processing. This can be observed in decreased cognitive-motor control and higher reaction times on the cost of accuracy during stress; it may also hold for intentional decision-making, making it harder to navigate through complex environments under pressure (Farhadbeigi et al., 2012; Yu, 2016). Studies using electroencephalographic (EEG) imaging investigated representations during joint actions and have shown that receiving cues about an upcoming joint action allows subjects to reduce unnecessary representations of irrelevant action plans for the benefit of focusing on representing goals and coordination plans for the remainder of the possible items in the joint action set. This leaves room to interpret that such mental representations do not only consider one’s own and the other’s actions but a different concept of a so-called we-representation, comprising a set of possible joint action-plans that contain sensorimotor predictions for both partners’ movements alike (Kourtis et al., 2019).

McClung et al. (2013) investigated the role of group membership for shared representations that may contribute to we-representative approaches. In a go-nogo task participants were instructed to press a button if a stimulus arrow points into a specific direction. In a joint version of the task, they were paired with a partner and received the same instructions, only this time, they knew that their partner was responsible for pressing their button when the arrow pointed in the other direction, which they were not responsible for. Measuring reaction times, the authors found that those were significantly slower in the joint task than in the individual task. These results suggest that the mere knowledge about their partners’ task increased their processing time for pressing the button in the correct stimulus and therefore, this may indicate co-representation of their partner’s task next to their own. However, participants were either paired with an in-group or an out-group partner, and the described effect solely occurred for subjects who were performing together with an in-group partner. Thus, with subjects performing their task independently from their out-group partners, these results suggest that out-group members were not represented as relevant agents for the task independently of the task properties but solely because their partners were part of an out-group.

Halevy et al. (2012) examined mental representations of conflict and negotiation using economic games. Mental representations of these situations and, hence, people's views of these conflicts, shape the behavior and cooperativeness of players involved in the game. Albeit examined for symmetric games, the authors found that a small set of games emerged out of the two players having two choices of action (i.e., cooperation or coordination by player 1 or player 2). These structures afford reasons for competition and cooperation akin to be found in asymmetric games such as the Ultimatum Game as well. One of these social affordances for example, is the one of self-promotion through exploitation and the desire to harm the opponent out of spite to reinforce norm adherence (Jensen, 2010; Columbus et al., 2019; Thielmann et al., 2020). These examinations, show that mental representations and the sense-making of individuals may influence their perceptions and motivations about a conflict. Depending on their perception and affordances of the payoff structure as well as their partner's behaviors, competitors may engage in perspective-taking to a higher or lower degree. However, what are the implications of the human ability to share representations of tasks and conflicts as a group and to what extent do groups benefit from joint decision-making and sharing task representations? The next section aims to combine knowledge from joint action with decision-making frameworks in order to understand the role joint deliberation within groups for making successful collective decisions.

2.2. Interacting Minds making Collective Decisions

As a field concerned with social systems and broadening the methodological lens from one to many individuals interacting with each other, social psychology has been facing the issue of not only studying how people act towards each other, but also how they interact with each other on group-external levels. Next to in-group and out-group effects, the field also investigated group decision-making processes since its early stages (Hertwig, 2012). While crowds in social psychology were often considered to be influenced by several biasing factors leading to poor group judgment, a number of studies report the aggregation of large individual decisions or the so-called *wisdom of crowds* to outperform individual judgments by the same agents, which has ever since been regarded as a supportive argument for democratic decision processes (Galton, 1907; Lorge et al., 1958; Surowiecki, 2005; Page, 2008; Bonabeau, 2009; List, 2012; Yi et al., 2012; Ratner et al., 2023). Subsequent literature suggests that this effect, albeit in a flattening curve and depending on the task, may increase with larger crowd sizes (Walter et al., 2022).

While the theory on the *wisdom of crowds* posits collectively aggregated decisions to outperform those of the group's individuals, one may differentiate between collective decision-making processes and the aggregation of large amounts of individual decisions. The latter does not involve immediate discussion and deliberation among the collaborating individuals (Mercier & Cladière, 2022). Literature on collective decision-making suggests that forming group consensus enhances the already present benefit of aggregating diverse perspectives. Nuanced exchanges of perspectives underlying individual judgements may add a crucial component to converge on the best integrated judgment as an active group process. These deliberation processes have improved group performance over their individuals' judgements in various tasks (Laughlin et al., 2002). This might be due to the dynamical deliberation processes and active filtering of actions and judgements, which may have an evolutionary advantage over the mere post-hoc aggregation of individual performances (Conradt & List, 2008). Bahrami et al., (2010) demonstrate that dyads may make Bayes optimal decisions and thus perform better than each of their members alone at visual oddball discrimination tasks. Their experiments show that dyads performed better if they had the opportunity to communicate freely within the decision time. Before joint decisions were made and discussed, participants declared their individual decisions first. Interestingly, they find that the communication of certainty to be a strong predictor of dyad performance as the more correct team member performs better than the dyad if levels of certainty are not communicated. Moreover, feedback about the correctness of the decision does neither improve individual nor dyad performance. The same data was explored by Fusaroli et al., (2012) for communicative patterns that benefit joint task performance, for which they find a correlation between linguistic alignment general and to the task environment and team performance over trials. Hence, they show that besides serving as a means to coordinate decisions, coordinating the use of language itself benefits collective decision-making. This combination of confidence ratings and linguistic alignment may be one of the nuances in collective deliberation that depict an improvement of post-hoc aggregations of individual judgments. Roy et al. (2021) find weighted averages of neural decision signals and individual confidence ratings to be a more accurate predictor for performance than the majority vote in a visual detection task. Dezecache et al. (2022) investigate the effects of collective deliberation for abstract decision-making with no immediate feedback, related actions, or outcomes on individual performance. Small groups were asked to make predictions about future geopolitical events as individuals and then again as groups after verbal deliberation. This deliberation increases the predictive accuracy of group decisions compared to the individual decisions made by their single members. These and other findings from previous studies demonstrate the benefits of aggregated consensus decisions over aggregated independently made decisions by individuals (Hamada et al., 2020). Moreover, if given the opportunity to correct their individual decisions after group discussions, individuals were found to make more accurate predictions than when they do not have this opportunity. This is in line with Robert & Carnevale (1997) demonstrating an improvement in individual performance in the Ultimatum Game after participants have played in groups.

However, despite the evidence supporting the benefits of collective decision-making, another line of research examining the impact of social influence suggests that the latter may impede the benefits of collective decision-making. Accordingly, social influence (i.e., the influence due to which individuals revise their initial opinion's) may either improve collective decision-making, given that the crowd's error is high, the average opinion underestimates the actual solution value. There is moderate social influence, or harm if given that the collective error is initially low. This suggests that collective opinions are sensitive to social influence and that the latter may correct unstable or destabilize stable collective assumptions (Mavrodiev & Schweitzer, 2021). Lorenz et al. (2011) argue that in order to access the benefits of collective decision-making, independent opinions are required for collective deliberation. However, in most cases, the subjects invited to experiments and surveys do not provide independent opinions as they mostly form their opinions based on information they obtained from the same networks. Therefore, their opinions are neither independent of the shared information network nor are subjects protected against the social influences they exert on each other. The wisdom of crowds is therefore sensitive to herding behavior. Thus, the mere exposure to others' opinions may be used to update beliefs towards those opinions, which may lead to convergence of perspectives even without the latter being more accurate than any of the individual estimates. Due to this adjustment to other opinions, polarizing estimates may prevail at the end of deliberation processes while individuals' estimates become more and more alike, and diversity of perspectives, which is the basis for collective decision-making, decreases at the same time. This herding behavior can be described as a perpetual process across multiple iterations of finding group consensus. Since different opinions complement each other, they are also influenced by their counterparts, which makes the group deliberation itself a process of belief updating which consequently forms the next cycle of group discussion. Especially for tasks where no correct value is available and individual confidence is low, group diversity becomes fragile in such situations (Navajas et al., 2022). Thus, the process of collective deliberation may actually negatively influence individuals' beliefs if the sample of opinions is not independent. In these scenarios, individual confidence may be beneficial or detrimental to the collective. High confidence serves as a predictor of accurate judgment prior to deliberation. However, this may develop into confirmation bias if social influence is present, as encountering similar opinions may quickly increase confidence and thus cause individuals to quickly arrive at consensus (Moussaid et al., 2013). In this sense, confidence may also boost error potential in scenarios where individual judgments are vulnerable to fallacies. Here, two heads were found to be less accurate than one and the accuracy of members with lower confidence was more reliable than that of more confident members who had the potential to distort group judgment (Koriat, 2012). These findings may be relevant for the Ultimatum Game, where group effects are present, which make deciders vulnerable to inequality aversion or spiteful punishment. Schkade et al. (2000) presented evidence for more amplified pre-existing judgments on punishment ratings after group deliberation.

Additionally, sociological processes can exert pressure on strongly identified group members and make them adopt group judgment, creating in-group norms. In light of adopting group opinions, subsequent emotional contagion may serve as an evolutionary mechanism of social learning and group adaptation to external stimuli (Stolle et al., 2024). Evolutionary theories on altruistic behavior posit that docile group members may engage in imitation as a social learning mechanism and, ultimately, in high expressions, a ground for herding. For example, neural activation in the insula, an early evolved limbic structure relevant to emotional processing, has been associated with altruistic behaviors in the Ultimatum Game. The benefits for adapting to immediate conflict may be a long evolved local optimum; however, for complex and stochastic environments such as economic decisions, these instinctive mechanisms may impede more beneficial Bayesian reasoning that requires careful deliberation. Neuroscientific evidence has associated risk-related financial decision-making with increased activations in structures of the limbic system, especially for situations with ambiguous information. Additionally, social and financial reward processing have been associated with the ventral striatum, to greater extent in cooperative interactions (Baddeley, 2010). Thus, thinking in quick, norm-adherent and

Bayesian deliberative decision-making may be a more suitable scheme to frame joint decision-making in the Ultimatum Game than the simplified view of separating behavior into rational and instinctive expressions as is done by traditional economics (Kahneman, 2011).

It is this conflict of frameworks that finds its way into our provided analysis of the joint Ultimatum Game in the following chapter. The conflict ranges between evidence that joint decision-making may increase joint task performance, which was also shown for the Ultimatum Game, and frameworks supporting the view that social influence on the ground of emotional responses to inequality may impede joint performance in complex situations without unequivocal solutions.

2.3. The Ultimatum Game

The Ultimatum Game is one of the classical games analyzed in game theory and resembles a distribution problem in the form of a bargaining situation.

While the situation of ultimatum bargaining only contains some properties of bargaining situations in general, which can be tainted by a myriad of complex interactions, ultimatum bargaining is nevertheless considered a useful tool to examine certain processes and behaviors that may resemble other bargaining situations in the field. The situational aspects of the Ultimatum Game are considered to have perfect information, as both agents have complete information about all possible as well as previous outcomes and decisions. This follows a paradigm of economics and behavioral economics to operationalize the action-space of the interactions to be measurable at the cost of the ecological validity of the situation. However, a large body of literature on economic games used in experimental psychology and cognitive science pleads for the usefulness of these situations to investigate cooperation and (pro-)sociality. Moreover, these limited action-spaces mainly focus on the question of what is supposed to be investigated using these paradigms (Güth et al., 1982).

The problem of the classical Ultimatum Game tackles the interaction of two players distributing an externally received amount of money between each other while being in an asymmetric dynamic. Player 1, the proposing party, can decide how much money of a given amount they want to share with their counterpart. Hence, the proposer must determine how much of the money they keep and how much they allocate to the responding party. The responder receives information about the original amount of money the proposer was supposed to allocate and receives an offer from the proposer. The responder then can decide to accept or reject the offer. If they accept, both the proposer and the responder receive the shares that the proposer distributed. However, if the respondent rejects the offer, not only does he or she receive nothing, but the proposer also loses their possible gain (Güth & Tietz, 1990).

While this makes the structure of the Ultimatum Game simple to provide a ground for studying game theoretical predictions, it also represents a space to study psychological processes that emerge through the affordances given by the game and the interpersonal social interactions of the players.

From a game theoretical view, the Ultimatum Game has two subgame-perfect equilibria based on the concept of *Nash equilibria* which can be identified by analyzing each player's best possible reaction to each possible state (i.e., each offer). Through a game theoretic lens, each player has a set of possible strategies to adopt with a Nash equilibrium representing such a set of strategies. The Nash equilibrium comprises a set where the strategies of all the players represent their best possible responses to the choices of their competitors. Hence, there is no incentive for an adaptation of strategies (Holt & Roth, 2004). In the Ultimatum Game a Nash equilibrium would represent an offer and a response where both the proposer and the respondent would not change their decisions if presented with the outcome Thielmann et al. (2021). propose the importance of the minimum acceptance level the responder decides on. One subgame-perfect equilibrium would then be the offer of 0 and the minimum acceptance level of also 0, implying that the proposer would best offer 0 due to the irrelevance of the responder's further actions. The second subgame-perfect equilibrium the authors propose is the offer of 10% with the equivalent minimum acceptance level, thus implying that the optimal choice for the responder is to set the minimum acceptance level equal to the offer as accepting a low offer leads to higher outcomes than rejecting it. While these two subgame-perfect equilibria are considered the optimal solutions according to game theory, they do not mirror experimental evidence since neither proposers nor responders offer and accept only 10% of the initial endowment. On the contrary, rejection rates are the highest for offers of 10% and half of the offers making up 20% of the initial pool are rejected as well. In experiments, the most frequent offers are 40% or 50% of the total amount, which are mostly accepted by responders. These values vary across populations but remain robust on average (Houser & McCabe, 2014). These findings can best be described as a recurring pattern that has been coined as inequality aversion which has repeatedly been replicated by a number of studies. This framework explains the occurrence of punishment of uncooperative co-actors at the expense of subjects' resources to maintain cooperation and why subjects reject unfairly perceived offers that should otherwise be accepted in one-shot Ultimatum Games (Fehr & Schmidt, 1999).

Therefore, the game theoretic Nash equilibria seem to be inapplicable for actual bargaining situations in the field. Schuster (2017) suggests an alternative approach for an ecologically more valid Nash equilibrium for the Ultimatum Game, the golden ratio, based on insights from justice research and the optimization of step-wise fractionation of values close to equal offers. Based on empirical evidence and hence on the assumption of effects of inequality aversion in the Ultimatum Game, the golden ratio models an ecologically valid equilibrium of 0.382 as an offer proportion, which yields higher utility for the proposer and is close enough to an equi-distribution to prevent the saliency of inequality in the distribution. Hence, the framework suggests this to be the limit that is still considered an approximation of equality by responders, whereas 0.35 is already mathematically closer to ⅔ and, therefore, makes the presence of an unequal distribution more salient. Given that the rules demand offers to be integers, the golden ratio indeed is supported by empirical evidence, which maps offers of 40%, apart from offers of 50%, as most desirable for responders (Camerer, 2003). Therefore, a more ecologically valid Nash equilibrium for the Ultimatum Game may amount to 40% of the initial endowment, given that the offers can only be integers.

Next to the Nash equilibrium, there are various equilibria in social dilemmas where joint decision-making is required. The Kantian equilibrium, for example, is a commonly used solution concept in symmetrical games such as the Prisoner's Dilemma as it suggests that each agent maximizes their payoff under the assumption that their competitors will act in the same way (Roemer, 2010). However, as the Ultimatum Game is an asymmetrical game and this thesis aims to highlight the differences and bridges between game theory and social coordination, we will mainly focus on the Nash equilibrium and its variations as it highlights the mentioned conflict for the dilemma of the Ultimatum Game.

2.4. The Ultimatum Game in Groups

With the continuing merging of behavioral economics and social psychology, increasing research has been conducted on the role of social cognition in economic situations. Therefore, the standardization of experimental conditions through paradigms of economic games has played a crucial role in understanding group decision-making concerning the distribution of resources. Moreover, these kinds of investigations have led to new research in social cognition, specifically on in-group interactions in collective decision-making.

Combining frameworks describing the differences between collective and individual decision-making and economic games, a body of literature examines the effects described in the framework of the wisdom of the crowd effect within the context of economic games. Parallely, in the course of the investigation of game theoretical scenarios by social psychology, the question of how groups perform and interact in economic games also became relevant for both social psychology and experimental economics.

Previous research in this area has suggested different hypotheses, leading to opposing results in some cases. However, one might argue that the issue of the replication crisis, as well as the different affordances among various economic games, may make it a challenge to draw proper conclusions for group performance in economic games in general (Columbus et al., 2019). Separate syntheses of replicable scenarios for each economic game would have to be compiled in order to draw exhaustive conclusions on this problem. While the challenges for such an endeavor are multifaceted, Kugler et al. (2012) have reviewed a large body of literature on interactive decision-making in groups throughout various economic games of sequential but also simultaneous nature. They conclude that most of the reviewed studies found that group behavior in economic games was more rational and selfish than the behavior of individuals. In line with the wisdom of the crowd effect, they draw connections to groups performing better in understanding game rules and structures through discussion and collective error correction. Furthermore, the authors underline assumptions predicting that fearing the decisions of opponents may cause groups to play more selfishly due to distrusting beliefs. However, the review only comprises studies of groups of equal size. Moreover, they suggest that through aggregated preferences, the social desirability and acceptance of greed are expressed higher and provide opportunities for more selfish behavior. Robert & Carnevale (1997) compared individuals and groups taking the role of the proposer as one of the first experiments and found groups to exhibit a more rational playstyle (i.e, made selfish rather than generous offers) than individuals as well. Moreover, offers to out-group members came closer to game theoretic predictions than to members of an in-group. The researchers also found the most competitive group member to have been the best predictor for the offer that groups ultimately made. Individuals who have played the game again after interacting in a group, adjusted their playstyles to their most competitive group member. Bornstein & Yaniv (1998) compared groups and individuals in a between-subject design for both the roles of responders and proposers and replicated these effects. However, while groups played in a more rational and selfish playstyle, they did not show to be more competitive than individuals. Groups offered less and also accepted lower offers compared to individuals which is in line with game theoretic predictions. Competition, on the other hand, would be associated with higher rejection rates for low offers, which was not observed by the researchers.

Albeit the limitations of these investigations, the aspect of groups being more rational but not more competitive represents a crucial implication for groups playing the Ultimatum Game, which is that their playstyle might be closer to game theoretic predictions than to inequity-averse (i.e., competitive) behavior shown by individuals. This indicates that the aggregation of collective preferences through discussion may in fact be beneficial to decision outcomes in bargaining situations. The benefit of aggregating collective preferences through a discussion process over simply collecting votes in a rule-based decision-making system was shown by Elbittar et al. (2011), who matched groups of proposers and responders in an Ultimatum Game where the offers were automatically accepted or rejected if responders did not come to a unanimous decision or accepted if the majority of the group voted to accept the offer. In contrast to Robert and Carnevale (1997) little difference was found in the voting behavior of individuals between one-on-one and group settings. On the individual level responders, unlike proposers, exhibited similar voting behavior across all conditions. However, direct group interaction was not possible, and individual votes were cast anonymously. Therefore, the similarity of votes between group members was measured rather than the aggregation of collective preferences. Moreover, the lack of discussion distinguishes this procedure from the previously described scenarios. Similarly, but for the role of the proposer, Messick et al. (1997) conducted experiments with groups of proposers in the Ultimatum Game and investigated whether expectations of voting differences among responders influenced the proposers' offers. Specifically, whether they had to meet at least one or all the responders minimally accepted offers. However, they did not find significant differences between these conditions either.

Additionally, the advantages of options for dynamic conversations not only come to light for communication among team members but also for communication between opposing parties. In the classic Ultimatum Game, the decisions of the responders (i.e., accepting or rejecting the offer) are the only source of information for proposers to form and update their beliefs about the responders next decisions. Rankin (2003) explored these dynamics for cases where responders have the opportunity to make requests and found requests to generally result in lower offers and higher rejection rates. He suggests that proposers may perceive requests as unfair since they mostly were for more than half of the monetary pool. This suggests that spiteful actions may also emerge on the proposer side. Several studies investigated the role of communication for distributive problems and economic games in relation to agent decision-making. Especially in feedback-based scenarios (i.e., when there is feedback from and to players), sanctions and punishments play an essential role as regulatory mechanisms. Andrighetto et al. (2016) investigated the synergy of punishment and communication in multiplayer public goods games. They show that in light of communicating norms, either explicitly through verbal communication or implicitly via in-game sanctions, punishment may regulate cooperation through norm signalling and even reduce rates of spiteful counter-punishment, when the initial punishment was norm adherent. Choi & Menghrajani (2011) find pre-bargaining discussions between proposers and responders to influence bargaining decisions towards more cooperative offers and consequently reduced rates of rejections. Moreover, pre-bargaining discussions seems to promote the emergence of shared cognition as well as shared identity and thus also higher performing collaborative bargaining.

These findings and alternative variations of the ultimatum game including multiple subjects collectively making decisions for one role as a team and pre- and in-game communication between team members and opponents suggest that collective decision-making may alter the dynamics of ultimatum bargaining processes. The further chapters elaborate on theories of collective decision-making and aim to locate its role for bargaining processes.

2.5. Social Norms & Spiteful punishment – More important than Money

According to game theoretical frameworks, responders in the Ultimatum Game should accept any offer higher than 0 since gaining something is better than paying the opportunity costs of rejecting offers (Thielmann et al., 2021). However, as outlined above, empirical evidence shows that people do not react this way and tend to reject lower offers more frequently even when they only have one chance of playing and thus only one chance of gaining money. The framework of inequality aversion describes these preferences, claiming that the aversion of asymmetric conditions is less favorable for people than the gain of small amounts of money (Fehr & Schmidt, 1999). While inequality aversion explains a wide range of behavioral patterns in ultimatum bargaining, the roots of the phenomenon itself are not fully understood yet. What is it that drives people to avoid finding themselves being treated unfairly and preferring symbolic status over real monetary incentives? Social psychological explanations base these phenomena on establishing and maintaining of social norms. Motives to exhibit so-called irrational behavior may not be locally optimal in one-shot bargaining but these preferences may be representative of more global power dynamics between populations beyond one-shot negotiations.

The social heuristic hypothesis presents defection in unfair situations as a useful way to enforce fairness by the counterpart. By making norms salient and punishing the transgression of implicit social norms, decision-makers lay the ground for more fair situations in future interactions. In the Ultimatum Game this translates to the rejection of unfair offers. In this framework, social heuristics serve as either evolved or culturally acquired implicit strategies that help agents model and hence navigate through social interactions efficiently. What may not be expressed optimal decision-making, may represent internalized information processing systems suited for more complex environments (Rand et al., 2014). Several empirical observations support the relevance of norms for decision-making in ultimatum bargaining. For example, evidence suggests that fast decisions are more likely to be acceptances and slow decisions are more likely to be rejections; Ferguson et al. (2014) link this to internalized social norms. They propose the degree of certainty over whether a situation is unfair or not to be predictive for cooperation or defection and show that offers close to a 50-50 split are accepted faster than more unfair offers but that rejection rates for these offers rise with longer delays before a decision is made. These internalized social norms may also be the ground for evidence showing that proposers may converge to equal splits by themselves (Bahry & Wilson, 2006). The results of Lia et al. (2023) are in line with these suggestions as they show that negative emotions towards injustice may drive rejections of unfair offers. Cognitive appraisal of unfair offers evokes emotions of anger but if there is only an uncertain range and no clear anchor to evaluate fairness, appraisal is difficult. Thus, rejection rates for ranges of possible offers are lower. However, if the prospect of receiving an unfair offer is salient enough (i.e., low numbers such as 0-2 are in the range with the range not being very large), then the uncertainty effect is cancelled out.

The rejection of offers can be viewed as a demand for higher offers as well as the underlying intention of enforcing norm-adherent behavior by the proposer. Depending on rejection rates, demands or desired norms by responders can be perceived in different intensities by proposers. What is communicated through emotional expressions in direct interactions, is translated to rejected offer values and rejection rates. Mental representations of individuals' own emotional states may underlie perceptions of counterparts and therefore facilitate altruistic behavior even without direct visual interactions. However, the same goes for social pressure which may as well be a product of mental representations of abstract economic information (Baddeley, 2010). In an extensive meta-analytic review, Arvanitis et al. (2019) demonstrated that responders were more likely to accept lower offers if they could express their anger or agreement with their response to the offer either pre or post decision-making. Similarly, proposers tended to justify low offers if given the chance to communicate after they had sent an offer and at the same time were found to make higher offers if they were required to justify. Other accounts have shown that apologies by proposers made them more trustworthy, while their awareness of other players having the opportunity to gossip about them made proposers increase their offers. Expression and communication of anger by the responder and guilt by the proposer furthermore were associated with the rejection of lower and higher offers by proposers and thus higher ratings in coordination and trustworthiness (van Rijk & De Dreu, 2021).

What these results have in common is the effects of rejections on the behavior of proposers. Combinations of economic games may show different cooperative preferences of subjects for different situations. The dictator game presents an interesting case of the Ultimatum Game where the proposer’s outcomes are independent of the responder’s decisions. Proposers can freely decide how much of the initial endowment they share with responders while the latter have no action possibilities. The average amount proposers give to responders in the dictator game is 20% of the initial pool value. From a purely rational standpoint, proposers should offer the smallest possible amount (Engel, 2011). However, theories concerning reputation offer an explanation that takes a population of proposers and responders into consideration. As proposers may find themselves within the population of responders in other situations, weighing the costs of their reputational value becomes important for repeated interactions or future encounters with individuals who observed their behavior (Frith et al., 2008). In the Ultimatum Game proposers adjust their offers if information about responders’ previous decisions and bargaining style is provided. In observed or repeated games, fairness emerges out of these iterative bargaining cycles. The playstyle and reputation of both parties serves as a basis for the convergence of fair offers and ultimately cooperation (Nowak et al., 2000). Following up on this, the effects of framing in bargaining situations have been addressed to emphasize the presence of norm-based decision-making in ultimatum bargaining. Eriksson et al. (2017) compared rejections of offers with reductions of the proposer's payoffs by the respondent by spending the money they previously earned from the negotiation and associated them with moral motives mentioned by responders. Upon receiving unfair offers, participants had the option to pay the received amount (i.e., take the opportunity costs) in order to reduce the proposer's earnings. Motives to punish the proposer were more frequently associated with harming the proposer via the reduction option and motives to be fair with the rejection option. Therefore, Eriksson and colleagues question the concept of altruistic punishment and punishment in ultimatum bargaining per se. Whether associations of behavior with self-reported moral judgments are sufficient to fully understand the roles of rejections for punishment in the Ultimatum Game may be questionable, this perspective nevertheless proposes an interesting perspective for non-evolutionary frameworks.

However, another explanation might be the function of moral judgments as representations of reputational regulations for such scenarios. One aspect of sentiments of fairness may result from relations between players and the reputations agents have among other interacting agents. Thus, reputation may serve as a psychological and regulatory function for social norms, which is operated by rejections in the ultimatum game (Zhang et al., 2023). While the underlying motives for punishment are discussed extensively among scholars studying prosociality, evidence suggests that punishment may be a product of various rather than only one motivation. Frameworks of behavioral biology and evolutionary psychology uphold the paradigm of punishment as a means to ultimately bring about benefits. In some cases, punishment can be costly for the punisher and still be functional in terms of creating delayed benefits, e.g., sanctioning transgressors in order to ensure future norm-adherence on their side. Costly punishment is considered functionally altruistic when punishers themselves do not receive direct benefits but rather incur self-inflicted costs for maintaining of collective cooperation, as is the case in evidence of costly punishment in Public Goods Games (Nikiforakis, 2008). In these scenarios multiple players each possess an amount of money from which they can add to the public pool each turn. The amount in the public pool will then be increased by some ratio, which makes cooperation by everyone the best possible outcome, but individual defection, also called free riding, is the best individual strategy for gaining the benefits of others' labor. Costly punishment by directly affected but also by third parties have both been observed in the Public Goods Game (Fehr & Gächter, 2000; Rand &

Nowak, 2011). While this kind of costly punishment was exhibited by third-party observers and directly affected players that do not receive any of the benefits, this is contrary to evolutionary frameworks that posit that require immediate or delayed benefits for punishment to be beneficial. Jensen (2010) proposes another category of socially motivated functional spite directed at others. However, for functional spite neither the outcome needs to be beneficial to the punisher, nor does cooperation need to be the desired goal behind the punishment act. While the altruistic benefits for the collective may be by-products, the intention may be to inflict emotional or material harm to the transgressor without ulterior motives. Instead, asserting dominance and gaining leverage may be the driving force. An alternative to functional spite is psychological spite, where the punisher psychologically benefits from the harm of the transgressor by restoring self-esteem after losing or maintaining psychological leverage without material benefits. This aligns with findings on inequality aversion despite opportunity costs (Fehr & Schmidt, 1999). However, there are accounts of ultimatum bargaining where responders rejected unfair offers regardless of the proposer‘s intentions which indicates that there were more complex mechanisms at work than sole norm-enforcements. People seemed to also have other-regarding preferences in which they compare themselves to others regardless of intentions and previous actions. Social desirability and norms may constitute inhibitions for psychological spite. Explanations may include that functional spite is grounded in evolutionary motives to inhibit reproduction of the othered population (i.e., other proposers) and therefore have a smaller pool of competitors. Extreme expressions of this process may be reducing everyone's payoff to achieve psychological equality (Jensen, 2010). Psychological spite may be underlying punishment in cases where the punished individual does not have the opportunity to learn from the sanctions and adjust their behavior. In cases where norm-transgressors did not have information about outcomes and, therefore, could not learn from sanctions, spiteful punishment was observed, nevertheless. Thus, while most situations afford costly punishment to remind transgressors about norms, their emotional harm can be an additional driving force next to norm-enforcement. This is in line with Arvanitis et al. (2019), who show that pre- and post-game communication may influence the behavior of both proposers and respondents, possibly to avoid emotional harm. People were found to not admit or at least were not aware of their spiteful motives and instead state motives of inhibiting further norm transgressions by the punished, similar to the conclusions of Jensen (2010). While subjects spent more resources on inflicting punishment and were more likely to punish if they could communicate the norms they wanted the transgressors to abide by, these patterns of punishment, despite no learning opportunities, are nevertheless interesting and suggest spiteful as well as vengeful motives. These results also suggest evolutionary motives of reducing competitor fitness as well (Crockett et al., 2014). However, even though such evolutionary motives may be omnipresent across and within different populations, individual differences still influence tendencies for different types of punishment and other mechanisms for maintaining cooperation. With the evermore presence of economic games in studying social interactions and prosociality, individual differences also have found their way into studies of social cognition using economic games (van Dijk et al., 2004). Among these individual differences, models of personality have been used to study differential playstyles, preferences and aspects of interdependent social interaction and decision-making. Specifically, the HEXACO model of personality has been found to be associated with the situational affordances some economic games provide (Zhao & Smillie, 2015; Lee & Ashton, 2018). One of the strongest personality traits related to situations of prosocial behavior found among multiple models of personality and other individual differences was the Honesty-Humility dimension of the HEXACO model. High scores in this dimension have been related to fairness towards others and cooperation despite opportunities to exploit. Nevertheless, tendencies towards fairness motives may still incline subjects scoring high in this dimension to strive for equal distributions in the face of exploitation as they would too, according to their ideals (Hilbig et al., 2018). However, Honesty-Humility was not the strongest predictor for the role of responders in the Ultimatum Game., which was most strongly associated with the Agreeableness dimension of the same model (Thielmann et al., 2020). Agreeableness comprises tendencies of forgiveness in situations affording retaliation. Thus, subjects scoring high on agreeableness tend to prefer cooperation even in the face of being exploited.

Following this line of thought, the question arises: What dynamics come into play when reputational advantages neither pose benefits nor prevent harm and thus also serve no regulatory function for responders? Moreover, while purposeful acts of paying a price to reduce the payoffs of transgressors for unfair offers may be associated with punishment and solely reactant rejections which are associated with fairness, how does this relate to cases where spiteful punishment is exerted despite seemingly missing learning effects on the transgressor's side? Do instances where rejections that neither serve functions of regulating behavior nor reputational value influence overall interaction with the opponent? If so, how are these influences related to different personality traits usually linked to allowing exploitation due to high cooperative tendencies in the classic Ultimatum Game?

The next sections comprise the methodological approaches of the present work and aim to investigate these questions in further detail.

3. The present study

Based on the findings of Arkes et al. (2017), next to groups of three or more players, groups of two also come closer to game theoretic optima for both teams of two responders and two proposers. Moreover, the framed situation represented power dynamics of even stronger asymmetry than in the original Ultimatum Game due to the assigned roles of management and labor. However, on the other hand, different findings show that spiteful punishment may arise when facing an opponent unwilling or unable to learn from norm-enforcements (Jensen, 2010; Crockett et al., 2014; Eriksson et al., 2017). Such effects, in combination with out-group perceptions, may enhance spiteful rejections for our study, where subjects face an algorithm simulating an unfair and stubborn proposer. Therefore, our first research question *RQ1* asks whether there is a difference between the joint and individual performance of responders in the Ultimatum Game against an uncooperative proposer that does not converge on cooperative dynamics and is insensitive to punishment of norm-violations. While considering evidence suggesting that spiteful punishment may arise against inflexible opponents, we nevertheless expect to find that subjects perform better (i.e., accumulate more rewards in line with the game-theoretic optimum) as dyads compared to when playing alone due to the large body of literature demonstrating the superiority of group decisions (Robert & Carnevale, 1997; Bornstein & Yaniv, 1998; Arkes et al., 2017). Therefore, our Hypothesis *H1* postulates that mean individually accumulated points in the repeated Ultimatum Game will be significantly smaller than accumulated points when deciding together.

Moreover, based on the findings of Robert & Carnevale (1997) as well as Dezecache et al. (2022) showing that individuals improve after group deliberation in the Ultimatum Game as well as non-interactive prediction tasks, we pose *RQ2* which asks whether this effect carries over to individual playstyles when facing an inflexible opponent as well, specifically whether subjects' individual playstyles and leads to subjects performing better individually after having decided jointly with their partner compared to their performance before making joint decisions in a dyad. *H2*, therefore predicts that individual rewards after having decided jointly with a partner will be significantly higher than individual rewards that were accumulated before having jointly made decisions with a partner.

The HEXACO model of personality has been frequently used to study individual differences in cooperation and interdependent interactions. The Honesty-Humility dimension of the HEXACO model was repeatedly found to be linked to the spectrum of cooperation and defection in social dilemmas, specifically to fairness preferences. That is, Honesty-Humility was found to be strongly linked to fairness-oriented behavior while being negatively linked to exploitation, therefore not allowing exploitative behavior while at the same time staying cooperative (Hilbig et al., 2018). Moreover, the Agreeableness dimension was found to negatively predict retaliation after being exploited due to its association to forgiveness in contrast to retaliatory motives (Hilbig et al., 2016). Based on these findings *RQ3* asks how levels of Honesty-Humility and Agreeableness are expressed in accumulated rewards and negotiation efficacy against an inflexible opponent, posing unfair offers. We predict that Honesty-Humility, consistent with how subjects high in Honesty-Humility would behave, relates to intolerance of exploitative behavior and thus higher rewards due to more aggressive negotiations with the proposer as a response to the unfairness. At the same time, we predict that Agreeableness will exhibit a negative relationship with offers received as lower offers will be accepted due to tendencies towards forgiveness instead of retaliation. Thus, *H3a* predicts a positive correlation between Honesty-Humility scores and mean rewards accumulated when playing alone, while *H3b* predicts a negative correlation between Agreeableness scores and mean offers when individually making decisions.

In addition to individual differences in personality, frameworks of collective decision-making assume differential perspectives to be one underlying factor driving the benefits of joint deliberation (Laughlin et al., 2008; Bahrami et al., 2010; Mercier & Claidière, 2022; Dezecache et al., 2022). *RQ4* poses the question what influence diversity in personality traits and performative diversity, i.e., differing amounts of rewards accumulated between partners from previous experiences with the task, have on jointly accumulated rewards of dyads. Differential perspectives in our study may be operationalized as differences in performance due to either different negotiation or cognitive strategies. *H4a* predicts that the joint benefit of dyads, i.e., dyads' mean increase in accumulated points compared to playing individually, is positively correlated with dyads' performance disparities, i.e., the difference of individually accumulated points between partners before having played together. In the same sense, *H4b* postulates a positive correlation between partner disparities in Honesty-Humility scores and joint performance and *H4c* predicts a positive correlation between partner disparities in Agreeableness scores and joint performance in dyads. We predict disparities of both personality facets to be correlated with joint performance as previous literature suggests that regardless of individual performance, cognitive diversity may enhance joint performance in tasks. That is, not the factors influencing individual performance, but the differences between partners are suggested to be the driving forces behind the benefits of joint deliberation, which inspire our final hypotheses.

4. Methodology

With the experiment at hand, we aimed at investigating the impact of joint decision-making in bargaining situations on bargaining performance. Our objectives were to examine how players perform in the repeated Ultimatum Game when they joint deliberate and make decisions compared to their performance as individual players. In the task, participants played the ultimatum game as the responder firstly alone (individual pre-baseline), then as a team (joint condition block) and then alone again (individual post-baseline) against an algorithm that simulated an extortionate proposer.

After completion of the task, participants were presented a questionnaire measuring two dimensions of the English 60-item version of the HEXACO Personality Inventory Revised, namely *Honesty-Humility* and *Agreeableness.* These comprised 20 questions in total, including one control question (Lee & Ashton, 2018; Lee & Ashton 2024; see Appendix A – Questionnaires). The experiment was conducted throughout April and May 2024 in the laboratories of the Department of Cognitive Science of the Central European University in Vienna under the project reference number 2024-01. Prior to execution, the study was pre-registered on April 6th on the open science pre-registration platform *AsPredicted* under the reference number #169334 (AsPredicted, 2024). Participants completed the tasks on computers of the laboratories of the Central European University in separate rooms. Both computers ran on the Microsoft operating system Windows 10 via the 2023.2.3 standalone version of the open-source software for experimental psychological research, *PsychoPy* (Peirce et al., 2019). Code files can be found under the dedicated repository *Joint-Ultimatum-Game* on GitHub (GitHub, 2024).

A funnelled questionnaire at the end served as deception checks to assess participants' beliefs about the likelihood that the proposer was an algorithm.

4.1. Experimental Design

We used a within-subjects design with pre- and post-baseline measurements. Participants were assigned to two different conditions (individual and joint) within 3 blocks (the individual pre-baseline, the joint block and the individual post-baseline) in which they played for 10 points for 20 rounds in each block. The points were converted to actual money after the experiment and were added to the regular reimbursement of € 10 as a bonus, with a possible bonus of € 1-5 per participant per experiment depending on the points accumulated throughout the experiment. This was communicated to participants in advance upon registration as well as upon being instructed. In each block, participants played against an algorithm simulating a proposer playing in a selfish and uncooperative fashion (i.e., to collect as many points as possible and give stubborn offers). However, subjects were told that they would play against a third participant who was introduced to them during the instructions but was played by a confederate.

The deception of participants with confederates can be a challenging task. Hence, this experimental manipulation is often a topic of discussion among experimental psychologists and especially among experimental economics. Although there exists a debate among experimental economists about the definition of deception as to whether there is a distinction between withholding information and actively misleading participants with false information, the discussion on the prohibition of deception in experimental economics is strongly rooted in the discipline (Ortmann & Hertwig, 2002). Arguments against deception include the ethical underpinnings as well as the argument for experimental control as potential suspicion might compromise experimental conditions and validity (Charness et al., 2022). While we mainly focussed on the behavioral patterns and the underlying dynamics of participants' interactions instead of their decision-making per se, we expect the impact of our deception to be tolerable with regard to our data quality. From an experimental psychological perspective, the use of confederates is common for studies concerning social cognition and dyadic behavior, especially when the manipulation of the independent variable results in the simulation and the measurement of responses to unusual behavior (Martin, 1970; Kuhlen & Brennan, 2013; Rauchbauer et al., 2023).

Moreover, we purposefully mainly used the confederate as a decoy for the algorithm and to maintain the asymmetric dynamics of the ultimatum game as the aspect of inequality is crucial to the economic game. While using a frequently invited confederate and asking them to play in an extortionate or egoistic manner might seem like a reasonable alternative, the issue of learning effects would compromise the study. Furthermore, by priming confederates to exhibit a certain bargaining style, we cannot control for standardized experimental conditions across dyads as well as across condition blocks. Therefore, we used an algorithm that was designed to make unfair offers (i.e., with higher weights on less than half of the pool) and follow certain probabilities for increasing or decreasing the offer dependent on the responders' responses to provide standardized experimental conditions. The algorithmic pattern consists of reducing or increasing the offer by 1 point with certain probabilities given the previous offer. Starting with an offer of 30% out of 10 points every phase. This way, we standardize starting offers for every phase to make them comparable between dyads as well as between conditions and between individuals. On the other hand, we expect 30% to be a reasonable starting offer for a greedy and selfish proposer as the maximum offer of 50% represents an equidistribution and 30% is the integer just under ecologically more valid nash equilibrium, or the golden ratio, proposed by Schuster (2017). Moreover, as we try to simulate a selfishly playing proposer, we expect the lower starting offer to mimic an additional stubborn attempt to force a lower ground for bargaining for cases where the preferred offers converged on higher than 30% in the previous phase. We do not expect offers to converge on lower values than 30% as a large body of literature shows that offers of 20% are mostly rejected. For cases where the previous phase has converged on an offer of 30%, the algorithm fits to the amount the parties implicitly agreed upon (Güth et al., 1982; Güth & Tietz, 1990; Gale et al., 1995; Fehr & Schmidt, 1999; Camerer, 2003; Balliet et al., 2011; Avrahami et al., 2013; Hourser & McCabe, 2014; Eriksson et al., 2017; Schuster, 2017; Thielmann et al., 2021).

The algorithm was designed on conditional probabilities as the literature shows skewed distributions across various offer values. Therefore, for the previous offer of 30%, the algorithm increases the offer by 1 point with a probability of 0.7 upon rejection and decreases by 1 point with a probability of 0.15 upon acceptance. We implemented offer reductions upon acceptance to simulate aggressive bargaining behavior on the proposer's side. Albeit not rational and equilibrium-oriented, we expect the option for reduction upon acceptance to represent an irrational proposer who stubbornly refuses to aim for an equi-distributive equilibrium with their bargaining opponents. We put special emphasis on a non-random distribution of conditional probabilities since we expect random offer changes to raise participants' suspicion about the authenticity of the proposer.

If the previous offer was 40%, the probability for increasing the offer was set to 0.45 upon rejection and to 0.3 for reduction upon acceptance. The probabilities for raising the offer were intentionally set lower here to match the selfish policy of the proposer, which would give offers of 50% but repeatedly attempts to enforce convergence on 40% and make responders cooperate. We have set a limit to the algorithm to stop increasing offers beyond the 50% mark as our simulation of a selfish proposer not only includes a maximization of monetary utility who would give in if responders were aggressive enough in rejecting offers to get 60% of the pool. Rather, our simulation of a selfish proposer attempts to combine the maximization of monetary utility in a seemingly irrational way that includes an inverse fairness affinity, namely, to insist on priority due to its higher ground in the bargaining dynamics. Therefore, our simulation of the proposer does not exceed offers beyond 50%, reaching equi-distribution as its maximum offer and trying to further push the offers down. Once 50% is reached, the probability for decreasing the next offer upon acceptance of 50%-offers lies at 0.6. While this at first seems like a rather cooperative policy for a selfish proposer, the goal here was to simulate believable, semi-cooperative behavior as always decreasing the next offer upon acceptance of 50% might be too predictable and not human-like. Therefore, the belief in short-term convergence on an agreement may facilitate the overall robustness of the deception.

For cases where the algorithm offers 20% due to decreasing from 30% upon acceptance, the probability of further decreasing the amount in the coming round was set to 0.1 as offers of 10% have a lower record in empirical observations. While 10% may signal spiteful punishment from the proposer’s side due to simulated frustration if respondents do not cooperate, we expect participants to suspect irregular behavior among the proposers' offers if they only amount to 10% given the initial belief about all three persons to be actual participants invited for the study. Therefore, we assume that spiteful offers, especially in a setting where two people play against one person, would not be expected by participants as they will see each other again for the debriefing after the study and socially unacceptable offers would harm common social norms and therefore cause discomfort among all participants in such a setting. These assumptions result from pre-pilot runs and informal interviews on subjects’ experiences about the procedure and the game.

Lastly, due to similar reasons, the probabilities of increasing the offer after 10% was offered in the previous round, were set to 0.9 to avoid multiple offers of 10% in a row and if at all simulate spiteful offers through single 10%-offers. Similarly, the probability of increasing from 20% to 30% again was set to 0.8 to simulate empirical observations which indicate that offers of 20% are frequently rejected. With this pattern we attempted to simulate persistence as well as a selfish playstyle from the proposer’s side that at the same time still simulates human behavior inspired by empirical evidence (Zhong et al., 2002; Milinski, 2022). However, more importantly, this allowed us to simulate behavioral change in the proposers' offers so that it becomes unpredictable and yet does not seem random to participants. Table 1 below lists conditional probabilities for changing offers proposed by the algorithm.

Table 1
Overview of conditional probabilities for changes in offers proposed by the algorithm

	Probabilities of X changing by 1 point		
Previous Offer (X)	**P(X+1)**	**P(X+0)**	**P(X-1)**
1	0.9	0.1	0
2	0.8	0.1	0.1
3	0.7	0.15	0.15
4	0.45	0.25	0.3
5	0	0.4	0.6

To further promote believability and human behavior, we tackled the issue of interaction time. Evidence shows that subjects may become aware of repetitively identical stimulus time lags. In order to prevent such effects and to simulate contemplation, we added randomized time lags in the range of 5 to 9 seconds before the offer was sent by the algorithm. During these time lags, loading screens with three recurring dots were displayed with messages such as "*waiting for other players…*" for either simulated response lags by the algorithm but also for actual response lags by the participants themselves (Erez et al., 2015; Leavitt et al., 2019). Computers as opponents are commonly used in studies using economic games both in settings where it was made transparent and untransparent that participants were playing against a human or computer (Crockett et al., 2014; Peterburs et al., 2017; Milinski, 2022). Additionally, the cost-efficiency of experiments also needs to be taken into consideration. Three-participant studies are a costly and risky endeavor as they are prone to errors by participants as well as to testing cancellations due to a high potential for no-shows. However, more importantly, this way we ensure standardized experimental conditions and prevent possible learning effects that may emerge if we alternatively repeatedly invite a confederate who would be told to play in a selfish manner. While there have been early approaches to recreate ultimatum bargaining via machine learning, specifically with Q-learning models, there has been significant progress ever since (Erev & Roth, 1998; Camerer, 2003). However, to our knowledge, only a small proportion of studies on the ultimatum game research the repeated ultimatum game where the relationship between proposer and responder is focused on. Thus, as far as we are aware of, literature on the repeated ultimatum game is even more scarce (Zhong et al., 2002). This represents a lack of sufficient training data for reinforcement learning approaches. Nevertheless, we are confident that experiments like this may provide a starting ground to experiment on machine learning approaches simulating human behavior and competing against real subjects.

After completing the personality questionnaire, which was conducted after the game was completed, participants were presented with another set of 12 questions about their experience of the experiment. In a funnel ranging from open questions about the purpose of the experiment, over their strategies, their descriptions of their partner's and the proposer's behavior, participants were finally directed to subjective ratings on 10-point scales about how selfish they perceived the proposer and the partner as well as how fairly they felt treated by the proposer and how fairly they felt their partner was playing. After asking two additional open questions whether they noticed anything unusual about their partner or proposer, the last question asked participants how likely they believed it is that the proposer was a computer, again on a subjective 10-point scale (see Appendix A – Questionnaires). Funnelled questions for assessing suspicion of deception are a common practice in experimental psychological studies as direct questions may reveal the deception and the question itself might prime subjects into suspecting deception, ultimately undermining the purpose of the deception check (Chartrand & Bargh, 1999). Deception checks about the game as well as personality questionnaires were purposefully placed at the very end to avoid priming subjects into confirming their behavior to their described personality structures as well as to avoid suspicion about the relevance of personality measures and behavioral responses for the study. That said, we are aware of the issues that come with personality questionnaires after situations of potentially fairness-related conflict. Participants who might feel mistreated by the proposer may perceive themselves as fair in contrast to their aggressor and therefore answer questionnaires in a way that puts them in a socially desirable moral high ground to avoid cognitive dissonance. However, we argue that asking these questions after the experiment may still be the most feasible point due to the potential conflicts mentioned above. Moreover, we are aware that the last question about the likelihood of the proposer being a computer may raise similar issues, namely possibly priming participants into thinking about the validity of the proposer in the first place or effects of social desirability about seeing through the deception and not wanting to admit that they did not think about that possibility. However, in order to quantitatively measure possible suspicions of deception, we argue this question funnel to be the most effective alternative (Bargh & Chartrand, 2014).

4.2. Participants

Participants were recruited via the Research Participation System of the Central European Universities Department of Cognitive Science where campus students as well as external interested parties can join the participant pool and voluntarily register for upcoming studies. Inclusion criteria for participation included being at the age of 18-65, proficiency in English and basic computer literacy. Upon recruitment, participants were told that there would be a variable bonus of minimum € 1 and maximum € 5 per participant as an additional bonus to their regular reimbursement of € 10, depending on their performance throughout the experiment. Participants were told that both them as well as the proposer may receive a bonus to their regular reimbursement depending on their performance, while in fact, only participants will receive any reimbursements.

All participants were presented an information sheet explaining the purpose of the study without disclosing the focus on team performance. Instead, we presented the study purpose as an attempt to better understand the mechanisms by which people interact with each other in social and economic contexts (see Appendix C - Participant Information Sheet) Participants were told beforehand that they will perform the experiment together with another participant and play economic games together and communicate through a chatroom. The study was approved by the Psychological Research Ethics Board of the Central European University in Vienna. All participants were informed of the usage and protection of their data. Furthermore, written informed consent according to the ethical requirements of the research ethics board was provided by all participants prior to participation. A total of 40 participants (i.e., 20 dyads) took part in the study. 10 participants were male and 30 were female. Participants were between 20 and 34 years old, with an average age of 25,83 years (SD = 3,57). The sample size was estimated with an a priori power analysis (using G-Power 3.1 for a two-tailed one sample t-test with $\alpha = 0.05$; $1-\beta = 0.8$ and an effect size dz = 0.5) which suggested that we need 34 participants, however we rounded this up to 40 as to ensure that we have sufficient power to detect our effect in case of exclusions. None of the dyad members knew their partner privately, except for one dyad which had performed a non-verbal task together in another experiment a day before measurement for our study. Despite our comparison of dyad performance across pairs, we treat each individual participant and their phase performances as individual data points as we also compare across conditions in individuals. Therefore, the 34 participants suggested by G-Power does not entail 34 dyads but, 34 individual participants in total, which we complemented by additional 6 participants, inviting 40 individuals in total.

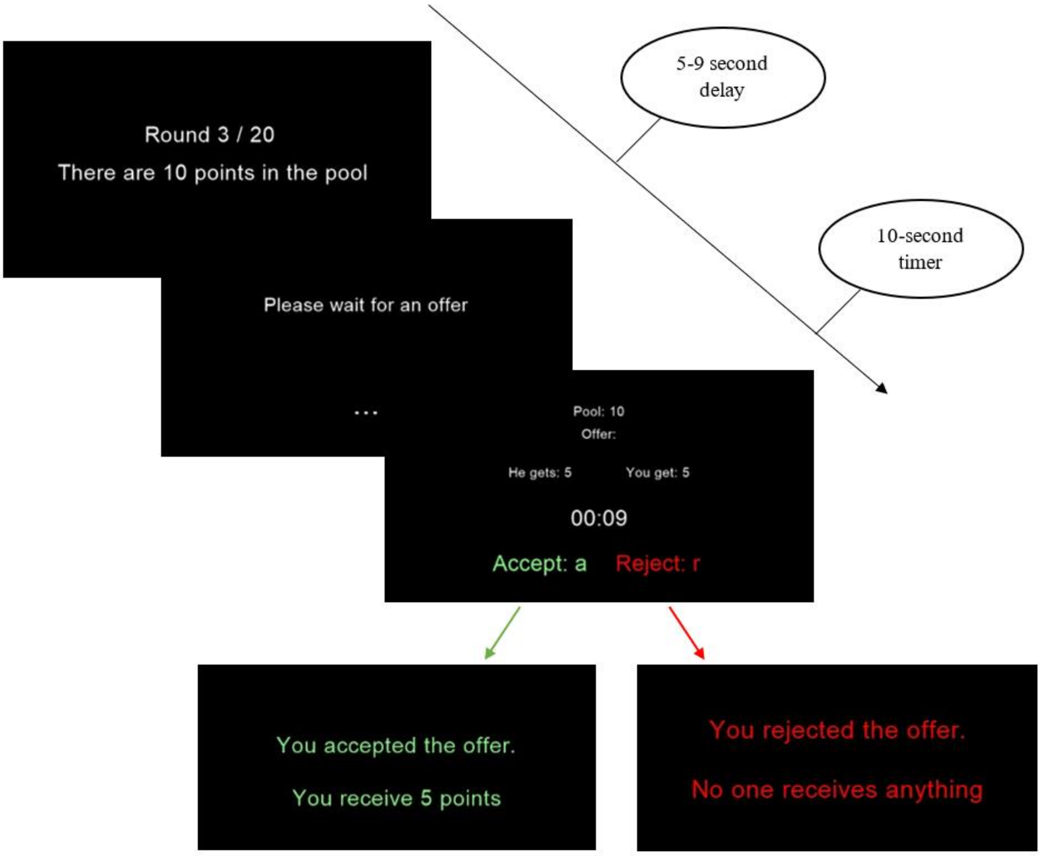
4.3. Procedure

The data for this study was collected in the laboratory facilities of the Department of Cognitive Science at the Central European University in Vienna. Upon arrival, participants filled out the necessary information and gave written informed consent for their participation in the experiment. Participants were provided with an information sheet with details of the experimental procedure and instructions for the process followed by a verbal description of the tasks by the experimenter. In line with the ethical requirements, participants were informed about the option to pose questions about the task or cancel the experiment at any time. The lab managers of the Social Mind and Body Lab managed the arrival, instructions, and the conduction of the experiment. As described in the instructions upon registration and in the instructions before the experiment, participants were invited in pairs of three, however, one of them was a member of the Social Mind and Body Lab acting as the confederate. Participants were assigned the role of the responder and were told that two of them would be playing against the third participant, who was assigned the role of the proposer. We decided to tell participants that the roles would be assigned by order of arrival and decided against simulating a biased randomizer as an explicit demonstration of apparent randomization might have raised suspicion among participating parties. First, the confederate was shown to their room where an open PsychoPy window was already displaying the instructions for the role of the proposer. Afterwards, participants were shown and seated into their respective rooms by the experimenter, where they also faced an already running PsychoPy script with the instructions for the role of the responder. Once every party was seated, the experimenter asked them to follow the instructions on the screen which once again described the task procedure and the confederate started to run the algorithm simulating the proposing party.

In the first block, the individual pre-baseline, participants played for 20 rounds individually. Each round participants received an offer from a (replenishing) pool of 10 points by the proposer on the display and the options to accept (in green color) or reject (in red color) as well as a 10 second countdown that indicated the time they had left to make their decision. Once a participant locked-in their choice by pressing the "a" key for accepting or "r" key for rejecting, they were notified about the outcome of the round and how many of the points they received and how many the proposer received. If a participant failed to lock-in their decision within the time limit a random choice was made for them and they were notified about the outcome of the respective round. After each round, the amount in the pool and the number of rounds left was indicated on the screen before the offer was presented. Prior to the first block, participants played one test round which did not count to the accumulated points. At the end of each block participants were informed about their accumulated rewards for the phase as well as about the next steps for the second phase. Participants played simultaneously while seated in separate rooms. Illustrations of the task process trajectories are depicted below in Figure 1.

Figure 1

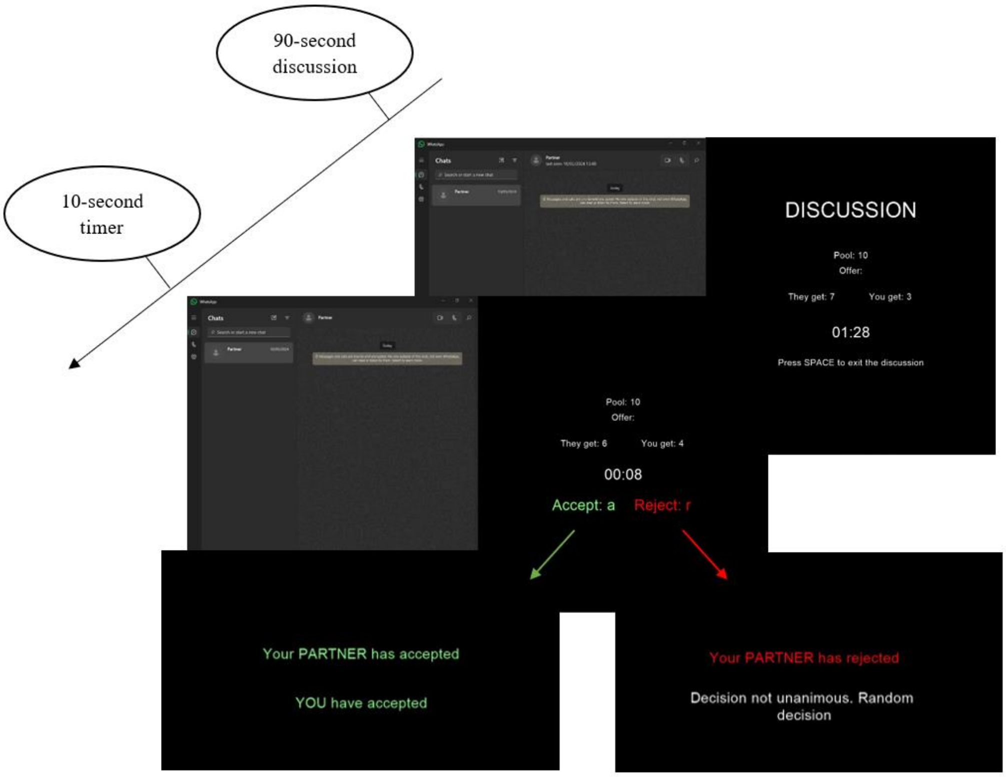
Process of the individual condition blocks for the pre- and post-baseline



In the second block, the joint condition, participants played as a team against the proposer and decided upon accepting or rejecting the offers together for another 20 rounds. Thus, participants were required to come to a unanimous decision. However, participants both received the accumulated points for this phase. For each round, after the offer was made, participants had 90 seconds (which was indicated by a countdown again) to discuss via a virtual chat room. We used the desktop version of the instant messaging service WhatsApp as the virtual chat room, as it poses a simple and stable software for a chat room. Furthermore, we expected most participants to be familiar with the services of WhatsApp. Moreover, as WhatsApp is a commonly used and familiar tool, we assume it to offer an ecologically more valid option in an otherwise rather artificial setting and allow for a clear visual discrimination between the messages. Two dummy accounts were created via newly created phone numbers prior to the experiment. Each dummy account only had access to one contact, which was named *Partner* on the respective accounts. No participants were unfamiliar with the format of instant messaging services and did not report any issues with using WhatsApp's chat environment. All participants were properly instructed in using the desktop app before the experiment. The chats were exported and cleared after each run, so the following dyad would start with an empty chat.

Participants had the option to manually end the discussion before their time ran out if they came to a unanimous decision beforehand by pressing the "space" key and if so, were directed to the accept/reject screen again with a 10 second timer to lock-in their decisions. In case participants did not end the discussion phase manually, the 10-second time limit for making the decisions was displayed after the discussion-countdown hit 0. If their decisions matched, the game continued to the next round. In case their decisions did not match or one of the participants ran out of time, a random decision was made for them by the computer and information as to why the random decision was made was provided. The second block of the experiment started with a test round as well which was not considered for the accumulated points for the phase. We ensured that participants would not communicate after the discussion screen had ended by implementing a function that would take global keyboard inputs regardless of the program of the activity on the computer. This way, inputs for accepting or rejecting would be taken by the program if the participants used the letters "a" or "r" while chatting outside of the decision time window. We disabled this function for the discussion screen. Figure 2 depicts the task process of the joint condition block.

Figure 2
Process of the joint condition block



After the second block, the first phase was repeated as a post-baseline. This phase was identical to the first one, but there was no test round for the third block. At the end of the experiment, participants are informed about how many points they had accumulated in total. Afterwards, participants were presented with a new window containing instructions for navigating through a questionnaire with questions regarding characteristics about themselves. The English 60-item version of the HEXACO Personality Inventory Revised, namely for *Honesty-Humility* and *Agreeableness* was used, thus comprising 20 questions (10 per dimension) and an additional control question, asking participants to press the key “4” after the first 10 questions. Participants were informed about the opportunity to go back and revise their responses until they had not submitted their final answers. After completing the personality questionnaire, participants were presented with another set of 12 questions about their experience of the experiment.

After finishing the experiment, participants were welcomed by the confederate upon arrival out of their rooms and were first asked whether they had suspected the confederate to be part of the laboratory staff. Subsequently, participants were fully debriefed about the purposes of the study as well as the role of the confederate and the algorithm after the study.

4.4. Measures and Analyses

To assess responses and offers across phases and trials, the data was recorded on a trial-by-trial basis throughout the experiment. Task performance was measured for each trial with the exception of each of the first runs in the pre-baseline (block1) and the joint condition (block2) respectively, which were labelled and treated as test runs for participants to become familiar with the virtual task environment and did not count to the overall accumulated rewards. This was communicated to subjects during the instructions. Therefore, a total of 60 valid trials were recorded per participant. To assess task performance, we measured rewards by points accumulated from turns in which participants accepted the presented offers. Our hypothesis claims that joint performance is significantly greater than overall individual performance. As simply comparing accumulated points between phases may not rule out learning effects from understanding the game dynamics and the algorithmic pattern throughout the experiment, we calculate an overall reward difference score ΔA between individual condition blocks and the joint condition block to make them comparable. After calculating the sum A of accumulated points a per trial i for each phase we subtracted this sum of accumulated points in the joint condition j_2 from the mean of individually accumulated points across the pre-baseline j_1 and post-baseline j_3 to assess an overall difference between individually and jointly accumulated rewards. Equation 1 below depicts the formalization of this procedure.

Equation 1

Overall reward difference score ΔA

$$\Delta A = \sum_{i=1}^{n=60} a_i, j_2 \quad - (\sum_{i=1}^{n=60} a_i, j_3 \quad + \sum_{i=1}^{n=60} a_i, j_1 \quad)/2$$

Difference scores of higher amounts signify greater deviations of collective from overall individual performance. If $\Delta A < 0$ then overall individual performance was higher than performance in the joint condition. Conversely, $\Delta A > 0$ indicates more accumulated points in the joint condition compared to overall points in individual condition blocks.

To test our first hypothesis *H1*, predicting that joint and individual performances differ significantly, with the mean rewards in the joint condition being significantly greater than overall individually accumulated rewards, a one-sample t-test was calculated for the computed reward differences scores across participants. 0 was taken as the comparison value since a mean reward difference of 0 would indicate no difference between the mean individual performances and joint performances. Moreover, a repeated measure one-way analysis of variance (repeated measures ANOVA) was calculated for the accumulated points across phases to examine whether there is an overall difference between performance across the phases. Additionally, follow-up mean comparison tests were calculated for combinations of accumulated rewards between phases, to test *H2* predicting that accumulated rewards in the post-baseline will be higher than in the pre-baseline. Furthermore, correlation coefficients were calculated for testing *H3a*, predicting a positive relationship between Honesty-Humility scores and mean rewards accumulated across the individual condition blocks as well as for and *H3b*, predicting a negative relationship between mean offers and Agreeableness scores across the individual condition blocks. Following these analyses, between-dyad level analyses were conducted. Performance differences within dyads were analyzed in order to test *H4a.* This measurement allows us to check whether performance disparities between partners are associated with their joint benefit, i.e., whether the differences in performance between participants in the pre-baseline are associated with their improvement or deterioration from the pre-baseline to the joint condition. For this, we calculated the difference of mean accumulated points in the pre-baseline A_1 between partners. Next, a Pearson correlation coefficient was calculated to measure the strength of the correlation between partner disparities in performance and the difference of individual performance between the pre-baseline and the joint condition block (i.e., the joint benefit).

Complementary to the above-mentioned analyses, correlations of performance and offers with measured personality traits were computed. Scores for individual questions were reversed and facet scores were calculated according to the scoring keys and instructions for researchers (Lee & Ashton 2024). None of the participants failed to correctly answer the control question. Next, correlation coefficients were calculated for computing whether the measured personality facet scores for *Honesty-Humility* or *Agreeableness* influence performance. In order to test *H4b* and *H4c* by examining whether cognitive differences in the *Honesty-Humility* and *Agreeableness* facet scores are associated with joint performance, correlation coefficients were computed for the absolute difference between the scores of each personality facet of partners within dyads and dyads' accumulated points in the joint condition block. Moreover, after these procedures, pairwise mean difference tests were computed for offers as well to examine how, next to the accumulation of points, negotiated offers significantly differed between condition blocks.

As will be reported below, we found significant differences in offers across condition blocks and based on these findings conducted further exploratory analyses to investigate potential factors driving these effects. In these exploratory analyse, a correlation coefficient was computed to examine the relationship between rejections and offers across condition blocks for participants as well as the relationship between rejections and trials within condition blocks. These analyses were conducted as they may serve as evidence for more detailed explanations of the dynamics across the whole experiment as well as within specific phases of the game. To assess deception suspicion, we additionally calculated the mean, median and mode of participants' beliefs on the likelihood that the proposer was a computer. Due to argumentations mentioned above about the possible skewing effects of asking a direct question about the realness of a co-participant, we postulate that, for this pilot study, a normal distribution of data for this measurement with a mean of 5 will be sufficient to assume believability of the proposer not being an algorithm. However, due to the funnelling structure of the deception check also asking questions about the proposer's and partner's behaviors, this does not only comprise suspicions about the the algorithm itself but the whole experimental setup including participants' prior experiences with confederates, economic games and deception in psychological experiments. By computing correlation coefficients for this variable with accumulated rewards and negotiated offers, we check whether people's suspicions might have influenced participants' playstyles and thus, our results.

5. Results

In the following section the within-subject and between-dyad comparisons are described. Mean comparison tests for repeated measurements checked for an overall difference in accumulated points between condition blocks. If significant, pairwise mean comparison tests followed to check significant differences in accumulated points between specific condition blocks. Additionally, a one-sample mean comparison test was computed to compare the difference of the mean reward-difference-score across participants with the value 0 to assess whether overall individual performance was significantly different from joint performance. Furthermore, performance disparity between partners in the pre-baseline were tested on correlations with their joint performance as dyads. Differences in personality facet scores within dyads were correlated with their joint performance as well as with performance on the individual level. Moreover, mean difference scores were computed for offers negotiated between condition blocks which served as a basis for further exploratory analyses. Requirements for statistical tests, specifically Shapiro-Wilk Tests to check assumptions of normal distributions and Levene's Test on homogeneity of variances for mean comparisons have priorly been tested and are not reported below but can be found below in the appendix (see Appendix B - Additional Tables and Results). If requirements were not met, this will be mentioned in the following chapters prior to the tests and respective non-parametric alternatives for the respective tests were computed. Reported p-values for all parametric pairwise mean comparison tests were Bonferroni-adjusted for multiple comparisons and p-values for repeated measure ANOVAs were Greenhouse-Geisser corrected for multiple comparisons. Analyses were conducted using Python version 3.11.

5.1. Within-subject comparisons

The one-way ANOVA for repeated measures for the accumulated points across all three condition blocks ($F(1,78) = 4.18$, $p = 0.02$, effect size: $\eta^2_p = 0.29$), revealed that the means significantly differed from another. Following these results, pairwise comparisons of means between specific phases were conducted. Table 2 lists the mean accumulated points and the reward difference score along with standard deviations.

Figure 3 illustrates a boxplot of mean accumulated rewards per condition block.

The pairwise comparison of means for the accumulated points across phases showed that the mean accumulated points for the pre-baseline ($M = 2.26$, $SD = 2.07$) were higher than the mean accumulated points in the joint condition block ($M = 2.02$, $SD = 2.31$), yet the difference was not found to be significant; t(39) = 1.88, p = 0.19. The effect size, as measured by Cohen's d, was d = 0.42, indicating a small effect. However, different results were found for the mean accumulated points between the pre-baseline and the post-baseline (i.e., the second individual block) (M = 1.92, SD = 2.17); t(39) = 2.91, p = 0.02, with an effect size as measured by Cohen's d, of d = 0.62, indicating a medium effect. The difference between mean accumulated points for the joint condition and the post-baseline was found to be insignificant; t(39) = 0.87, p = 1.00. The effect size, as measured by Cohen's d, was d = 0.17, indicating a very small effect. A one-sample t-test was computed, comparing the difference of the mean reward-difference-score across participants (M = -0.06, SD = 0.66) with the value 0. This difference was found to be insignificant; t(39) = -0.61, p = 0.54. The effect size, as measured by Cohen's d, was d = 0.1, indicating a very small effect. Table 3lists calculated statistics described above.

Table 2

Overview of reward means and standard deviations per condition block and overall reward difference

Variable	Mean	SD
Aj_1	2.26	2.07
Aj_2	2.02	2.31
Aj_3	1.92	2.17
ΔA	-0.06	0.66

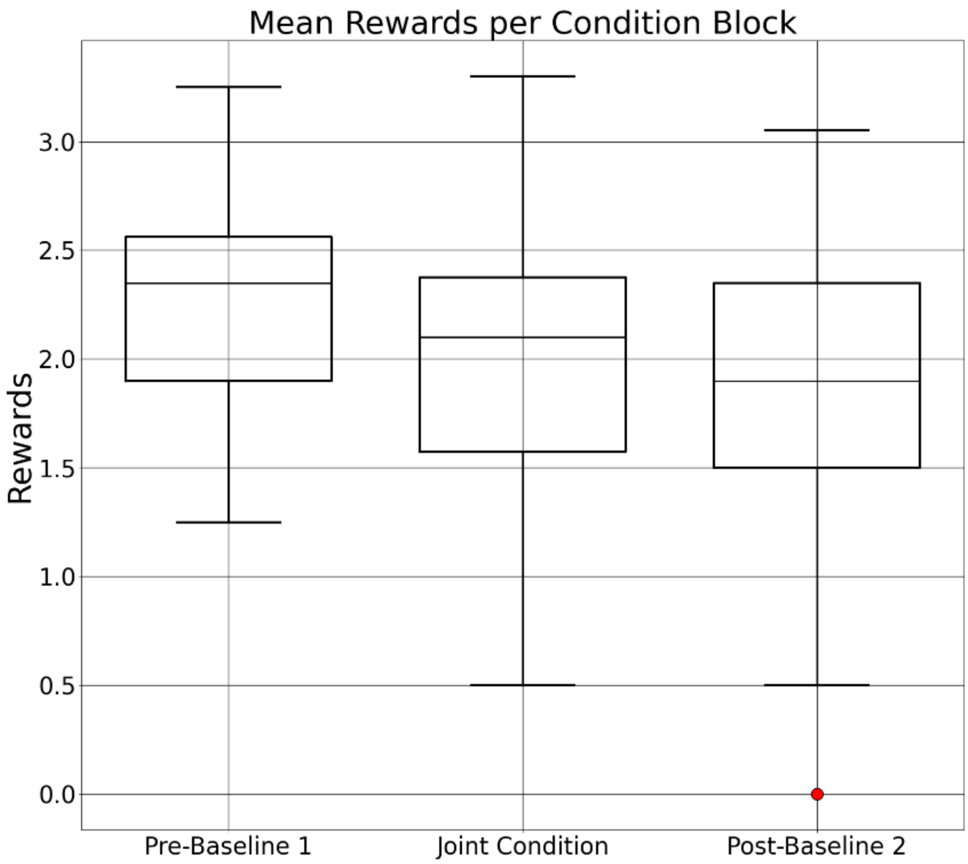
Table 3

Overview of calculated statistics for pairwise comparisons of rewards between condition blocks and overall reward difference

Measurement	Significance level	Test Statistic	Effect size (Cohen's d)
$Aj_1 - Aj_2$	p = 0.19	t = 1.88	d = 0.42
$Aj_1 - Aj_3$	p = 0.02	t = 2.91	d = 0.62
$Aj_2 - Aj_3$	p = 1.00	t = 0.87	d = 0.17
0 - ΔA	p = 0.54	t = -0.61	d = 0.1

Figure 3

Boxplot of mean accumulated rewards by participants per condition block



Like for the rewards, pairwise mean comparison analyses were conducted for offers x between condition blocks by using Wilcoxon signed rank tests due to non-normal distributions in the data. Table 4 shows mean offers and the offer difference score along with standard deviations. Figure 4 illustrates a boxplot of the mean offers per condition block.

The Wilcoxon tests showed that the mean offers for each participant in the first condition block (M = 3.82, SD = 0.77) were significantly lower than the mean offers in the joint condition (M = 4.23, SD = 0.65); z = -4.09, $p < 0.01$. Similar results were found for the mean offers in the pre-baseline and the post-baseline (M = 3.99, SD = 0.76); z = -2.74, $p < 0.01$. The difference between mean offers in the joint condition and the individual post-baseline was found to be significant as well; z = -3.26, $p < 0.01$. A one-sample Wilcoxon signed rank test was computed, comparing the difference of the mean offer-difference-score ΔX across participants (M = 0.32, SD = 0.45) with the value 0. This difference was found to be significant as well; z = -4.02, $p < 0.01$.

Table 5 lists calculated statistics described above.

Table 4

Overview of offer means and standard deviations per condition block and overall offer difference

Variable	Mean	SD
Xj_1	3.82	0.77
Xj_2	4.23	0.65
Xj_3	3.99	0.76
ΔX	0.32	0.45

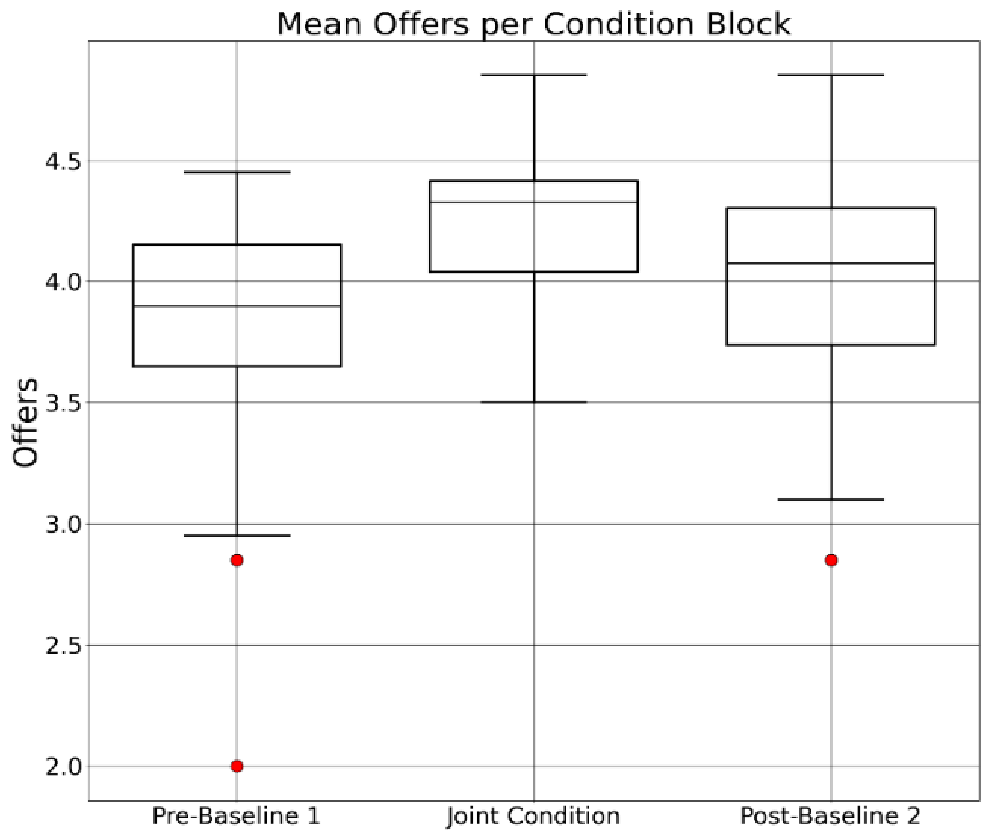
Table 5

Overview of calculated statistics for pairwise comparisons of offers between *condition blocks and overall offer difference*

Variable	Significance level	Test Statistic
$Xj_1 - Xj_2$	$p < 0.01$	$z = -4.09$
$Xj_1 - Xj_3$	$p < 0.01$	$z = -2.74$
$Xj_2 - Xj_3$	$p < 0.01$	$z = -3.26$
$0 - \Delta X$	$p < 0.01$	$z = -4.02$

Figure 4

Boxplot of mean offers by participants per condition block



The influence of Agreeableness on negotiated offers and Honesty-Humility for accumulated rewards was examined by computing Pearson correlation coefficients for facet scores and mean accumulated rewards or offers respectively across both individual baselines. The first correlation coefficient revealed no significant relationship between the *Honesty-Humility* facet and mean rewards across both individual baselines; $r(38) = 0.22$, $p = 0.18$. Significant, yet weak, negative correlation coefficients between the *Agreeableness* facet and mean negotiated offers across both individual baselines were found; $r(38) = -0.32$, $p = 0.04$.

5.2. Between-dyad correlations

For the between-dyad tests, first, a Pearson correlation coefficient was calculated for partner disparities in mean accumulated points in the pre-baseline A_1 (i.e., the performance disparity in the pre-baseline) and mean difference of accumulated points in the pre-baseline A_1 and accumulated points in the joint condition A_2 (i.e., the joint benefit). There was no significant correlation for the two variables; $r(18) = -0.34$, $p = 0.14$. The influence of differences in personality facets within dyads on accumulated points in the joint condition was computed with an additional Pearson correlation coefficient for the Agreeableness facet; $r(18) = 0.17$, $p=0.47$ and a Spearman correlation coefficient for the Honesty-Humility facet; $r(18) = -0.20$; $p = 0.40$, due to a non-normal distribution of the Honesty-Humility disparities between partner. Both correlation coefficients were statistically insignificant.

5.3. Exploratory Analyses

The stepwise decrease in rewards per condition block shows participants gradually accumulated less rewards in each phase of the game, indicating that acceptance rates in the joint condition and post-baseline decreased in comparison to the pre-baseline. Considering that the mean comparison analyses described above have shown that performance in the pre-baseline ($M = 2.26$) was significantly greater than in the post-baseline ($M = 1.92$) but not significantly different from joint performance ($M = 2.02$) and at the same time, offers increased significantly from the pre-baseline ($M = 3.82$) to the joint condition ($M = 4.23$), the question arises what influences this increase in offers and decrease in rewards. Our algorithmic pattern increases offers after previous offers were rejected with certain conditional probabilities. Therefore, we first compared differences in rejection rates between condition blocks with pairwise t-tests and found a significant increase in rejection rates from the pre-baseline ($M = 0.43$); $t(39) = -3.67$, $p < 0.01$ to the joint condition ($M = 0.56$) but not a significant difference between the joint condition and the post-baseline ($M = 0.54$); $t(39) = 0.38$, $p = 1.00$. Consequently, the difference between rejection rates in the pre- and post-baseline were found to be significant as well; $t(39) = -3.65$, $p < 0.01$. The p-values were Bonferroni-corrected for multiple comparisons. While these differences may give insights into the link between rejection rates and offers across phases, they do not provide any information about offer trajectories within condition blocks. Therefore, we explored whether rejection rates decrease over trials played in each condition block and found Pearson correlations to be significant for the pre-baseline; $r(18) = -0.54$, $p = 0.01$, the joint condition; $r(18) = -0.62$, $p < 0.01$ and the post baseline; $r(18) = -0.71$, $p < 0.01$. Moreover, we found rejection rates to be positively correlated with offers across participants for all condition blocks; $r(38) = 0.7$, $p <0.01$. Beliefs about the likelihood of the proposer being a computer were also evaluated and descriptive statistics for the deception check can be found in Table 6. Due to a non-normal distribution, a Spearman correlation coefficient was computed for measuring the relationships between the variable and accumulated rewards as well as negotiated offers. The Spearman correlation coefficient yielded significant results but very weak negative associations for the deception check with both accumulated rewards; $r(38) = -0.04$, $p = 0.02$, and negotiated offers; $r(38) = -0.06$, $p < 0.01$.

Table 6

Overview of descriptive statistics for the deception check about the proposer being a computer

Mean	Median	Mode	SD
5.38	5.50	7.00	2.57

6. Discussion

The present work investigated the effects of joint decision-making for responders in the repeated Ultimatum Game. Specifically, we investigated whether making decisions jointly with a partner results in an overall improvement of accumulated points compared to deciding upon offers alone as individuals and possible factors underlying these dynamics. Our predictions were based on literature on group decision-making in the Ultimatum Game as well as literature on the effects of collective deliberation for joint decision-making. Below we will discuss possible reasons underlying our results, limitations of our study and implications for future research.

6.1. Interpretation of Results

Our results do not replicate previous findings showing that groups in the Ultimatum Game come closer to the game theoretic optimum, exhibiting more rational playstyles in the form of accepting smaller offers and thus accumulating more rewards (Bornstein & Yaniv, 1998; Kugler et al., 2012; Arkes et al., 2017). In contrast to these findings, albeit not to a significant degree, subjects accumulated less points in the joint condition compared to the overall rewards in individual blocks. Hence, our results do not support our first hypothesis which predicted an overall higher performance in the joint condition compared to the individual condition. From the perspective of collective deliberation frameworks postulating that joint deliberation increases individual task performance, our results for differences in accumulated rewards between the pre- and the post-baseline do not replicate these findings either as performance in the post-baseline was found to be significantly lower. Moreover, we could neither find any significant relationship between increase in performance and cognitive diversity among partners in terms of differences in personality facet scores nor between increase in performance and individual performance disparities in the pre-baseline. Thus, our results do not match with previous literature claiming that cognitive diversity in personality and differences in individual task performance may have benefits for joint task performance (Bahrami et al., 2010; Dezecache et al., 2022). As for the relationship between personality scores and offers, we predicted *Honesty-Humility* to be positively associated with individually accumulated rewards as well as to find a negative relationship between *Agreeableness* and individually negotiated offers in the pre-and post-baseline. Although our findings for Honesty-Humility and individually accumulated rewards, our results for the negative relationship between Agreeableness and individually negotiated offers were significant and thus in line with postulations and findings in previous literature (Hilbig et al., 2016; Thielmann et al., 2020).

Out of line with our predicted framework, our analysis of the negotiated offers shows the pattern we were expecting for the accumulated rewards. While performance of accumulating points decreased across condition blocks, negotiated offers increased in the joint condition block after subjects had played individually after the pre-baseline. Albeit, decreasing by a smaller amount after the joint condition block, they still significantly decreased in the post-baseline. Moreover, with the offer difference score being significantly greater than 0, we found that offers negotiated jointly were significantly higher in the joint condition compared to overall negotiated offers when playing alone. These results come closer to patterns shown in previous literature, not examining the Ultimatum Game, which propose that participants exhibited higher task performance when jointly deliberating and carried over these effects to their individual decision-making styles (i.e., exhibited different strategies post-deliberation compared to pre-deliberation) (Dezecache et al., 2022). Due to our results matching this pattern only for offers negotiated and not for points accumulated, we will discuss possible reasons that drove effects found in our analyses.

Different from approaches in previous studies, our stimulus altered the experimental situation by aiming to simulate an uncooperative proposer. To our knowledge there are no accounts yet that match dyads of respondents against uncooperative proposers sending selfish offers. Moreover, the studies we are aware of that set subjects in the role of the responder as teams did either solely compare individually to jointly deliberated responses or, if using a pre-joint-post design, did not do so in the repeated Ultimatum Game (Robert & Carnevale, 1997; Bornstein & Yaniv, 1998; Kugler et al., 2012; Arkes et al., 2017). It is however important to note that including the aspect of repeated interaction and therefore creating possible temporal conflicts between short-term outcomes from accepting small offers and long-term goals to negotiate higher ones and achieve equality in distributions, may crucially change situational affordances and game dynamics (Thielmann et al., 2020). Therefore, one possible explanation for our contrasting results with previous literature could be that perceived dynamics with the alleged proposer influence perceived self-efficacy in regulating the proposer’s offers once participants become aware of the proposer’s stubbornness. This might explain why there was a significant decrease of accumulated points between the pre- and post but not between the pre-baseline and the joint condition block, while for offers, there was a significant increase between the pre-baseline and the joint condition block as well as a significant increase between the pre- and post-baseline. With rejections being positively associated with offers, these results may reflect a gradual increase in rejections across conditions. However, whether these changes in rejections and offers emerged from deliberating together with a partner or simply from the unfairness of the opponent is unclear since no significant difference between the rejections in the joint condition and the post-baseline were found. A stepwise increase in rejections between the phases of the game, which is the case for rejections in the joint condition compared to the pre-baseline, would reflect a gradual increase in spitefulness when repeatedly facing the same unfair offers. However, as our results show, rejections even decrease in the post-baseline compared to the joint condition, albeit to an insignificant amount. Therefore, a gradual development of dysfunctional spite due to an uncooperative and stubborn proposer seems less likely. Rather than progressively increasing spite, one possible explanation could be increasing inequality aversions and functional spite towards only accepting fair offers (Jensen, 2010). Previous accounts propose comparing one's own choices to decisions of others as a predictive factor for behavioral adaptation in socially interdependent scenarios. Avrahami et al. (2013) found players who were given the opportunity to compare their last decision to the median behavior of other players and expressed regret about their decision, to converge to equi-distributions faster, resulting in higher offers from proposers and higher rejection rates from responders. Thus, experiences of others may have an orienting function for changes in strategy. Although performance disparities within a dyad in the pre-baseline were not correlated with the joint benefit of playing together, joint deliberation with a partner that has made similar experiences may have increased higher minimum acceptance thresholds and therefore striving for more fairness-oriented distributions in the next phase of interactions. Next to benefits of joint deliberation of differential perspectives, group-membership effects may have contributed to strategy changes. While previous studies on groups playing the ultimatum game have mostly matched competing groups of equal size, the two versus one setting may have resulted in increased intergroup biases as framing effects may exert influence of underlying motivations and behavioral responses in the Ultimatum Game (Arkes et al., 2017; Eriksson et al., 2017). Previous evidence also shows that negative attitudes against an out-group member may result in ignorance of subjects' own utility maximization if the rejection of unfair offers also affect out-group members (Bella & Sacchi, 2018). Moreover, retaliatory motives emerging from unequal conditions may result in decreasing willingness for cooperation (Chierchia et al., 2021). However, whether the changes in decision strategies found between condition blocks were influenced by an increase in inequality aversion driven by partner deliberation or in-group effects that carried over to the post-baseline requires further analysis.

An explanatory factor that should not be left unmentioned is that participants could have become suspicious of the proposer's role throughout the experiment. This could have led subjects to emotionally disengage from the task. However, while we found our perception check to be significantly correlated with offers and accumulated rewards, the coefficient was close to 0. What may be may be of more importance is the non-normal distribution of the data with a left skewness due to the mode being two integers greater than the mean of 5.38 out of 10, with 10 representing that participants found it very likely that the proposer was a computer. On the other hand, the problem of direct queries for the belief of being deceived in combination with the possible tendency to circumvent cognitive dissonance by not admitting to having been deceived in a game of emotional arousal tainted by unfairness remains. Moreover, there might be a social desirability to have seen through the deception and additionally deception checks have been argued to be suggestive (Hauser et al., 2018). However, we do not argue about the probable influence of participant suspicion but at the same time critically evaluate self-reported deception checks. Nevertheless, the data distribution of the believability of the algorithm's behavior may have influenced our results and will be discussed, along with possible improvements, in the limitations section below.

A different, more interpretable aspect concerning the algorithmic design that may explain increasing rejections could be that, while rejection rates may increase across conditions but decrease over time within condition blocks, participants converged upon specific offers after a number of negotiations and that the last agreements converged upon in the previous block were higher than 30% which was the starting point for each condition block for comparable measurements. Such violations of previous agreements may have increased the retaliatory motives and driven participants to strive for a fair distribution even more. Arvanitis et al. (2019) propose a novel approach to the Ultimatum Game in general that moves away from the decision-making paradigm towards a joint negotiation process of agreeing to rules converged upon. A gradual rule-making process and context-based norms highlight interpersonal coordination as the core mechanism of negotiations such as the Ultimatum Game. Accounts of increased rewards for both proposers and responders due to decisions accompanied by short messages of agreement or frustration support this claim (van Rijk & De Dreu, 2021). This approach framing negotiations as social coordination processes could serve as a framework for further investigating how violating consensually negotiated agreements may influence further negotiation dynamics. However, more detailed analyses of behavioral, specifically temporal, responses would be necessary to make sound conclusions about underlying motives of increasing rejections during joint deliberation carrying over to individual playstyles. Limitations of the study and possible implementations for future research to answer these questions are discussed in more detail in the following section.

6.2. Limitations and Implications for Future Research

As a pilot study for investigating behavioral dynamics between partners playing the repeated Ultimatum Game as responders against an uncooperative proposer, the present work faces a couple of limitations leaving open questions for further investigation. Firstly, while our simulation of an uncooperative proposer was inspired by previous research on proposer offers and responder responses in the Ultimatum Game, the probabilities underlying offers in our model were mostly derived by behavioral reports of one-shot Ultimatum Games. Behavioral data on repeated Ultimatum Games is highly influenced by situational affordances, temporal differences in response times and memory of interactions in prior rounds and available to lesser degrees. Therefore, behavioral data form observed repeated Ultimatum Games was only possible to implement to a certain degree in our rather simple conditional probabilities model. While the use of algorithms as confederates, opponents and partners is a common procedure in studies using economic games, artificially simulated players mostly serve for control conditions in previous studies (Crockett et al., 2014; Peterburs et al., 2017; Milinski, 2022). The use of algorithms as main interaction partners for direct measurements in research on interpersonal coordination and joint action is yet to be standardized while there are already novel approaches for simulating spatial and temporal intricacies relevant for cooperative behavior (e.g., see McEllin et al., 2023). Although research on more complex artificial agents playing economic games has been conducted, most of this literature concerns emergent dynamics evolving from interacting artificial agents for multiple epochs, evolutionary algorithms, or theory-based models (Erev & Roth, 1998; Zhong et al., 2002; Camerer, 2003; Bo & Yang, 2010). Hence, this emphasizes once more the need for large training data sets of repeated interactions in the (repeated) Ultimatum Game, containing not only data on offers and responses but also temporal dimensions such as reaction and deliberation times as well as the need for properly trained and validated reinforcement learning models to simulate more complex human playstyles. Additionally, one limiting factor is the limitation given by deception checks for algorithmic confederates.

As described in the methodology section, checking for suspicion about being deceived is a demanding task for psychological studies, especially for samples that frequently partake in social interaction experiments that use confederates or deception. Thus, next to the training of more complex simulation models, the development of behaviorally measured deception checks for playing against algorithms may benefit future studies to avoid suggestive questions (Hauser et al., 2018).

Secondly, partly due to the same reasons, our study is limited by the fact that all three condition blocks were conducted immediately after another. One of the reasons we decided for a standardized electronic confederate was the benefit of a controlled task environment which rules out learning effects of a human confederate and avoids deviations in offers by the proposer throughout data collection. Hence, with the motivation of standardized conditions across dyad, we also needed to make within-subject comparisons comparable by starting each condition block with the same opening offer. However, since the aspects of repeated interactions, and therefore long-term negotiations, are crucial tenets for our research questions and study design, the conflict between making condition blocks comparable and keeping the stimulus ecologically valid was one of the main challenges to overcome. While even stubborn proposers may try to negotiate offers down again when starting a new phase of the game, this may have implications for reactions by responding parties. Further analysis would be necessary to investigate the possible effects of such dynamics and examine the aspect of converged agreement for negotiation processes as Arvanitis et al. (2019) propose.

Furthermore, larger sample sizes for higher statistical power may be required to find expected, significant correlations between personality traits and accumulated rewards or offers. The Central European University's participant registration system offers a diverse sample due to the university accommodating mostly international students of diverse cultural backgrounds. Nevertheless, pre-screening participants on individual differences and matching them accordingly as dyads would pose an interesting, more statistically valid, opportunity to test whether differences in personality traits influence the benefit of joint deliberation.

For future research on deliberation dynamics within dyads, chat data and temporal dimensions of responses such as discussion and response times might offer interesting and relevant data for investigating the specifics underlying team dynamics in joint decision-making in such situations. Linguistic differences and alignment, dominance in opinion formation, expressed sentiments as well as durations of partner interactions and decision-time synchrony may be influential factors for team performance and partner alignment (Fusaroli et al., 2012; Arkes et al., 2017). Thus, future research of analyzing these textual data with methods such as sentiment analyses may provide valuable insights to identify specific mechanisms behind joint decision-making dynamics and to investigate frameworks akin to those of Arvanitis et al. (2019) who postulate negotiations to be joint rule-making processes that build upon interpersonal coordination that exceed the limits of current decision-making paradigms.

7. Conclusion

The present thesis aimed to investigate the role of joint decision-making and punishment in ultimatum bargaining by understanding the dynamics underlying collective strategies for cooperating against uncooperative opponents in further detail. In this sense, another purpose of the study was to examine whether frameworks of joint deliberation can be applied to situations of amplified asymmetry in the repeated Ultimatum Game.

While our experiment does not show improvements in accumulated rewards during and after joint deliberation, but rather deterioration of performance, we find that negotiated offers increase with joint deliberation and that these effects carry over to individual post-deliberation strategies.

The specifics of mechanisms underlying cooperation have been subject to many years of research in multiple disciplines such as cognitive science, economics, psychology and biology, to only name a few. With these mechanisms already being characterized by complex phenomena, a crucial aspect present in daily human cooperation further increases the challenges of fully understanding the dynamics of negotiations and interactions between and within groups during negotiations. Social interaction and human cooperation are crucially defined by predicting intentions, actions and needs of partners and opponents. However, the question how dynamics of acting alone and within groups change when facing an extortionate and uncooperative opponent is yet to be answered.

With this pilot project we attempt to demonstrate possibilities for novel methodological approaches to understand the dynamics at work in joint decision-making in asymmetric situations in greater detail. By combining frameworks that investigate human interaction from different angles, zooming in on the role of low-level processes for joint decision-coordination may offer new perspectives to understand negotiation processes as communication and coordination rely at the basis for converging to and creating new contextual norms beyond ultimatum bargaining.

References

Andrighetto, G., Brandts, J., Conte, R., Sabater-Mir, J., Solaz, H., Székely, Á., & Villatoro, D. (2016). Counter-Punishment, Communication, and Cooperation among Partners. *Frontiers in Behavioral Neuroscience, 10*, Article 53. http://dx.doi.org/10.3389/fnbeh.2016.00053

Arkes, H. R., Kagel, J. H., & Mezhvinsky, D. (2017). Effects of a Management–Labor Context and Team Play on Ultimatum Game Outcomes. *Southern Economic Journal*, 83(4), 993-1011. https://doi.org/10.1002/soej.12187

Arvanitis, A., Papadatou-Pastou, M., & Hantzi, A. (2019). Agreement in the ultimatum game: An analysis of interpersonal and intergroup context on the basis of the consensualistic approach to negotiation. *New Ideas in Psychology*, *54*, 15-26. https://doi.org/10.1016/j.newideapsych.2018.12.005

AsPredicted 2024, *AsPredicted website*, Wharton Credibility Lab of the University of Pennsylvania, accessed 9 March 2024, https://aspredicted.org

Avrahami, J., Güth ,W., Hertwig, R., Kareev, Y., & Otsubo, H. (2013). Learning (not) to yield: An experimental study of evolving ultimatum game behavior. *The Journal of Socio-Economics*, *47*, 47-54. dx.doi.org/10.1016/j.socec.2013.08.009

Bahrami, B., Olsen, K., Latham, P. E., Roepstrorff, A., Geraint, R., & Frith, C. D. (2010). Optimally Interacting Minds. *Science*, *329*, 1081-1085. https://doi.org/10.1126/science.1185718

Bahry, D. L., & Wilson, R. K. (2006). Confusion or fairness in the field? Rejections in the ultimatum game under the strategy method. *Journal of Economic Behavior and Organization*, *60*(1), 37-54. https://doi.org/10.1016/j.jebo.2004.07.005

Balliet, D., Mulder, L. B., & Van Lange, P. A. M. (2011). Reward, punishment, and cooperation: A meta-analysis. *Psychological Bulletin*, *137*(4), 594–615. https://doi.org/10.1037/a0023489

Bargh, J. A., & Chartrand, T. L. (2014). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 311–344). Cambridge University Press.

Bella, M., & Sacci, S. (2018). Not fair but acceptable…for us! Group membership influences the tradeoff between equality and utility in a Third Party Ultimatum Game. *Journal of Experimental Social Psychology*,*77*, 117-131. https://doi.org/10.1016/j.jesp.2018.04.007

Bo, X., & Yang, J. (2010). Evolutionary ultimatum game on complex networks under incomplete information. *Physica A: Statistical Mechanics and its Application*, *389*(5), 1115-1123. https://doi.org/10.1016/j.physa.2009.11.026

Bonabeau, E. (2009). Decisions 2.0: The power of collective intelligence, *MIT Sloan Management Review*, *50*(2), 45-52

Bornstein, G., Yaniv, I. (1998). Individual and Group Behavior in the Ultimatum Game: Are Groups More “Rational” Players? *Experimental Economics*, *1*, 101-108. http://dx.doi.org/10.1023/A:1009914001822

Brañas-Garza, P., Espín, M. A., Exadaktylos, F., & Herrmann, B. (2014). Fair and unfair punishers coexist in the Ultimatum Game. *Scientific Reports*, *4*, Article 6025. https://doi.org/10.1038/srep06025

Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction.* Princeton: Russell Sage Foundation.

Charness, G., Samek, A., & van de Ven, J. (2022). What is considered deception in experimental economics? *Exp Econ*, *25*, 385-412. https://doi.org/10.1007/s10683-021-09726-7

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910. https://doi.org/10.1037/0022-3514.76.6.893

Chierchia, G., Parianen Lesemann, F. H., Snower, D., & Singer, T. (2021). Cooperation across multiple game theoretical paradigms is increased by fear more than anger in selfish individuals. *Scientific Reports, 11*, Article 9351. https://doi.org/10.1038/s41598-021-88663-0

Choi, D., & Menghrajani, E. (2011). Can group discussion promote cooperative ultimatum bargaining?. *Group Processes & Intergroup Relations*, *14*(3), 381-398. https://doi.org/10.1177/1368430210376404

Columbus, S., Thielmann, I., & Balliet, D. (2019). Situational Affordances for Prosocial Behaviour: On the Interaction between Honesty–Humility and (Perceived) Interdependence. European Journal of Personality, *33*(6), 655-673. https://doi.org/10.1002/per.2224

Conradt, L., & List, C. (2008). Group decisions in humans and animals: a survey. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*(158), 719-742. https://doi.org/10.1098/rstb.2008.0276

Crockett, M. J., Özdemir, Y., & Fehr, E. (2014). The value of vengeance and the demand for deterrence. Journal of experimental psychology. General, *143*(6), 2279–2286. https://doi.org/10.1037/xge0000018

Diaconescu, A. O., Mathys, C., Weber, L. A. E, Kasper, L., Mauer, J., & Stephan, K. E. (2017). Hierarchical prediction errors in midbrain and septum during social learning. *Social Cognitive and Affective Neuroscience, 12*(4), 618-634. https://doi.org/10.1093/scan/nsw171

De Vicariis, C., Chackochan, V.T., Sanguineti, V. (2022). Game theory and partner representation in joint action: toward a computational theory of joint agency. *Phenomenology and the Cognitive Sciences.* https://doi.org/10.1007/s11097-022-09819-5

Dezecache, G., Dockendorff, M., Ferreiro, D.N., Deroy, O., & Bahrami, B. (2022). Democratic forecast: Small groups predict the future better than individuals and crowds. *J Exp Psychol Appl.*, *28*(3), 525-537. https://doi.org/10.1037/xap0000424

Elbittar, A., Gomberg, A., & Sour, L. (2011). Group Decision-Making and Voting in Ultimatum Bargaining: An Experimental Study. *The B.E. Journal of Economic Analysis & Policy*, *11*(1), 53. http://dx.doi.org/10.2202/1935-1682.2631

Engel, C. (2011). Dictator games: a meta study. *Experimental Economics*, *14*, 583-610. https://doi.org/10.1007/s10683-011-9283-7

Erev, I., & Roth, A. E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review*, *88*(4), 848-881

Erez, A., Schilpzand, P., Leavitt, K., Woolum, A. H., & Judge, T. A. (2015). Inherently relational: Interactions between peers' and individuals' personalities impact reward giving and appraisal of individual performance. *Academy of Management Journal*, *58*(6), 1761–1784. https://doi.org/10.5465/amj.2011.0214

Eriksson, K., Strimling, P., Andersson, P. A., & Lindholm, T. (2017). Costly punishment in the ultimatum game evokes moral concern in particular when framed as payoff reduction. *Journal of Experimental Social Psychology*, *69*, 59-64. http://dx.doi.org/10.1016/j.jesp.2016.09.004

Farhadbeigi, P., Bagherian, F., & Khorrami Banaraki, A. (2012). The effect of stress on accuracy and speed of judgment. *Procedia - Social and Behavioral Sciences*, *32*, 49-52. https://doi.org/10.1016/j.sbspro.2012.01.009

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *The American Economic Review*, *90*(4), 980-994. https://www.jstor.org/stable/117319

Fehr, E., & Schmidt, K.M. (1999). A theory of fairness, competition and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817-868. https://doi.org/10.1162/003355399556151

Fehr, E., & Schurtenberger, E. (2018). Normative foundations of human cooperation. *Nature Human Behavior*, *2*, 458-468. https://doi.org/10.1038/s41562-018-0385-5

Friston, K. J., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement Learning or Active Inference? *PLOS One*, *4*(7), Article 6421. https://doi.org/10.1371/journal.pone.0006421

Frith, C., & Singer, T. (2008). The role of social cognition in decision making. *Philosophical Transactions of The Royal Society B*, *363*, 3875-3886. https://doi.org/10.1098/rstb.2008.0156

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science*, *23*(8), 931–939. https://doi.org/10.1177/0956797612436816

Galton, F. (1907). Vox Populi. *Nature*, *75*, 450-451. https://doi.org/10.1038/075450a0

Gale, J., Binmore, K. G., & Samuelson, L. (1995). Learning To Be Imperfect: The Ultimatum Game. *Games and Economic Behavior*, *8*(1): 56-90. https://doi.org/10.1016/S0899-8256(05)80017-X

GitHub 2024, *GitHub Repository: Joint-Ultimatum-Game*, created 27 May 2024, https://github.com/HauchTuell/Joint-Ultimatum-Game.git

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4): 367-388. https://doi.org/10.1016/0167-2681(82)90011-7

Güth, W., & Tietz, R. (1990). Ultimatum Bargaining Behavior. A survey and comparison of experimental results. *Journal of Economic Psychology* *11*(3), 417-449. https://doi.org/10.1016/0167-4870(90)90021-Z

Halevy, N., Chou, E. Y., & Murnighan, J.K. (2012). Mind Games: The Mental Representations of Conflict. *Journal of Personality and Social Psychology*, *102*(1), 132-148. https://doi.org/10.1037/a0025389

Hamada, D., Nakayama, M., & Saiki, J. (2020). Wisdom of crowds and collective decision-making in a survival situation with complex information integration. *Cognitive Research Principles and Implications*, *5*(48). https://doi.org/10.1186/s41235-020-00248-z

Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are Manipulation Checks Necessary? *Frontiers in Psychology*, *9*. https://doi.org/10.3389/fpsyg.2018.00998

Hertwig, R. (2012). Tapping into the Wisdom of the Crowd-with Confidence. *Science*, *336*(6079), 303-304. https://doi.org/10.1126/science.1221403

Holt, C. A., & Roth, A. E. (2004). The Nash equilibrium: A perspective. *Proc Natl Acad Sci USA*, *102*(12), 3999-4002. https://doi.org/10.1073/pnas.0308738101

Houser, D., & McCabe, K. (2014). Experimental Economics and Experimental Game Theory. In P. W. Glimcher, & E. Fehr (Ed.), *Neuroeconomics: Decision Making in the Brain (2nd Edition, pp. 19-34).* Elsevier Academic Press. https://doi:10.1016/b978-0-12-416008-8.00002-4

Hilbig, B. E., Thielmann, I., Klein, S.A., & Henninger, F. (2016). The two faces of cooperation: On the unique role of HEXACO Agreeableness for forgiveness versus retaliation. *Journal of Research in Personality*, *64*, 69-78. https://doi.org/10.1016/j.jrp.2016.08.004

Hilbig, B. E., Kieslich, P.J., Henninger, F., Thielmann, I., & Zettler, I. (2018). Lead Us (Not) into Temptation: Testing the Motivational Mechanisms Linking Honesty-Humility to Cooperation. *European Journal of Personality*, *32*(2), 116-127. https://doi.org/10.1002/per.2149

Isomura, T. (2022). Active inference leads to Bayesian neurophysiology. *Neuroscience Research*, *175*, 38-45. https://doi.org/10.1016/j.neures.2021.12.003

Jensen, K. (2010). Punishment and spite, the dark side of cooperation. *Philosophical Transactions of the Royal Society B.*, *365*, 2635-2650. https://doi.org/10.1098/rstb.2010.0146

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Koriat, A. (2012). When Are Two Heads Better than One and Why? *Science*, *336*(6079), 360-362. https://doi.org/10.1126/science.1216549

Kourtis, D., Woźniak, M., Sebanz, N., & Knoblich, G. (2019). Evidence for we-representations during joint action planning. *Neuropsychologia*, *131*, 73-83. https://doi.org/10.1016/j.neuropsychologia.2019.05.029

Knoblich, G, Butterfill, S., & Sebanz, N. (2011). Psychological Research on Joint Action: Theory and Data. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 59-101). Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-385527-5.00003-6

Kuhlen, A. K., & Brennan, S. E. (2013). Language in dialogue: when confederates might be hazardous to your data. *Psychonomic Bulletin & Review*, *20*, 54-72. https://doi.org/10.3758/s13423-012-0341-8

Laughlin, P.R, Bonner, B.L., & Miner, A.G. (2002). Groups perform better than the best individuals on Letters-to-Numbers problems. *Organizational Behavior and Human Decision Processes*, *88*(2), 605-620. https://doi.org/10.1016/S0749-5978(02)00003-1

Leavitt, K., Qiu, F., & Shapiro, D. L. (2019). Using Electronic Confederates for Experimental Research in Organizational Science. *Organizational Research Methods*, *24*(1), 3-25. https://doi.org/10.1177/1094428119889136

Lee, K., & Ashton, M. C. 2024, *HEXACO-PI-R website*, Kibeom Lee & Michael C. Ashton, accessed 9 March 2024, https://hexaco.org

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, *25*(5), 543-556. https://doi.org/10.1177/1073191116659134

Lia, Y., Li, Y.J., & Zhang, H. (2023). Uncertainty reduces rejections of unfair offers in the ultimatum game. *Current Psychology*, *42*, 17977-17984. https://doi.org/10.1007/s12144-022-03004-2

List, C. (2012). Collective Wisdom: Lessons from the Theory of Judgment Aggregation. In H. Landemore, & J. Elster (Eds.) *Collective Wisdom: Principles and Mechanisms* (pp. 203-

230). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511846427

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, *108*(22), 9020-9025. https://doi.org/10.1073/pnas.1008636108

Lorge, I., Fox, D., Davitz, J., & Brenner, M. (1958). A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. *Psychological bulletin*, *55*(6), 337–372. https://doi.org/10.1037/h0042344

Martin, J.D. (1970). Suspicion and the Experimental Confederate: A Study of Role and Credibility. *Sociometry*, *33*(2): 178-192. https://doi.org/10.2307/2786328

Mavrodiev, P., & Schweitzer, F. (2021). The ambigous role of social influence on the wisdom of crowds: An analytic approach. *Physica A: Statistica Mechanics and its Applications*, *567*, Article 125624. https://doi.org/10.1016/j.physa.2020.125624

McEllin, L., Fiedler, S., & Sebanz, N. (2023). Action planning and execution cues influence economic partner choice. *Cognition*, *241*, 105632. https://doi.org/10.1016/j.cognition.2023.105632

McClung, J. S., Jentzsch, I., & Reicher, S. D. (2013). Group Membership Affects Spontaneous Mental Representation: Failure to Represent the Out-Group in a Joint Action Task. *PLOS One*, *8*(11), Article 79178. https://doi.org/10.1371%2Fjournal.pone.0079178

Mercier, H., Claidière, N. (2022). Does discussion make crowds any wiser? *Cognition*, *222*, Article 104912. https://doi.org/10.1016/j.cognition.2021.104912

Messick, D.M., Moore, D.A., & Bazerman, M.H. (1997). Ultimatum Bargaining with a Group: Underestimating the Importance of the Decision Rule. *Organizational Behavior and Human Decision Processes*, *69*(2), 87-101. https://psycnet.apa.org/doi/10.1006/obhd.1997.2678

Michael, J., Sebanz, N., & Knoblich, G. (2016). Observing joint action: Coordination creates commitment. *Cognition*, *157*, 106-113. https://doi.org/10.1016/j.cognition.2016.08.024

Milinski M. (2022). Extortion - A voracious prosocial strategy. *Current opinion in psychology*, *44*, 196201. https://doi.org/10.1016/j.copsyc.2021.08.033

Moussaid, M., Kämmer, J. E., Analytis, P. P., & Neth, H. (2013). Social Influence and the Collective Dynamics of Opinion Formation. *PLOS One*, *8*(11). https://doi.org/10.1371/journal.pone.0078433

Navajas, J., Armand, O., Moran, R., Bahrami, B., & Deroy, O. (2022), Diversity of opinions promotes herding in uncertain crowds. *Royal Society Open Science*, *9*(6). http://dx.doi.org/10.1098/rsos.191497

Nikiforakis, N. (2008). Punishment and counter-punishment in public good games: Can we really govern ourselves? *Journal of Public Economics*, *92*, 91-112. https://doi.org/10.1016/j.jpubeco.2007.04.008

Nowak, M. A., Page, K. M., & Sigmund, K. (2000). Fairness Versus Reason in the Ultimatum Game. *Science*, *289*(5485), 1773-1775. https://doi.org/10.1126/science.289.5485.1773

Ortmann, A., & Hertwig, R. (2002). The Costs of Deception: Evidence from Psychology. *Experimental Economics*, *5*, 111-131. https://doi.org/10.1023/A:1020365204768

Page, S. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton University Press.

Pesquita, A., Whitwell, R. L. &, Enns, J. T. (2018). Predictive joint-action model: A hierarchical predictive approach to human cooperation. *Psychonomic Bulletin & Review*, *25*, 1751-1769. https://doi.org/10.3758/s13423-017-1393-6

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*, 195-203. https://doi.org/10.3758/s13428-018-01193-y

Peterburs, J., Voegler, R., Liepelt, R., Schulze, A., Wilhelm, S., Ocklenburg, S., & Straube, T. (2017). Processing of fair and unfair offers in the ultimatum game under social observation. *Scientific Reports*, *7*, 44062. https://doi.org/10.1038/srep44062

Ratner, N., Kagan, E., Kumar, P., & Ben-Gal, I. (2023). Unsupervised classification for uncertain varying responses: The wisdom-in-the-crowd (WICRO) algorithm. *Knowledge-Based Systems*, *272*, Article 110551. https://doi.org/10.1016/j.knosys.2023.110551

Rankin, F. W. (2003). Communication in Ultimatum Games. *Economic Letters, 81*(2), 267-271. https://doi.org/10.1016/S0165-1765(03)00191-5

Rauchbauer, B., Jank, G., Dunbar, R.I.M., & Lamm, C. (2023). Only empathy-related traits, not being mimicked or endorphin release, influence social closeness and prosocial behavior. *Scientific Reports*, *13*, Article 4072. https://doi.org/10.1038/s41598-023-30946-9

Rand, D. G., & Nowak, M. A. (2011). The evolution of antisocial punishment in optional public goods games. *Nature Communications*, *2*, Article 434. https://doi.org/10.1038/ncomms1442

Rand, D., Peysakhovich, A., Kraft-Todd, G., Newman, G.E., Wurzbacher, O., Nowak, M.A., & Greene, J. D. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, *5*, Article 3677. https://doi.org/10.1038/ncomms4677

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, *22*(4), 1694-1703. http://dx.doi.org/doi:10.1016/j.neuroimage.2004.04.015

Robert, C., & Carnevale, P. J. (1997). Group Choice in Ultimatum Bargaining. *Organizational Behavior and Human Decision Processes*, *72*(2), 256-279. https://doi.org/10.1006/obhd.1997.2738

Roemer, J. E. (2010). Kantian equilibrium. Scand. J. Econ, *112* (1), 1–24. https://doi.org/10.1111/j.1467-9442.2009.01592.x

Samuelson, L. (2016, Fall). Game Theory in Economics and Beyond. *Journal of Economic Perspectives, 30(4)*, 107-130. DOI: 10.1257/jep.30.4.107

Schkade, D., Sunstein, C.R., & Kahneman, D. (2000). Deliberation about Dollars: The Severity Shift Empirical Study. *Columbia Law Review*, *100*(4), 1139-1175. https://doi.org/10.2307/1123539

Schuster, Stefan (2017). A New Solution Concept for the Ultimatum Game leading to the Golden Ratio. *Scientific Reports, 7*, 5642. https://doi.org/10.1038/s41598-017-05122-5

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How two share a task: corepresenting stimulus-response mappings. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1234-1246. https://doi.org/10.1037/0096-1523.31.6.1234

Sebanz, N., Rebbechi, D., Knoblich, G., Prinz, W., & Frith, C. D. (2007). Is it really my turn? An event-related fMRI study of task sharing. *Social Neuroscience*, *2*(2), 81-95. https://doi.org/10.1080/17470910701237989

Stolle, C.M., Gula, B., Yu, R., & Huang, Y. (2024). The impact of diversity on group-decision making in the face of the free-rider problem. *Judgment and Decision Making*, *19*, e4. https://doi.org/10.1017/jdm.2023.47

Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.

Thaler, R.H. (2016). Behavioral Economics: Past, Present, and Future. *American Economic Review*, *106*(7), 1577-1600. DOI: 10.1257/aer.106.7.1577

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and Prosocial Behavior: A Theoretical Framework and Meta-Analysis. *Psychological Bulletin*, *146*(1), 30-90. http://dx.doi.org/10.1037/bul0000217

Thielmann, I., Böhm, R., & Ott, M. Hilbig, B.E. (2021). Economic Games: An Introduction and Guide for Research. *Collabra: Psychology, 7(1),* 19004. https://doi.org/10.1525/collabra.19004

Tenenbaum, G., & Land, W. M. (2009). Mental representations as an underlying mechanism for human performance. *Progress in Brain Research*, *174*, 251-266. https://doi.org/10.1016/S0079-6123(09)01320-X

van der Wel, R.P.R.D., Becchio, C., Curioni, A., & Wolf, T. (2021, April). Understanding joint action: Current theoretical and empirical approaches. *Acta Psychologica*, *215*, Article 103285. https://doi.org/10.1016/j.actpsy.2021.103285

van Dijk, E., De Cremer, D., & Handgraaf, M.J.J. (2004). Social value orientations and the strategic use of fairness in ultimatum bargaining. *Journal of Experimental Social Psychology, 40*(6), 697-707. https://doi.org/10.1016/j.jesp.2004.03.002

van Dijk, E., & De Dreu, C.K.W. (2021). Experimental Games and Social Decision Making. *Annual Review of Psychology, 72*, 415-438. https://doi.org/10.1146/annurev-psych-081420-110718

van Ginkel, W. P. &, van Knippenberg, D. (2008). Group information elaboration and group decision making: The role of shared task representations. *Organizational behavior and Human Decision Processes*, *105*(1), 82-97. https://doi.org/10.1016/j.obhdp.2007.08.005

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., Hristova, D., Karlinsky, A., McEllin, L., Nijssen, S.R.R., Schmitz, L., & Wahn, B. (2017). Joint Action: Mental Representations, Shared Information and General Mechanisms for Coordinating with Others. *Frontiers in Psychology, 7*, Article 2039. https://doi.org/10.3389/fpsyg.2016.02039

Walter, V., Kölle, M., & Collmar, D. (2022). Measuring the Wisdom of the Crowd: How Many is Enough? *Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 90*, 269-291. https://doi.org/10.1007/s41064-022-00202-2

Wiese, W. (2017). What are the contents of representations in predictive processing? *Phenomenology and the Cognitive Sciences*, *16*, 715-736. https://doi.org/10.1007/s11097-016-9472-0

Yi, S.K.M., Steyvers, M., Lee, M.D., & Dry, M.J. (2012). The Wisdom of the Crowd in Combinatorial Problems. *Cognitive Science*, *36*, 452-470. https://doi.org/10.1111/j.1551-6709.2011.01223.x

Yu, R. (2016). Stress potentiates decision biases: A stress induced deliberation-to-intuition (SIDI) model. *Neurobiology of Stress*, *3*, 83-95. https://doi.org/10.1016/j.bbih.2024.100766

Zhang, Y., Yang, S., Chen, X., Bai, Y., & Xie, G. (2023), Reputation update of responders efficiently promotes the evolution of fairness in the ultimatum game. *Chaos, Solitons & Fractals*, *169*, Article 113218. https://doi.org/10.1016/j.chaos.2023.113218

Zhong, F., Kimbrough, S. O., & Wu, D.J. (2002). Cooperative Agent Systems: Artificial Agents Play the Ultimatum Game. *Group Decision and Negotiation*, *11*, 433-447. https://doi.org/10.1023/A:1020687015632

Zhao, K., & Smillie, L. D. (2015). The Role of Interpersonal Traits in Social Decision Making: Exploring Sources of Behavioral Heterogeneity in Economic Games. *Personality and Social Psychology Review*, *19*(3), 277-302. https://doi.org/10.1177/1088868314553709

Appendix A – Questionnaires

HEXACO Questionnarie

Answer options ranged from 1 (not at all) - 5 (very much):

1. I rarely hold a grudge, even against people who have badly wronged me.

2. People sometimes tell me that I am too critical of others.

3. If I knew that I could never get caught, I would be willing to steal a million dollars.

4. People sometimes tell me that I'm too stubborn.

5. Having a lot of money is not especially important to me.

6. People think of me as someone who has a quick temper.

7. My attitude toward people who have treated me badly is 'forgive and forget'.

8. I wouldn't use flattery to get a raise or a promotion, even if I thought it would succeed.

9. I am usually quite flexible in my opinions when people disagree with me.

10. If I want something from someone, I will laugh at that person's worst jokes.

11. For this question, please press 4.

12. I tend to be lenient in judging other people.

13. I would never accept a bribe, even if it were very large.

14. When people tell me that I'm wrong, my first reaction is to argue with them.

15. Most people tend to get angry more quickly than I do.

16. I think that I am entitled to more respect than the average person is.

17. I wouldn't pretend to like someone just to get that person to do favors for me.

18. Even when people make a lot of mistakes, I rarely say anything negative.

19. I'd be tempted to use counterfeit money, if I were sure I could get away with it.

20. I would get a lot of pleasure from owning expensive luxury goods.

21. I want people to know that I am an important person of high status.

Funnelled Suspicion Check

Answer options for rated questions ranged from 1 (not at all) - 10 (very much)

1. What do you think the purpose of the experiment was? What did we try to study?

2. What were you trying to do when responding to the offers?
   Did you have any particular goal or strategy?

3. How did your strategies differ when you were deciding alone vs. together?

4. How would you describe the behavior of the proposer?

5. How would you describe the behavior of your partner?

6. On a scale from 1-10 how selfish did the proposer seem to you?

7. How fairly do feel you were treated by the proposer? (1-10 rating scale)

8. How selfish did your partner seem to you? (1-10 rating scale)

9. How fairly do you feel your partner played? (1-10 rating scale)

10. Did you notice anything unusual about the proposer?

11. Did you notice anything unusual about your partner?

12. How likely do you think it is that the proposer was a computer? (1-10 rating scale)

Appendix B - Additional Tables and Results

The Tables below list statistical results of statistical tests that serve as requirements for

Table 7 lists the results of Levene's Tests for the assumption of homogeneity of variance. These were conducted as requirements for parametric (pairwise) mean comparison tests, i.e., repeated measures ANOVAs and t-tests. If the assumption of homogeneity of variance was not met, Wilcoxon signed-rank tests were computed instead of t-tests.

Table 8 to Table 13 list the results of Shapiro-Wilk Tests that were performed to test the assumption of normality of the data as requirements for parametric (pairwise) mean comparison tests, i.e., repeated measures ANOVAs and t-tests as well as for parametric correlation coefficient computations such as the Pearson correlation coefficient. If the assumptions of homogeneity of variance or the assumptions of normally distributed data were not met, Wilcoxon signed-rank tests were computed instead of t-tests. Likewise, Spearman correlation coefficients were computed as alternatives to the Pearson correlation coefficient if the Shapiro-Wilk tests indicated non-normally distributed data for the respective variables.

Table 7

Levene's Tests for homogeneity of variances in accumulated rewards, negotiated offers and rejection rates across all condition blocks

Variable (all blocks)	W	p-value	equality of variance
Rewards	0.73	0.48	True
Offers	1.35	0.26	True
Rejection Rates	0.86	0.43	True

Table 8

Shapiro-Wilk Tests for reward difference scores, offer difference scores, joint benefits and reward disparities between partners in the pre-baseline

Variable	W	p-value	normality of data
Reward Difference Scores	0.97	0.5	True
Offer Difference Scores	0.91	0.0	False
Joint Benefits	0.97	0.48	True
Reward Disparieites in Block1	0.94	0.27	True

Table 9

Shapiro-Wilk Tests for accumulated rewards in each condition block

Phase	W	pval	normal
Pre-Baseline 1	0.98	0.56	True
Joint Condition	0.95	0.09	True
Post-Baseline 2	0.95	0.09	True

Table 10

Shapiro-Wilk Tests for offers in each condition block

Phase	W	pval	normal
Pre-Baseline 1	0.85	0.0	False
Joint Condition	0.94	0.03	False
Post-Baseline 2	0.93	0.02	False

Table 11

Shapiro-Wilk Tests for rejection rates in each condition block

Phase	W	pval	normal
Pre-Baseline 1	0.95	0.07	True
Joint Condition	0.96	0.16	True
Post-Baseline 2	0.97	0.42	True

Table 12

Shapiro-Wilk Tests for personality scores and partner disparities in personality scores

Variable	W	p-value	normality of data
Honesty-Humility	0.97	0.3	True
H-H Disparities	0.88	0.02	False
Agreeableness	0.97	0.45	True
Agree Disparities	0.93	0.15	True

Table 13

Shapiro-Wilk Tests for ratings on the likelihood that proposer may be a computer

W	pval	normal
0.96	0.0	False

Appendix C - Participant Information Sheet

Appendix C - Participant Information Sheet

Information Sheet (Making Decisions Together)

Purpose of the study: This research will enable us to better understand the mechanisms by which we interact with each other in social and economic contexts.

Procedure: If you decide to participate, you will receive more detailed instructions about the specific task. You will be asked to complete a computer task in which you will play economic games with two other people for points for multiple rounds. Depending on how many points you earned during the game, you may receive an extra bonus of € 1 – 5 per person in addition to your regular reimbursement.
If one of the other participants does not show up, you will receive a show-up fee of € 5. You will also be asked several questions which will check your understanding of and attention to the task. Unsatisfactory responses to these checks may lead to your exclusion from the study. At the end of the task, you will be asked some questions regarding some characteristics about yourself. These data will be completely pseudonymized and it will be made sure they do not trace back to your identity.
Lead researcher: Natalie Sebanz

Contact information: ghorabt@ceu.edu, mcellinl@ceu.edu, sebanzn@ceu.edu

Potential risks: The risks associated with participation in this experiment are equivalent to those of using a computer. The study does not include any unpleasant sounds or images.

Your rights as a participant. Please note that by signing this form you are willingly consenting to participating in this study. Your participation is voluntary, and you are free to withdraw yourself and your data from this study at any time without any reason by telling one of the researchers you want to abort or by leaving the room. After completing the tasks, you will receive 10 EUR per hour as a compensation for your participation and a bonus from 1-5 EUR which will be calculated based on your performance in the task.

Data storage and protection. In this study, we will record information about your task performance (for example, how many points you collected). During the task you may be asked to communcate with your partner via a set up WhatsApp chat. We will not ask you to use any of your personal information but provide you with an arbitrary account that we previously have created. At the end of the task, you will be asked some questions regarding some characteristics about yourself. All these data are confidential. That means we use a number code to identify the data. The data will not be directly linked to your name or any other identifying information. Information about your identity is kept strictly separate from the code and data. The same is true for questionnaire data. You may be given a questionnaire asking you about your age, gender and about your experience of the task. A code will be assigned to these questionnaires, and they will not be directly linked to your name. We take secure storage of the collected data very seriously. The recorded data are accessible only to the researchers of the study and those to whom he or she explicitly grants access rights. Your data will be stored for 10 years. More information about your rights can be found at https://www.ceu.edu/privacy.

Further information and questions. Please feel free to ask further questions about the study. It is important to us that you have received all the information you need to decide whether to participate in this study. If you would like to participate, please sign the attached consent form at the corresponding rows attached to this Information Sheet. Otherwise, please tell the researchers that you would like to withdraw from the study.

Contact information	
Lead researcher: Natalie Sebanz	Researchers: Tiam Ghorab, Luke McEllin
Email: SebanzN@ceu.edu	Email: GhorabT@ceu.edu, McellinL@ceu.edu
Institute: Central European University Private University, Department of Cognitive Science	