MASTERARBEIT | MASTER'S THESIS

Titel | Title

Constructing Conflict Templates: Task Format and Individual Differences Affect the Mental Representation of Social Dilemmas

verfasst von | submitted by

Tiemo Jakob Steinmann B.Sc. BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2026

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 840
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Psychologie
Betreut von | Supervisor:	Univ.-Prof. Dipl.-Psych. Dr. Robert Böhm

Constructing Conflict Templates: Task Format and Individual Differences Affect the Mental Representation of Social Dilemmas

Abstract

People do not always respond to social dilemmas based on an objective payoff structure alone, but based on how they mentally represent the interdependence involved. Building on research on conflict templates, the present preregistered study examined both whether individuals converge on a small set of prototypical 2x2 game structures and how these representations are shaped by task characteristics and individual differences. A total of 268 participants generated a conflict scenario and constructed its payoff structure under varying conditions of abstraction level (payoff matrix vs. verbal ranking) and perspective sequence (self-first vs. partner-first), with an additional simultaneous payoff-matrix control group. Consistent with the perceptual convergence hypothesis, participants disproportionately often created structures corresponding to Prisoner's Dilemma, Chicken, Assurance, and Maximizing Difference games. At the same time, the way the task was operationalized influenced the resulting representations: separating perspectives reduced symmetry, the payoff matrix format produced more symmetric and more competitive game structures than verbal ranking. In contrast, perspective sequence did not affect competitiveness. Individual differences showed weaker effects: social value orientation was positively associated with game harmony, although this effect was small and less robust in sensitivity analyses, whereas beliefs in a zero-sum game were not significantly related to competitiveness. Overall, the findings suggest that conflict templates are neither arbitrary nor fixed, but constrained mental representations that are still sensitive to situational cues and how interdependence is elicited.

Keywords: social dilemmas; conflict templates; interdependence; economic games; social value orientation

Table of Content

Constructing Conflict Templates: Task Format and Individual Differences Affect the Mental Representation of Social Dilemmas

Introduction

Some of the most significant problems facing modern society, ranging from the interpersonal to the global level, have their roots in a conflict between personal and collective interests (Kollock, 1998; Van Lange et al., 2013). On a global scale, humanity faces existential threats such as the climate crisis, caused by greenhouse gas emissions, overpopulation, and resource scarcity (Dawes, 1980; Harring & Krockow, 2021; Van Lange et al., 2013). Equally urgent are major public health crises, including the COVID-19 pandemic, in which individuals had to weigh their personal inconvenience of mask-wearing, physical distancing, and vaccination against the collective benefit of the measures reducing viral spread (Korn et al., 2020; Van Lange & Rand, 2022), as well as the rise of antimicrobial resistance, where the overuse of antibiotics threatens the global effectiveness of medications (Harring & Krockow, 2021). Furthermore, society is increasingly confronted with misinformation, which requires individuals to make cognitive efforts to verify facts before sharing them online (Van Lange & Rand, 2022). Beyond these global challenges, similar dynamics appear in everyday life: deciding whether to commute by car or public transport (Fujii et al., 2001), choosing whether to donate to public journalism and charity organisations (Kollock, 1998; Komorita, 2019), or voting in elections (Van Lange et al., 2013).

All of these scenarios are examples of social dilemmas, that is, mixed-motive situations characterized by a fundamental tension between individual and collective rationality generated by a partly negative interdependence (Deutsch, 1949) between individual and collective outcomes (Capraro, 2013; Kollock, 1998). Specifically, a social dilemma is characterized by two key features: each individual receives a higher payoff for a socially deviant, selfish decision than for a socially cooperative decision, regardless of what others do.

Yet ultimately, everyone is worse off if everyone acts deviantly than if everyone cooperates (Dawes, 1980). Essentially, individually rational behaviors lead to a deficient equilibrium and a suboptimal collective outcome, demonstrating that the pursuit of typically short-term individual interests is fundamentally at odds with often long-term collective interests (Kollock, 1998; Van Lange, 2014; Van Lange et al., 2013).

An important mechanism creating a social dilemma is interdependence, the pattern of mutual control that two or more individuals have over their own and each other's outcomes (Balliet & Lindström, 2023; Columbus & Molho, 2022). It describes how one person's behavior directly affects the results experienced by others in a social interaction (Columbus et al., 2021) and thereby contributes to the severity of the social dilemma.

While social interactions vary in their degree of interdependence, ranging from complete independence to total mutual dependence, social dilemmas are specifically characterized by negative interdependence, where individual and collective interests are at odds (Balliet et al., 2017; Kelley et al., 2009; Kelley & Thibaut, 1978). However, it is not the objective structure of interdependence but the involved people's subjective beliefs and inferences about interdependence that guides their behavior and allows them to differentiate between social situations (Columbus et al., 2021; Gerpott et al., 2018; Thielmann et al., 2021). Because individuals rely on differing subjective mental models, two people facing the exact same situation with the same objective incentives may perceive their interdependence entirely differently. As a result, they are essentially tailoring their behavior to different structures with different rules and underlying reasoning.

Existing research suggests that people rely on subjective mental models, or conflict templates, when navigating social dilemmas (Halevy et al., 2012; Halevy & Katz, 2013), but less is known about the cognitive processes or the influence of situational and personal factors involved in forming these representations. The present preregistered study extends this work by examining not only which conflict templates people use to mentally represent social interactions, but also how these templates are constructed. Therefore, it tests whether the task format, specifically abstraction level (payoff matrices vs rankings of verbal descriptions), splitting perspectives and varying perspective sequence (self first vs. partner first), influence the resulting representation of conflict. In addition, the study examines whether individual differences, specifically social value orientation and beliefs in a zero-sum game, are associated with systematic differences in how individuals construe a situation, ranging from more competitive to more harmonious forms of interdependence. By linking cognitive evaluation processes with stable individual differences, the present study aims to clarify why objectively identical social dilemmas may be perceived differently across individuals, and how these differences are reflected in the economic game structures people construct. If individuals systematically differ in how they construe the same objective conflict structure, then findings based on fixed economic games may also capture variation in how the situation itself is mentally represented in addition to cooperative tendencies. The following sections outline the theoretical foundations of social dilemmas and conflict templates, and develop the specific hypotheses tested in this study.

Theoretical Background

Social Dilemmas and Economic Games

To investigate human cooperative behavior in social dilemmas, researchers frequently model these situations using economic games drawing on two complementary frameworks, game theory and interdependence theory (Thielmann et al., 2021). Economic games are abstract, mathematical models of strategic interactions specifying the involved players with their available choices, as well as the payoffs or utilities associated with every possible choice combination (Camerer, 2003b; Thielmann et al., 2021). These formal structures provide the basis for analyzing how individuals' outcomes depend on each other's choices, which is central to interdependence theory. While classical game theory sets the framework for the analysis of economic games from a rational mathematical standpoint (Aumann, 2018;

Fudenberg & Tirole, 2010; Neumann & Morgenstern, 2007) and was recently generalized to also explain experimentally observed violations using psychological principles (behavioral game theory; see Camerer, 2003b, 2018), interdependence theory provides a psychological lens for analyzing strategic interactions. One of the key ideas in interdependence theory, which is highly relevant for this study, is that individuals do not simply react to the objective payoffs in a game. Rather, they psychologically transform this “given” matrix of outcomes into an “effective” matrix based on their subjective perceptions, personal values and motives, which then dictates their behavior (Kelley et al., 2009; Kelley & Thibaut, 1978). To examine such subjective construals of interdependence in a formally tractable way, the present study focuses on 2x2 economic games.

Games consisting of two players, each of whom has exactly two alternatives or choices of strategy are called 2x2 games and make for the simplest representation of a strategic relationship or conflict between parties (Kilgour & Fraser, 1988; Robinson & Goforth, 2006). The cross-product of these two sets of alternatives is the strategy space, hence a 2x2 game mathematically yields exactly four possible strategy combinations or outcomes. These games are typically represented in the so called normal form using a bi-matrix (the payoff matrix, see Figure 1) visually mapping the players’ independent strategy options to the specific payoffs associated with each of the four possible outcomes (Robinson & Goforth, 2006; Thielmann et al., 2021). Strategically important solution concepts to characterize the underlying incentive structure of a given 2x2 game, are dominant strategies and Nash equilibria (Robinson & Goforth, 2006; Thielmann et al., 2021). A dominant strategy is a player’s generally best strategy yielding the best outcome regardless of the other player’s decisions (Halevy et al., 2012; Thielmann et al., 2021), while a Nash equilibrium describes the strategy combination in which no player can obtain a higher outcome by changing their individual strategy given the other player sticks to their chosen strategy (Kreps, 2018).

Figure 1

Normal Form Representation of a 2x2 Game Structure

		Player 2	
		Option 1	Option 2
Player 1	Option 1	4,3	2,1
	Option 2	1,2	3,4

Note. The first value in each cell represents player 1's payoff, the second value represents player 2's payoff.

Other concepts to describe the properties of economic games, looking at the complete payoff structures instead of single strategies or strategy combinations, include evaluating symmetry and game indices. A 2x2 economic game is considered symmetric when both players face the same strategic problem and are thereby entirely interchangeable (Kilgour & Fraser, 1988). In a matrix form, this means that both players have the exact same set of available strategies and receive the same associated payoffs for corresponding combinations of choices (Robinson & Goforth, 2006; e.g. the game in Figure 1 is symmetric). Game indices are functions of the specific payoffs in a game, condensing the entire payoff matrix into single metrics to quantify underlying characteristics like cooperativeness, how conflictual or harmonious the players' interests are, efficiency measures, and components of interdependence (Gächter et al., 2024; Stivers et al., 2025).

Although 2x2 games represent the simplest form of strategic interaction (Robinson & Goforth, 2006), there is vast complexity within this space, given that there are 726 possible options how two players can rate the favorability of the four resulting outcomes in an ordinal manner (Kilgour & Fraser, 1988). This is why 2x2 games are the most commonly employed models in research on human conflict and cooperation (Gächter et al., 2024; Halevy et al., 2012; Kilgour & Fraser, 1988; Villiger et al., 2022) and why this study also focuses only on one-shot 2x2 games. However, even within this constrained space of 2x2 games, individuals may construe the same objective situation in different ways, highlighting the importance of understanding how people mentally represent interdependence and conflict.

Conflict Templates and Mental Representation of Interdependence

While economic games are widely used to operationalize real world scenarios in research on cooperative and competitive behavior (Thielmann et al., 2021), individuals still need to form their own mental representation of the situation in order to guide their behavior. This directly parallels the distinction between the given and effective matrices in interdependence theory. Because the process of building these representations involves reducing complexity, individuals may misrepresent the objective structure of a given situation. However, it is their subjective representation that ultimately guides their behavior (Devetag & Warglien, 2008). Accordingly, it is important to examine how exactly individuals form mental representations of conflict situations and which factors shape them. Otherwise, conclusions about cooperative behavior drawn from research using fixed economic games may be constrained, such that a given game may represent the situation well for some, but not for others. This concern is reflected in one of the most prominent questions in game theory: what game people think they are playing when confronted with a social situation (Camerer, 2003a).

More generally, behavior in social situations is strongly shaped by beliefs about interdependence (Gerpott et al., 2018), meaning that the same scenario may be understood differently by different individuals (Halevy & Phillips, 2015). It is therefore important to understand how such beliefs about interdependence are formed and how both environmental cues and individual differences contribute to this process.

People naturally map their subjective mental representations of outcome interdependence in social interactions on so called conflict templates (Halevy & Katz, 2013; Halevy & Phillips, 2015). They capture how a person perceives the relationship or structural contingencies between their own actions, their partner's actions, and the resulting outcomes for both parties. Focusing on which specific conflict templates people use, Halevy et al. (2012) proposed the perceptual convergence hypothesis, which assumes that people converge to a small set of only four 2x2 games when forming their mental representations. This hypothesis is based on three key concepts: preference for symmetry (e.g. in art, faces, and objects; see Freyd & Tversky, 1984; Humphrey, 1997; Reber, 2002; Reber et al., 2004; Rhodes et al., 1998), preference for cooperative partners (derived from people costly punishing uncooperative individuals; see Fehr & Gächter, 2000, 2002), and correspondence between scholars' and lay people's theories of conflict (Cargile et al., 2006; Levy et al., 2006). Combining these concepts results in four prototypical games fulfilling the criteria that are found to be used as conflict templates disproportionately often: Maximizing Difference (MD; detailed description in Marwell et al., 1969), Stag Hunt/Assurance (detailed description in Skyrms, 2004), Chicken (detailed description in Rapoport & Chammah, 1966), and Prisoner's Dilemma (PD; detailed description in Rapoport & Chammah, 1965). Corresponding payoff matrices of these four games can be found in Figure 2.

Strategically they are characterized as follows: in MD, cooperation is a dominant strategy for both players, mutual cooperation is the only Nash equilibrium; in Assurance, copying the partner is the best strategy, resulting in two Nash equilibria in mutual cooperation and mutual defection; in Chicken, doing the opposite of the partner is the best strategy, resulting in two Nash equilibria in both of which one player cooperates and the other defects; in PD, defection is a dominant strategy for both players, mutual defection is the only Nash equilibrium.

Figure 2

Payoff Matrices of the Four Prototypical Games

		Player 2: Cooperate	Player 2: Defect
Player 1	Cooperate	4,4	2,3
	Defect	3,2	1,1

Maximizing Difference

		Player 2: Cooperate	Player 2: Defect
Player 1	Cooperate	4,4	1,3
	Defect	3,1	2,2

Assurance

		Player 2: Cooperate	Player 2: Defect
Player 1	Cooperate	3,3	2,4
	Defect	4,2	1,1

Chicken

		Player 2: Cooperate	Player 2: Defect
Player 1	Cooperate	3,3	1,4
	Defect	4,1	2,2

Prisoner‘s Dilemma

However, limiting the problem space to only these four conflict templates may constitute an overly strong reduction. As Balliet et al. (2017) stated, the perceptual convergence hypothesis overlooks the possibility of asymmetric interdependence, even though detecting asymmetry in social situations is highly beneficial for the actors. Moreover, prior work has primarily examined which conflict templates people endorse but has paid less attention to how task features shape the construction of these templates (Halevy et al., 2012, 2014; Halevy & Katz, 2013; Halevy & Phillips, 2015). The present study therefore also considers how specific aspects of the task format and instructions influence the mental representation of interdependence.

Variations in Task Format: Abstraction Level and Perspective Sequence

In economic games, people are susceptible to framing. For example, they make different decisions in what is, mathematically speaking, the same game, depending on how the game is presented (Camerer, 2003b). This suggests that not only the content of a conflict template, but also the way it is presented, may influence how people represent interdependence. Varying the task format when asking people to evaluate outcomes of a conflict scenario might have effects in a similar way as framing and therefore influence which conflict templates people use. This study specifically focuses on potential effects of the abstraction level in the way outcomes are evaluated and the split of perspectives, also varying the order in which their own and their partner's perspective are evaluated by people.

The preference for symmetry as described in the perceptual convergence hypothesis relates primarily to symmetry as a perceptual feature (Halevy et al., 2012; Halevy & Chou, 2014). The authors evaluated conflict templates by letting people fill in empty payoff matrices, a format in which symmetry is visually present. Hence, it is possible that the preference for symmetry is not grounded in the interpretation of interdependence itself, but rather reflects a preference for visual symmetry in a payoff matrix. Halevy et al. (2012) themselves stated that the preference for symmetry exists at a high level of abstraction. Thus, reducing abstraction and making visual symmetry less available by not giving people empty payoff matrices but instead asking them to order verbal outcome descriptions might also reduce their preference for symmetric conflict templates. Similarly, splitting the evaluation of outcomes to separate perspectives, deconstructing the full payoff matrix with eight outcome values into two separate ones with four outcome values each for oneself and the interaction partner also reduces the availability of symmetry as a visual cue and might therefore reduce people's preference for symmetry in constructing their conflict templates as well.

Beyond abstraction level and visual symmetry, the sequence in which perspectives are evaluated may further shape how conflict templates are constructed. Perspective taking, the ability to represent the thoughts, feelings, and intentions of another person accurately (Epley & Caruso, 2009) has been found to work as a catalyst for prosocial behavior and cooperation (Erle & Topolinski, 2017; Ortiz-Riomalo et al., 2021), reducing egocentric biases in judgment (Epley et al., 2005), and fostering cooperative attitudes in international relations and workplace environments (Casler & Groves, 2023; Škerlavaj et al., 2018). Vice versa, cooperative interactions can enhance the ability for flexible perspective taking (Li et al., 2019). However, the relationship between perspective taking and cooperation is highly context-dependent, as in competitive environments it can ironically even escalate conflict (Epley et al., 2005). This dual nature of perspective taking might directly influence the formation of conflict templates, because assessing game-theoretic payoffs inherently engages perspective taking and mentalizing processes (Chang et al., 2023). In 2x2 games, evaluating outcomes requires individuals to consider their interaction partner's likely intentions and thoughts in addition to their own consequences. Given that conflict templates have been shown to be temporarily malleable when different motivational goals are induced (Halevy & Katz, 2013), one possible mechanism is that construction of a conflict template depends on when and how perspective taking is prompted during evaluation. Specifically, requiring participants to evaluate their partner's outcomes first may induce perspective taking earlier in the process, potentially reducing egoistic motives and promoting more harmonious representations. In contrast, evaluating one's own outcomes first may foster a more self-centered construal, with perspective taking entering only at a later stage.

Individual Differences: Zero-sum Beliefs and Social Value Orientation

Halevy & Phillips (2015) found some level of consistency in which conflict templates people use across time and different situations, suggesting that individual differences contribute to shaping how people mentally represent conflict scenarios. Adding to this, Balliet & Lindström (2023) also state that people make the necessary inferences about their interdependence in social interactions by using environmental cues as well as general prior beliefs they have about interdependence.

Therefore, in addition to environmental cues in task format and priming, individual factors might also play a role in people's perception of social interactions and therefore in the construction of conflict templates. This study focuses on two aspects of personality, which were already shown to correlate with cooperative behavior and conflict templates in a meaningful way, and which were also examined in the original study on the perceptual convergence hypothesis (Halevy et al., 2012): social value orientation (SVO) and beliefs in a zero-sum game (BZSG).

SVO is a stable personality trait reflecting the magnitude of concern people have for others (Balliet et al., 2009; Bogaert et al., 2008). It is broadly studied within research on human cooperation and was shown to be a fundamental predictor of prosocial behavior, showing the highest correlation with cooperative behavior across all games in a comparison of over 50 personality constructs (Thielmann et al., 2020). This is attributed to more prosocial individuals evaluating interactions through a lens of collective rather than individual rationality, which less prosocial individuals tend to use (Bogaert et al., 2008). This idea is consistent with interdependence theory (Kelley & Thibaut, 1978), which proposes that people's social preferences and norms, such as prosociality, shape the transformation of the given matrix into the effective matrix. Therefore, it is not surprising that SVO was found to negatively correlate with the perception of conflict (Gerpott et al., 2018). Accordingly, in this study, SVO is expected to influence the cooperativeness of the conflict templates people construct, as captured by game indices. Although Halevy et al. (2012) did not find a correlation between SVO and the endorsement of the four prototypical games, this study extends this work by allowing participants to construct conflict templates themselves and by quantifying their properties using multiple game indices.

Another important construct in research on human cooperation are zero-sum beliefs, which reflect the subjective assumption that one party's gain is only possible at another's expense, and function either as domain-specific judgments or even as a generalized worldview (Andrews Fearon & Götz, 2024; Davidai & Tepper, 2023). Davidai et al. (2022) demonstrate that zero-sum beliefs are fundamentally tied to expectations of interpersonal conflict, aggression, and animosity, suggesting that people with higher BZSG might generally see social interactions as more conflictual. BZSG operate as an internal "lay game theory" dictating how people subjectively construe their strategic interdependence with others (Andrews Fearon & Götz, 2024), leading people with strong BZSG to mentally transform ambiguous social interactions into highly competitive conflict templates – specifically the PD and Chicken games (Halevy et al., 2012). The present study extends this research on the relation between BZSG and conflict templates by not only examining correlations with the prominence of a small set of predefined games, but by quantifying the degree of conflict in every possible conflict template people might construct.

Taken together, these situational and dispositional factors provide a framework for predicting how people construct mental representations of conflict in terms of economic games, which is formalized in the following hypotheses.

Overview & Hypotheses

Based on the current status of research on conflict templates and the highlighted gap, this study aims to investigate in detail how individuals mentally represent conflict situations in terms of economic game structures. Accordingly, the present study examines conflict template construction at two levels: first, it tests whether individuals disproportionately often converge on a small set of prototypical game structures when representing conflict situations. Second, it investigates whether the structural properties of these constructed templates, particularly symmetry and competitiveness or harmony, vary as a function of task characteristics and individual differences. This approach allows the study to address both which conflict templates people construct and how these representations differ structurally. The following hypotheses can be derived:

H1 Across all conditions, participants will disproportionately more often create payoff structures corresponding to Prisoner’s Dilemma, Chicken, Assurance and Maximizing Difference games than other possible structures.

H2 Participants who complete payoff ratings for self and partner separately will create more asymmetric payoff structures than participants in the control group, who rate both perspectives simultaneously.

H3 Participants using the more abstract numeric format of ranking outcomes (i.e., filling out payoff matrices) will create more symmetric payoff structures than participants using the verbal format (i.e., ranking their verbal outcome descriptions).

H4 Participants who evaluate their own perspective first will create more competitive payoff structures (e.g., Prisoner's Dilemma, Chicken) than participants who evaluate their partner's perspective first.

H5 Participants with higher BZSG will create more competitive payoff structures (e.g., Prisoner's Dilemma, Chicken) than participants with lower BZSG.

H6 More prosocial participants will create less competitive structures (e.g., Prisoner's Dilemma, Chicken) than less prosocial participants.

Method

Ethics and Open Science

The study received a positive ethics vote from the Departmental Review Board of the Department of Occupational, Economic, and Social Psychology at the University of Vienna (DRB number 2025/W/022). All participants provided informed consent before taking part in the study. The hypotheses, methods, materials, sample size and underlying rationale, exclusion criteria, and planned analyses of the study were preregistered on the Open Science Framework (https://osf.io/832de) before recruiting participants. The preregistration, anonymized data, and analysis code are available online at https://osf.io/5bzwg.

Experimental Design

To test the hypotheses, the study employed a randomized 2x2 factorial between-participants design with a separate control group, manipulating the factors abstraction level (payoff matrix or verbal ranking) and perspective sequence (self-first or partner-first). This resulted in five groups, participants were randomly assigned to: the control group (CG), the verbal ranking self-first group (VS), the verbal ranking partner-first group (VP), the payoff matrix self-first group (PS), and the payoff matrix partner-first group (PP).

In all conditions, participants were asked to come up with a 2x2 conflict scenario and label the two options each of the two players has. Participants in the payoff matrix condition (groups PS, PP) were presented with empty payoff matrices and asked to fill in payoffs from one to four for both themselves and their interaction partner. Participants in the verbal ranking condition (groups VS, VP) were asked to describe each of the four outcomes of their self-generated 2x2 scenario in a short sentence. They were then presented with these outcome descriptions and asked to rank them from most favorable to least favorable for both themselves and their interaction partner.

In both abstraction level conditions, participants in the self-first condition evaluated the favorability of their own outcomes first, and their partner's second, while participants in the partner-first condition evaluated their partner's outcomes first and their own ones second.

Following the original design of Halevy et al. (2012), participants in the separate CG evaluated both perspectives simultaneously using the payoff matrix format.

Sample

An a priori power analysis for detecting medium-sized effects with a power of .80 and an alpha level of .05 was conducted for the preregistered analyses, using *G*Power* (Faul et al., 2007). The power analysis resulted in a target sample size of 212 participants, double-weighting the CG relative to each of the four experimental groups and including an additional 10% to account for potential dropouts or poor data quality.

Participants were recruited from October to December 2025 through personal and professional networks as well as on online survey exchange platforms SurveyCircle and SurveySwap. A total of 354 participants completed the survey. Applying the preregistered exclusion criteria, 55 participants were excluded for unrealistically low completion times of less than five minutes, five were excluded for wrong answers in the attention check, 20 were excluded for reporting having not understood the instructions fully, and six were excluded for illogical answers in the conflict scenario imagination task.

The 268 participants included in the analyses ranged in age from 18 to 65 years, with a mean age of 26.7 years (*SD* = 6.8), 171 (63.8%) participants were female, 95 (35.4%) were male, and two (0.7%) were diverse. Participants came from 20 different countries, Germany (57.8%), Austria (23.9%), and the United States (3.0%) being the most frequent ones. Forty-nine point three percent of the sample had a bachelor's degree, 26.5% finished high school, and 17.9% had a master's degree. In regards to game theory, 59.3% reported having at least some knowledge. Ninety-three participants were assigned to the CG, 53 were assigned to the VS, 43 were assigned to the VP, 38 were assigned to the PS, and 41 were assigned to the PP group.

Materials & Measures

Conflict Representation Task

The conflict representation task started with an introductive text to explain the general concept of 2x2 conflict scenarios involving two players with two distinct behavior options, whose choice for one of the options influences the outcome for both players. Three examples from different contexts were given and participants were asked to briefly describe a 2x2 scenario they came up with, as well as the two options both players have to choose from. Afterwards, participants were asked to imagine being one of the players and evaluate the favorability of the four resulting outcomes for both themselves and their partner, with the evaluation format and perspective sequence depending on the group they are assigned to: participants in the verbal ranking condition verbally described the four outcomes in short sentences each and were asked to rank these sentences in order of their favorability. Participants in the payoff matrix condition were presented with empty payoff matrices and were asked to fill in payoff values from one to four. For both conditions, participants in the self-first condition evaluated the outcomes from their own perspective first and from their partner's perspective second, participants in the partner-first condition evaluated the outcomes vice versa. Participants in the CG evaluated the outcomes for both perspectives simultaneously in the payoff matrix format.

To ensure comparability of the outcome favorability evaluations across conditions, verbal rankings were transformed into payoff values, assigning a payoff of four to the first rank, a payoff of three to the second rank and so on, and individual-perspective payoff matrices were combined, yielding a normal form representation (see Figure 1) of the game structures participants created as a result of the conflict representation task in every condition.

The full instruction texts and examples of the conflict representation task can be found in the appendix.

Game Properties

To evaluate the game structures' properties, several game indices were used: a self-developed asymmetry index, averaging the absolute values of differences between the players' payoffs across outcomes, was used to quantify the level of asymmetry between the players' payoffs with higher values indicating more asymmetry. Symmetry was also coded as a dichotomous variable with all game structures yielding an asymmetry index of zero being coded as symmetrical. The Rapoport *K* index (Rapoport & Chammah, 1965) and the game harmony index (Zizzo & Tan, 2007, 2011) were used to evaluate the game structures' cooperative potential, with higher values indicating more cooperative potential and less competition. Correspondence (Kelley & Thibaut, 1978; Kelley et al., 2009) and conflict of interest (Axelrod, 1967) were used to measure the degree of goal compatibility. Game riskiness (Au et al., 2012; Ng & Au, 2016; Ng et al., 2024) was used to measure the players' perceived uncertainty in choosing the cooperative behavior option. Actor control, partner control, joint control, and interdependence (Kelley & Thibaut, 1978; Kelley et al., 2009;

named following Stivers et al., 2025) were used to disentangle the interdependence structure into its components. Finally, price of anarchy and price of stability (Mak & Rapoport, 2013) were used to evaluate the structure of equilibria in the created games. All indices designed for symmetric games and in that relying on only one set of four payoffs equal for both players were computed for both perspectives separately and then averaged to allow for comparison of asymmetric game structures as well.

The game structures participants created were classified as matching one of the four games from the perceptual convergence hypothesis using the payoff hierarchy and the underlying strategic properties of the game structures. The games were also grouped into thematic categories using Microsoft Copilot version 1.25121.84.0. The AI was prompted to go through the scenario description and behavior option inputs participants gave, to search for thematic clusters, and to assign every scenario to one of the clusters. The assignment was manually checked, cases in which participants referred to one of the given examples instead of giving their own description were manually assigned to one of the clusters.

Detailed information on how all used indices were computed and how games were classified and clustered can be found in the appendix.

Zero-sum Beliefs

To measure participants' BZSG, the BZSG scale developed by Różycka-Tran et al. (2015) was used. The BZSG scale is designed to assess the perceived antagonistic nature of social relations, measuring how strongly people believe that the world contains only a finite amount of goods, making one person's success only possible at the expense of others. Originally containing 12 items, the scale was reduced to eight positively worded items after psychometric analyses (Różycka-Tran et al., 2015). Items (e.g. "Successes of some people are usually failures of others") are rated on a 7-point Likert scale ranging from "strongly disagree" to "strongly agree" and averaged to one scale score afterwards, with higher scores indicating higher BZSG. Internal consistency and factor structure were evaluated in the present sample with the eight-item scale showing a good reliability, Cronbach's α = .83, McDonalds's ω = .89. More information on the scale's items and psychometric properties can be found in Różycka-Tran et al. (2015).

Social Value Orientation

To assess participants' prosociality, the SVO slider measure developed by Murphy et al. (2011) was used. It quantifies the magnitude of concern individuals have for the welfare of others in allocating resources independently. The measure was developed to overcome the limitations of categorical measures, now treating SVO as a continuous construct, and consists of six primary and nine optional secondary items, in each of which participants have to choose their favorite of nine resource allocations between themselves and an imaginary other. The results are scored by calculating the SVO angle with higher angles indicating a more prosocial SVO. The angles can also be mapped to four categories: altruistic, prosocial, individualistic, and competitive (Murphy et al., 2011). For this study, only the six primary items were used, transitivity of participants' answers was evaluated. More information on the calculation of the SVO angle and the categorization can be found in Murphy et al. (2011).

Procedure

After giving informed consent, participants were randomized to one of the five groups, with the CG being double-weighted relative to the four experimental groups to ensure sufficient group size for the planned analyses. The general structure of 2x2 conflict scenarios was explained by providing examples from different domains and participants were asked to generate a scenario of their choice, shortly describe the scenario and name the two behavior options each of the two players have.

After their input, participants in the payoff matrix condition (groups CG, PS, PP) were presented with empty payoff matrices and asked to fill in payoffs from one to four for both themselves and their interaction partner. In the CG, participants evaluated both perspectives simultaneously, entering two values for each cell of the payoff matrix. In contrast, participants in the self-first (PS) and partner-first conditions (PP) completed the perspectives separately: they first evaluated either their own or their partner's payoffs, respectively, and were then presented with a second empty payoff matrix to provide the remaining perspective. In between the two matrices, participants had to answer an attention check, that also served as a distractor to facilitate the perspective change. Participants in the CG had to complete the attention check before filling in the payoff matrix.

Participants in the verbal ranking condition (groups VS, VP) were asked to describe each of the four outcomes of their self-generated 2x2 scenario in a short sentence. They were then presented with these outcome descriptions and asked to rank them from most favorable to least favorable for themselves first and for their partner second (VS) or vice versa (VP). They had to complete the same attention check/distractor item as the groups in the payoff matrix condition in between the rankings for the two perspectives.

After the conflict representation task, participants' SVO and BZSG were assessed across all groups. Afterwards, participants were asked if they had fully understood all tasks and questions, and gave sociodemographic information on gender, age, country of origin, education, occupation, and prior knowledge in game theory.

Finally, participants had the opportunity to leave comments and were shortly debriefed.

Data Preparation and Analyses

To test the six preregistered hypotheses, the preregistered analysis methods were used. However, due to specific properties of some of the game structures participants created, there were some deviations from the preregistration.

First, in evaluating the outcomes' favorability, participants in the payoff matrix condition were specifically instructed to use every number from one to four exactly once per perspective to ensure strictly ordinal payoff structures. However, 85 participants did not follow this instruction but evaluated two or more payoffs equally in at least one of the perspectives. Additionally, 38 participants in the verbal ranking condition used the exact same phrase twice to describe the two mixed strategy outcomes, in which the two players choose different behavior options. This made it impossible for them to distinguish between the two when they were presented with their outcome descriptions and were asked to rank them, and made their payoff hierarchy ambiguous. Taken together, it does not seem appropriate to rely solely on the concise payoffs and their hierarchy in analyzing the game structures' properties using indices computed directly from payoffs and game types classified based on the payoff hierarchy. Instead, the underlying strategic structure of the games should also be considered, analyzing dominant strategies and equilibria. This also changes the probabilistic base rate of the four games in the perceptual convergence hypothesis from four in 576 possible combinations of two strictly ordinal rankings of four outcomes each to four in 78 strategically distinct games (Kilgour & Fraser, 1988; Robinson & Goforth, 2006).

Second, the self-developed asymmetry index was constructed in line with the preregistration. However, the index did not meet basic criteria for meaningful interpretation (e.g. continuity, normality) and was therefore discarded. Deviating from the preregistration, all reported analyses on symmetry of the game structures are based only on the dichotomous symmetry coding. Finally, some of the preregistered game indices were originally developed only for specific games, mostly for the characterization of PD games. These are Rapoport *K* (Rapoport, 1967), conflict of interest (Axelrod, 1967), and game riskiness (Au et al., 2012). Applying these indices for games other than their original intent yields values outside their usual boundaries or makes computation impossible (e.g., due to division by zero), severely limiting interpretability. Therefore, deviating from the preregistration, those indices were not used for any analyses.

All analyses were performed using R statistical software version 4.5.2 (R Core Team, 2025). To evaluate the internal consistency and factor structure of the BZSG scale, the packages psych version 2.6.1 (Revelle, 2026) and lavaan version 0.6-21 (Rosseel et al., 2025) were used. For robust statistical methods, the package WRS2 version 1.1-7 (Mair & Wilcox, 2020) was used.

Results

Descriptive Statistics

Random assignment resulted in comparable groups. Demographic variables and individual-difference measures (BZSG, SVO) were evenly distributed across conditions (see Table 1). After confirming that the randomization was successful and the groups were comparable, the structural properties of the game representations participants produced were examined. Table 2 summarizes these properties, including game types, symmetry, and game harmony.

Hypothesis 1 – Distribution of Conflict Templates

The hypothesis that participants would create structures of Chicken, Assurance, MD, and PD disproportionately often was tested with a binomial test comparing the observed frequency with the theoretical base rate.

Using the payoff hierarchy of the game structures participants created and applying the payoff matrices given by Halevy et al. (2012) to identify the four expected games PD, MD, Chicken, and Assurance, 34 of the 268 game structures were classified as one of the four expected games. A binomial test showed that the proportion was significantly higher than the base rate of 4/576, $p < .001$, with an observed proportion of 12.7%, 95% CI [9.5%, 100%], supporting the hypothesis.

Table 1

Participant Characteristics by Experimental Condition and Control Group

Variable	Experimental conditions				Control group	Total
	Payoff matrix		Verbal ranking			
	Self-first	Partner-first	Self-first	Partner-first		
	n = 38	*n* = 41	*n* = 53	*n* = 43	*n* = 93	*N* = 268
Age in years	26.0 (4.8)	25.7 (5.5)	26.1 (5.1)	27.7 (8.8)	27.4 (7.8)	26.7 (6.8)
Gender	19 (50.0%)	27 (65.9%)	36 (67.9%)	28 (65.1%)	61 (65.6%)	171 (63.8%)
BZSG score	3.99 (1.21)	3.70 (1.01)	3.80 (1.28)	3.84 (1.06)	3.85 (1.25)	3.84 (1.18)
SVO angle	27.77 (12.93)	30.39 (12.14)	30.99 (11.68)	29.42 (12.79)	28.01 (13.01)	29.15 (12.55)
Game theory knowledge	23 (60.5%)	28 (68.3%)	29 (54.7%)	27 (62.8%)	52 (55.9%)	159 (59.3%)

Note. Continuous measures are reported as *M* (*SD*), gender refers to female participants, reported as *n* (percentage), game theory knowledge refers to participants who reported having at least some knowledge, reported as *n* (percentage).

Table 2

Game Structure Properties by Group

Game property	Experimental condition				Control group	Total
	Payoff matrix		Verbal ranking			
	Self-first	Partner-first	Self-first	Partner-first		
	n = 38	*n* = 41	*n* = 53	*n* = 43	*n* = 93	*N* = 268
Game type						
Assurance	5 (13.2%)	21 (51.2%)	25 (47.2%)	15 (34.9%)	21 (22.6%)	87 (32.5%)
Chicken	12 (31.6%)	1 (2.4%)	-	3 (7.0%)	8 (8.6%)	24 (9.0%)
MD	1 (2.6%)	-	4 (7.5%)	4 (9.3%)	6 (6.4%)	15 (5.6%)
PD	2 (5.3%)	2 (4.9%)	-	1 (2.3%)	5 (5.4%)	10 (3.7%)
Symmetric	17 (44.7%)	20 (48.8%)	13 (24.5%)	15 (34.9%)	47 (50.5%)	112 (41.8%)
Game harmony	.13 (.70)	.33 (.69)	.51 (.58)	.43 (.70)	-.05 (.68)[a]	.22 (.70)

Note. Categories are reported as *n* (percentage), game harmony is reported as *M* (*SD*). For game types, only the four games expected in the perceptual convergence hypothesis are listed. Counts refer to the games matching the strategic structure considering dominant strategies and equilibria.

[a] In four cases, game harmony could not be computed, therefore this descriptive is based on *n* = 89 valid observations of the CG.

Deviating from the preregistration, not only the payoff hierarchy but also the resulting strategic structure of dominant strategies and equilibria was used to classify the game types. This allows for games with two or more equal payoffs in one of the perspectives to still be classified as a PD, MD, Chicken or Assurance game, if the structure of dominant strategies and equilibria matches (detailed information on how the games were classified can be found in the appendix). This game classification method revealed 136 of the 268 game structures matching one of the four expected game types. Another binomial test showed that this proportion was significantly higher than the changed probabilistic base rate of 4/78 that follows from considering strategically unique games, $p < .001$, with an observed proportion of 50.7%, 95% CI [45.5%, 100%], supporting the hypothesis as well.

Hypotheses 2 & 3 – Effects of Operationalization on Symmetry

As stated above, the self-developed asymmetry index was discarded for not meeting basic criteria for meaningful interpretation, and only the dichotomous symmetry coding was used to analyze effects on symmetry, deviating from the preregistration.

A chi-square test of independence showed that the pooled experimental groups (VS, VP, PS, PP) and the CG differed significantly in symmetry of the created game structures, $\chi^2(1) = 4.48$, $p = .034$, $\varphi = -.13$. In the CG, 50.5% of participants created a symmetric game structure, compared with 37.1% within the pooled experimental groups (see Table 2). This indicates a relationship of small effect size between separating perspectives in the outcome evaluation and less symmetry of the game structures created, supporting the second hypothesis.

A binary logistic regression was conducted to examine the effects of perspective sequence and abstraction level on symmetry of the game structures. The model was statistically significant $\chi^2(2) = 6.88$, $p = .032$, with the Nagelkerke $R^2 = .05$ indicating that the model explained a small proportion of the variance in symmetry. The odds of creating a symmetric game structure were significantly higher in the payoff matrix abstraction level compared with the verbal ranking condition, *OR* = 2.10, 95% CI [1.13, 3.95], $p = .020$, supporting the third hypothesis. Perspective sequence was not a significant predictor of symmetric game structures, *OR* = 1.39, 95% CI [0.75, 2.62], $p = .300$, indicating no clear difference in symmetry odds between evaluating the own or the partner's perspective first.

Hypothesis 4 – Effects of Operationalization on Competitiveness

To test the hypothesis that evaluating payoffs from the own perspective first will lead to the creation of more competitive game structures compared to evaluating the partner's perspective first, a 2 (perspective sequence) x 2 (abstraction level) between-participants analysis of variance (ANOVA) of game harmony was performed.

The ANOVA indicated no significant main effect of perspective sequence, $F(1, 171) = 0.38$, $p = .540$, $\eta_p^2 = .002$, a significant main effect of abstraction level, $F(1, 171) = 5.55$, $p = .020$, $\eta_p^2 = .031$, and no significant interaction, $F(1, 171) = 1.91$, $p = .168$, $\eta_p^2 = .011$. As the Q-Q-Plot and Shapiro-Wilk-Test ($W = 0.85$, $p < .001$) of residuals indicated significant non-normality, a robust two-way ANOVA on 20% trimmed means and adjusted critical values was performed using R package WRS2 version 1.1-7 (Mair & Wilcox, 2020). The robust ANOVA also indicated no significant main effect of perspective sequence (*Value* = 0.69, $p = .411$), a significant main effect of abstraction level (*Value* = 5.66, $p = .020$), and no significant interaction (*Value* = 1.18, $p = .281$). Combining these results with the group-level descriptive statistics (see Table 2) of game harmony, participants filling out payoff matrices created significantly more competitive and less harmonic game structures compared to participants ranking verbal outcome descriptions, but it does not make a significant difference from which perspective payoffs are evaluated first. Therefore, hypothesis four is not supported.

Hypotheses 5 & 6 – Effects of Individual Differences on Competitiveness

Concerning the effects of individual differences on the competitiveness of the created game structures, it was hypothesized that participants with higher BZSG create more competitive structures than participants with lower BZSG, while more prosocial participants create less competitive structures compared to less prosocial participants. To test the hypotheses, a multiple linear regression was conducted to predict game harmony from the participants' BZSG score and SVO angle.

The scale showed a good reliability with Cronbach's $\alpha = .83$ and McDonalds's $\omega = .89$, but the confirmatory factor analysis (CFA) indicated poor model fit (CFI = .741, TLI = .638, RMSEA = .204, SRMR = .103), suggesting that the scale did not adequately represent the data. Therefore, analyses involving the BZSG scale should be interpreted with caution.

The linear regression model, $F(2, 261) = 2.05$, $p = .131$, $R^2 = .015$, adjusted $R^2 = .008$, showed that SVO angle significantly predicted game harmony with a small effect, while BZSG score was nonsignificant (see Table 3 and Figure 3, for descriptive statistics of SVO angle and BZSG score see Table 1), partially supporting the hypotheses.

Table 3

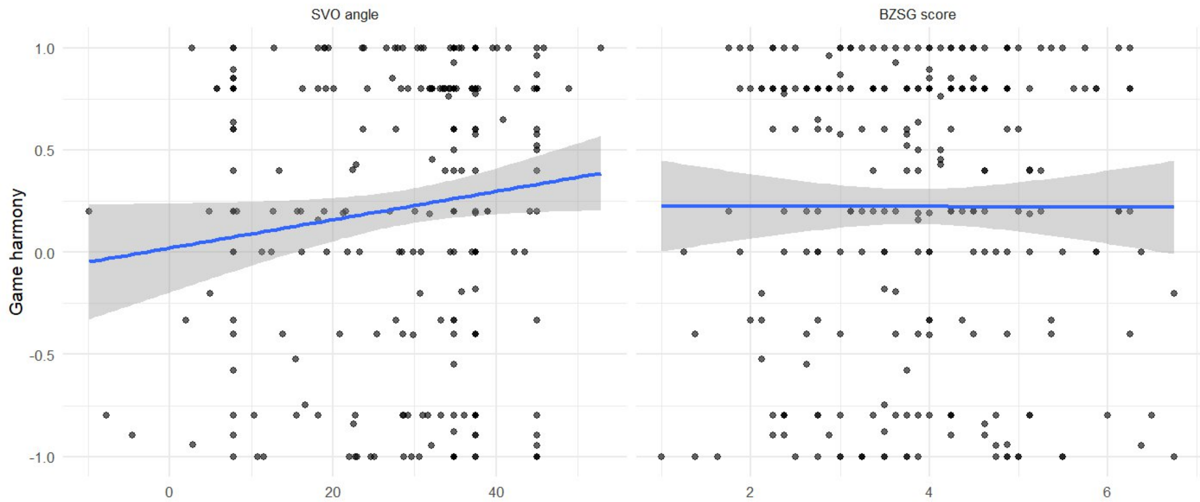
Regression Coefficients of Game Harmony Predicted by BZSG Score and SVO Angle

Effect	Estimate	SE	95% CI		*p*
			LL	*UL*	
Intercept	-0.04	0.20	-0.43	0.35	.845
BZSG score	0.01	0.04	-0.06	0.09	.714
SVO angle	0.01	0.004	0.00	0.01	.044

Note. 264 valid observations of game harmony, CI = confidence interval; *LL* = lower limit; *UL* = upper limit.

Figure 3

Relationship of Game Harmony With SVO Angle and BZSG Score



Note. Scatterplot of 264 valid observations of game.harmony, linear regression line in blue with standard error band in grey.

As a robustness check, the 25 participants with intransitive responses to the SVO items were excluded, and a linear regression was conducted to predict game harmony from the remaining participants' SVO angle only. In this model, $F(1, 238) = 2.05$, $p = .084$, $R^2 = .012$, adjusted $R^2 = .008$, SVO angle was not a significant predictor of game harmony, $b = 0.006$, $SE = 0.004$, 95% CI [-0.001, 0.013], $p = .084$, limiting support for hypothesis six.

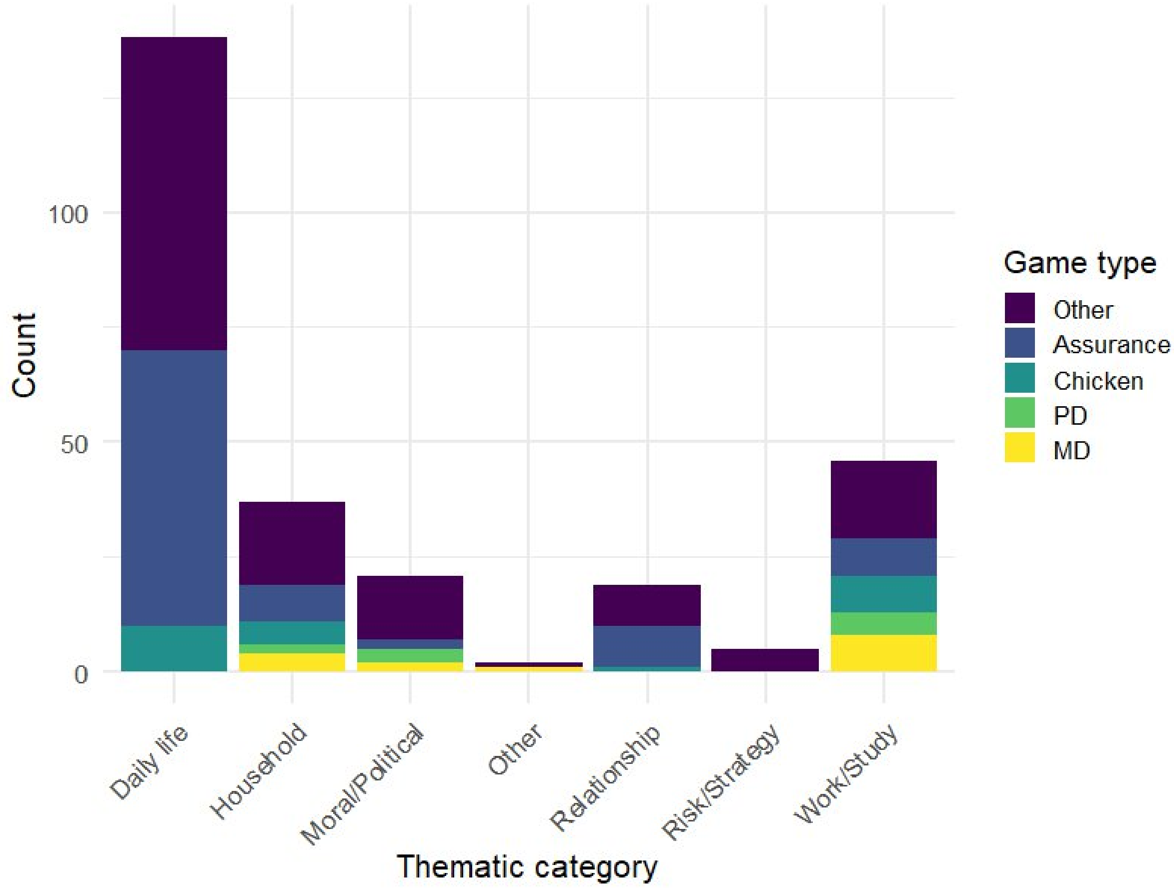
Exploratory Analyses

Thematic Categories

For exploratory analyses, participants' conflict scenarios were clustered into six thematic categories: daily life (51.5%, $n = 138$), work/study (17.2%, $n = 46$), household (13.8%, $n = 37$), moral/political (7.8%, $n = 21$), relationships (7.1%, $n = 19$), risk/strategy (1.9%, $n = 5$). Two of the scenarios did not fit in any of the six categories and were named as 'other' (0.7%, $n = 2$).

Figure 4

Distribution of Game Types Across Thematic Categories



Note. All game structures participants created that did not match one of the four expected games of the perceptual convergence hypothesis are combined under "Other".

Due to sparse cell counts and resulting violations of chi-square assumptions, the relation between thematic categories and game types was analyzed descriptively, indicating that the domain of the conflict scenario has an influence on the game structures created (see Figure 4). Inferential analyses focused on continuous structural properties of the games (e.g., game harmony), which allow for more robust statistical testing. A one-way between-subjects ANOVA revealed significant differences in game harmony across thematic categories, $F(6, 257) = 3.24$, $p = .004$, $\eta_p^2 = .07$. A contrast comparing the relationship scenarios, descriptively showing the highest average game harmony (see Table 4), against all other categories was significant, $t(257) = 2.76$, $p = .006$, indicating higher game harmony in relationship conflicts compared to other domains (contrast estimate = 0.52, $SE = 0.19$).

Table 4

Descriptive Statistics of Game Harmony by Thematic Category

Variable	Daily life	Household	Moral/ Political	Other	Relationship	Risk/ Strategy	Work/ Study
n	138	37	21	2	19	5	46
Game harmony	0.28 (0.70)	0.06 (0.65)	-0.22 (0.67)	0.40 (0.57)	0.63 (0.53)	-0.12 (0.87)	0.24 (0.70)

Note. Game harmony is reported as *M* (*SD*) for all subgroups.

As the Q-Q-Plot and Shapiro-Wilk-Test ($W = 0.91$, $p < .001$) of residuals indicated significant non-normality, a robust one-way ANOVA on 20% trimmed means was performed, $F(6, 9.64) = 7.83$, $p = .003$, Effect size = .57, still indicating significant differences in game harmony between thematic categories.

Equilibria

To examine potential effects of the manipulations of abstraction level and perspective sequence on the created game structures' equilibria, exploratory analyses with the game indices price of anarchy and price of stability were performed. Due to the strongly skewed and bounded distribution of the price of anarchy and price of stability measures, and the failure of transformations to approximate normality, group differences were analyzed using robust ANOVAs on 20% trimmed means and adjusted critical values. For price of anarchy, the robust ANOVA showed a significant main effect of perspective sequence (*Value* = 8.19, $p = .006$), no significant main effect of abstraction level (*Value* = 0.39, $p = .537$), and no significant interaction (*Value* = 2.29, $p = .134$). For price of stability, the robust ANOVA showed no significant main effect of perspective sequence (*Value* = 0.14, $p = .715$), no significant main effect of abstraction level (*Value* = 0.14, $p = .715$), and no significant interaction (*Value* = 0.14, $p = .715$). This indicates that there is no difference in the inefficiency between the best outcome and the best stable outcome between groups, but a significant difference in the inefficiency between the best outcome and the worst stable outcome depending on the perspective sequence. Full descriptives of the two indices by group can be found in Table 5.

Table 5

Price of Anarchy and Price of Stability by Group

Index	CG $n = 93$	PS $n = 38$	PP $n = 41$	VS $n = 53$	VP $n = 43$	Total $N = 268$
Price of anarchy	1.20 (0.50)	1.09 (0.17)	1.25 (0.36)	1.11 (0.17)	1.17 (0.23)	1.17 (0.35)
Price of stability	1.05 (0.13)	1.05 (0.14)	1.07 (0.17)	1.03 (0.10)	1.04 (0.13)	1.05 (0.13)

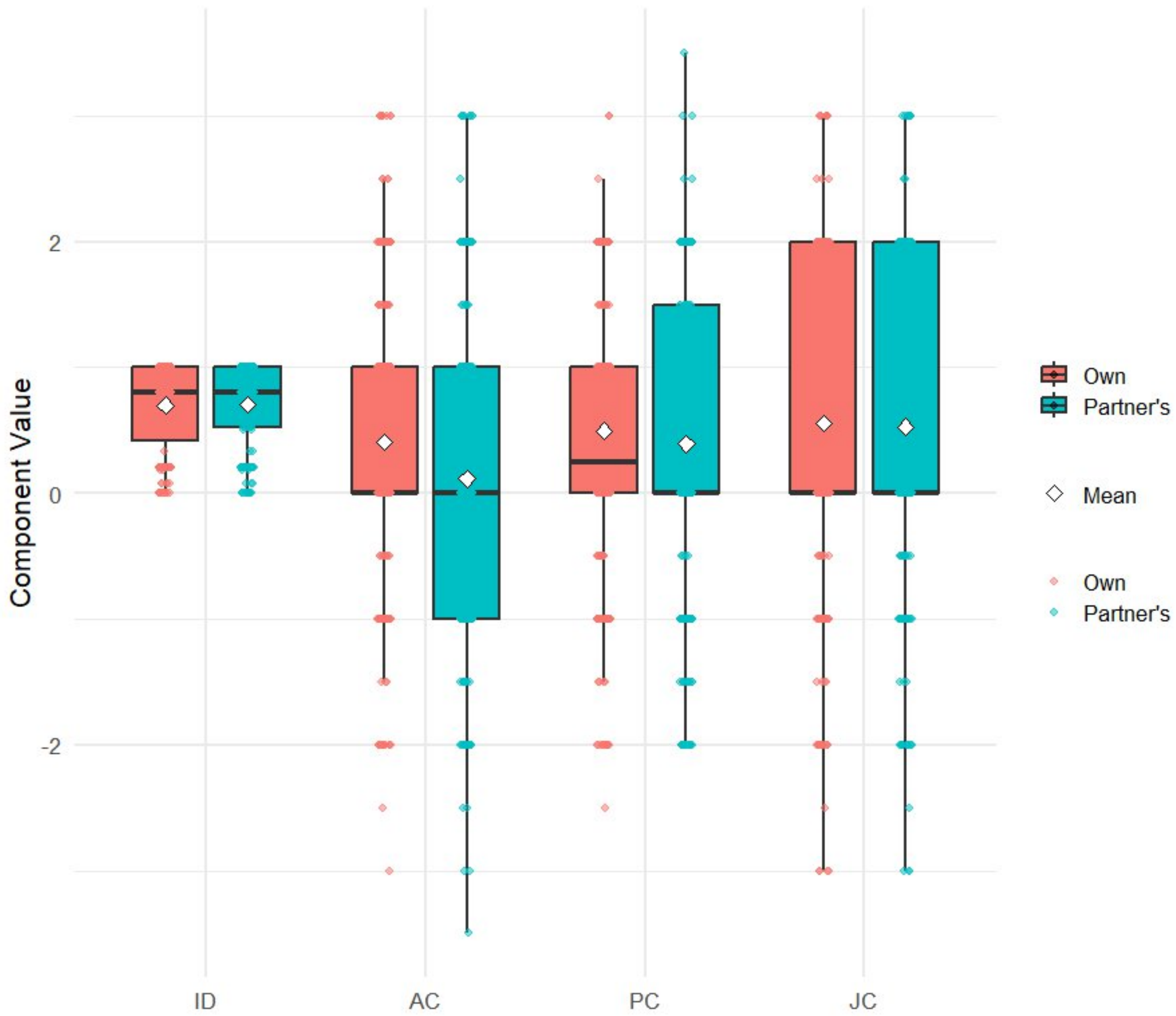
Note. All values are reported as *M* (*SD*).

Components of the Payoff Structures

To disentangle the payoff structures into their components, actor control, partner control, joint control, and interdependence were analysed. As all of them are computed using the both players' individual sets of four payoffs each separately, paired samples *t*-tests were performed to compare the payoff structure components between the own perspective and the partner's perspective.

Figure 5

Payoff Structure Components by Perspective



Note. Boxplots show payoff structure components by perspective (own vs. partner's). Boxes represent interquartile ranges with medians; whiskers extend to 1.5**IQR*. Points indicate individual observations and diamonds mark group means. ID = Interdependence, AC = Actor Control, PC = Partner Control, JC = Joint Control.

The tests showed a significant difference in the actor control of themselves and their partners in the game structures participants created, $t(267) = 3.27$, $p = .001$, mean difference = 0.29, indicating that participants created structures in which they were more in control of their own outcomes than their partners were in control of their outcomes, respectively. For interdependence, partner control, and joint control, there were no significant differences between the participants' own perspective and the partners' perspective (in all cases $p > .164$). A visualization of individual observations and descriptives can be found in Figure 5.

Discussion

Summary of Main Findings

The present study examined how individuals mentally represented self-generated conflict situations constructing 2x2 economic games. Consistent with prior work on conflict templates (Halevy et al., 2012), participants disproportionately often generated the four expected game structures of PD, MD, Chicken, and Assurance despite the vast theoretical space of possible games, supporting the perceptual convergence hypothesis. Moreover, the operationalization of the task significantly shaped the game structures participants created: separating perspectives in the outcome evaluation reduced symmetry, numeric abstractions (filling payoff matrices) increased both symmetry and competitiveness compared to less abstract ranking of verbal descriptions. In contrast, individual differences played a comparatively smaller role, with prosociality predicting more harmonious, less competitive structures, while BZSG showed no significant association.

The findings suggest that people do not construct their representations of conflict at random. Instead, they show partial convergence on a limited set of prototypical game structures, whereas the specific form of these representations remains sensitive to how the task is operationalized. Individual differences contributed as well, but less consistently than expected.

Implications for Conflict Templates and Interdependence Theory

The results support the idea that interdependence is mentally represented through simplified templates, but also that task structure shapes these templates. In that, mental representations are not fixed personality expressions but situationally constructed.

The pattern that the majority of people map conflicts on four archetypal games (Halevy et al., 2012) is replicated in this study, even with self-generated scenarios, strengthening ecological validity. This is consistent with the basic perceptual convergence hypothesis suggesting that mental representations are not arbitrary and that people do not represent conflict with unlimited diversity. However, the proportion of the four prototypical perceptual convergence hypothesis games found in this study is much lower than in Halevy et al. (2012). This might be due to a less constraining design of task instruction and measurement of the conflict templates in the present study: while in one of their studies they did not specify a particular interaction as well, they still asked participants to think specifically about a negotiation (Halevy et al., 2012). In the present study, however, participants were asked to imagine a conflict, more generally speaking. While this seems to only be a small difference, it might still have a considerable effect on people as they are especially receptible for little cues in complex situations like the conflict representation task (Cortes et al., 2023; Payne et al., 1992; Shah & Oppenheimer, 2008).

Adding to this, participants in prior work only filled in empty payoff matrices while a significant part of the present study was to alter this task format to examine potential effects of operationalization. The same broad task yielded different structural properties depending on whether the players' outcomes were evaluated jointly or separately, and whether the format was numeric or verbal. These significant operationalization effects suggest that conflict templates, although prominently reduced to a small set of fixed structures, are still actively constructed, cue-sensitive and therefore method-dependent. This is strengthening the link to interdependence theory as it confirms that people do not merely read off the objective structure of interdependence in a given situation but actively transform it into an effective matrix (Kelley & Thibaut, 1978; Thielmann et al., 2021).

Together, this means that conflict templates may be best understood as constrained mental representations. They are neither fixed reflections of objective structure nor fully individual inventions.

Role of Operationalization vs. Individual Differences

Effects of Abstraction and Symmetry Cues

Participants who filled in numeric payoff matrices constructed more symmetric and more competitive game structures than people who ranked verbal outcome descriptions. Numeric matrices may put the formal strategic structure of games in focus, encouraging participants to think in more abstract, game-theoretic ways. The matrix format making tradeoffs, opposition, and strategic comparability especially salient could explain both higher symmetry and greater competitiveness. In contrast to this, the verbal format may preserve more of the semantic and contextual richness of a scenario, potentially reducing the salience of pure structural competition and allowing participants to represent situations in more relational or less formally opposed terms. Taking together, participants may attend more strongly to relative payoffs, thereby amplifying competitive interpretations of interdependence, when they operate on matrices rather than narratives. Lastly, the symmetry result also supports the theoretical argument that symmetry preferences may partly reflect not only a preference for symmetric interdependence but also a response to visual features of the format. The data also suggest that decomposing the matrix by perspective reduces symmetry, which also reduces the availability of visual cues for symmetry. At the same time, asking participants to evaluate each player separately might also make differences between players more cognitively available. In other words, asymmetry may often be overlooked (a part of the perceptual convergence hypothesis that was criticized, see Balliet et al., 2017) not because people do not perceive it but because some formats make representation of it easier than others.

Effects of Perspective Sequence

Contrary to expectations derived from perspective taking research that showed positive correlations between cooperative behavior and perspective taking in multiple domains (Casler & Groves, 2023; Erle & Topolinski, 2017; Li et al., 2019; Ortiz-Riomalo et al., 2021), the order in which participants evaluated perspectives did not significantly influence competitiveness of the conflict templates participants constructed. Contrastingly, the data suggest that merely evaluating the partner first is not sufficient to shift the overall construal of conflict in a consistent direction. This may indicate that sequence alone as a manipulation was too subtle compared with the more global effects of the task format. Furthermore, the scenario imagination task may have fixed the conflict frame already before the perspective manipulation occurred, and both sequence conditions still required evaluation of both perspectives limiting the possibility for divergence. Taken together, perspective taking may alter judgment only under stronger affective or motivational induction, and not merely through order of evaluation.

Individual Differences

Despite SVO being the strongest predictor for prosocial behavior in social dilemmas (Thielmann et al., 2020) and correlating negatively with the perception of conflict (Gerpott et al., 2018), it only showed a small and fragile association with the constructed game structures being more harmonious. This direction of effect fits interdependence theory suggesting that people with stronger prosocial preferences may transform the same given situation into a more cooperative effective matrix (Kelley & Thibaut, 1978). However, as the significance of the association did not withhold the robustness check of excluding intransitive SVO slider response sets, conclusions on the role of SVO in the construction of conflict templates should be made with caution. This pattern is in line with the finding of Halevy et al. (2012) who did not find any relation between SVO and endorsements of the four prototypical game types.

While the absence of significant BZSG effects in this study contrasts with prior theoretical expectations (Andrews Fearon & Götz, 2024; Davidai et al., 2022; Davidai & Tepper, 2023), the interpretation of the null finding is limited as the BZSG scale showed a poor fit of the measurement model in the confirmatory factor analysis with the present sample. It therefore remains unclear whether the absence of significant effects reflects insufficient construct validity or a true lack of association with conflict template construction.

Different Areas of Conflict

The findings of the present study show that conflict templates are not only person-driven or method-driven but also sensitive to content. Considering the differences in game harmony dependent on the thematic category or area of conflict that were found, the results of this study are in line with prior research showing that relationship interactions are mostly characterized by moderate mutual dependence and corresponding interests, resulting in low conflict (Columbus et al., 2021). This is also reflected in the game types participants created depending on the topic of their scenario (Figure 4): while participants who gave conflict scenarios from daily life or relationship contexts mostly created structures resembling Assurance or games that do not fit any of the four prototypical structures, participants who evaluated scenarios in the work and study context created game structures matching the generally more competitive PD and Chicken games (Columbus et al., 2021; Halevy & Phillips, 2015; Van Lange, 2014). Even though participants freely generated own scenarios, the domain of the interaction appeared to matter for the resulting game structure. This suggests that conflict templates, while to some extent shaped by general cognitive tendencies and dispositional priors, also reflect domain-specific expectations about how interests align in different areas of life.

Properties of Strategy and Interdependence

Interestingly, while Halevy et al. (2014) found more variation in best strategy beliefs than in worst strategy beliefs, this study did not find differences in the price of stability (which is based on the best equilibria) but only in price of anarchy (based on the worst equilibria) of the game structures, associated with the variation of perspective sequence. This suggests that the perspective sequence manipulation may have influenced only specific aspects of the conflict template generation, namely how participants represented the worst stable consequences of the interaction, but not the best stable outcome. In other words, the manipulation may have affected the downside structure of the games more than their positive aspects.

Looking at the results of the perspective comparison of interdependence components (Figure 5), participants represented themselves as having more control over their own outcomes than their partners over their own ones. However, they do not represent themselves as having more control over their partners' outcomes than their partners have over theirs. That may reflect a slight egocentric bias in construing interdependence, meaning that participants may anchor more strongly on their own action-outcome contingencies, even when they are explicitly representing both players separately. This complements the null result in the expected effects of perspective sequence on the competitiveness of the games structures: while partner-first evaluation of outcomes did not differ from self-first evaluation in competitiveness of the overall structures, the payoff component analysis still suggests some general asymmetry in the representation of the self and the interaction partners.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, while improving ecological breadth of the findings, the self-generated scenarios reduce standardization. Differences between participants may reflect differences in the scenarios they chose to imagine to a significant proportion. The introduction of three examples in the task instruction might also have introduced bias, as many participants chose one of those scenarios. Second, there were some limitations in the measurements that also led to substantial deviations from the preregistration: many participants violated instructions in giving responses with no strict ordinality in the payoff matrix condition, which suggests some indifference between the outcomes. However, this potential indifference could not be expressed in the verbal description condition. In this condition however, some participants gave duplicate descriptions for outcomes, making it impossible for themselves to distinguish between them when ranking the descriptions. As this distinction was made impossible partly by the design of the study (no indication of the strategy combination of an outcome when ranking the descriptions), it remains unclear if the duplicate descriptions reflect indifference between outcomes and could therefore not be considered meaningfully in the analysis. Furthermore, due to the payoff structures and the broad spectrum of games participants created, some of the game indices could not be used as intended. Overall, although efforts were made to preserve the greatest possible comparability between the planned and final analyses, this comparability remains limited. Third, as noted above, the BZSG measurement showed a poor fit in the confirmatory factor analysis, limiting the interpretability of the findings on the hypothesized relation between BZSG and the competitiveness of the game structures. Fourth, several findings relied on small effects or exploratory analyses with sparse cells, calling for replication. Lastly, the sample came largely from convenience recruitment and was mostly young, well educated, and had substantial prior knowledge in game theory, which may matter for how easily participants think in strategic structures.

Future Directions

Future research should build on the present findings on conflict templates by addressing the limitations of the current study and further examining the observed relationships with task format and individual differences. First of all, standardized scenarios could be used alongside the self-generated scenarios. Further, participants could be instructed to imagine multiple scenarios of different domains to separate variability in the imagined conflict domain from variability in mental representation. Second, as some of the results do indicate perspective differences but the perspective sequence manipulation did not show any significant effects, stronger manipulations of perspective taking should be tested. For example, not only could the order of evaluation be manipulated, but participants could also be given explicit partner-role instructions, an empathy induction could be used, or participants could be instructed to only evaluate one perspective which they are randomly assigned to. Third, the measurement of BZSG should be refined or more domain-specific BZSG should be tested to see whether they predict construction of competitive conflict templates more reliably. Lastly, the prediction of actual downstream behavior by the constructed conflict templates could be tested by integrating a subsequent incentivized game, connecting representation directly to choice of a strategy.

Conclusion

The present findings suggest that conflict templates are best understood as constrained but malleable mental representations. Although participants frequently converged on a limited set of prototypical game structures, these representations also varied as a function of task characteristics and, to a lesser extent, individual differences. This pattern underscores the importance of subjective construal in social dilemmas and supports the view that behavior is guided by an effective rather than purely given structure of interdependence. Accordingly, research using economic games may benefit from greater attention to how individuals mentally represent the conflict situations they face.

References

Andrews Fearon, P., & Götz, F. M. (2024). The zero-sum mindset. *Journal of Personality and Social Psychology*, *127*(4), 758–795. https://doi.org/10.1037/pspa0000404

Au, W. T., Lu, S., Leung, H., Yam, P., & Fung, J. M. Y. (2012). Risk and Prisoner's Dilemma: A Reinterpretation of Coombs' Re-parameterization. *Journal of Behavioral Decision Making*, *25*(5), 476–490. https://doi.org/10.1002/bdm.743

Aumann, R. J. (2018). Game Theory. In Macmillan Publishers Ltd (Ed.), *The New Palgrave Dictionary of Economics* (pp. 5014–5053). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95189-5_942

Axelrod, R. (1967). Conflict of interest: An axiomatic approach. *Journal of Conflict Resolution*, *11*(1), 87–99. https://doi.org/10.1177/002200276701100107

Balliet, D., & Lindström, B. (2023). Inferences about interdependence shape cooperation. *Trends in Cognitive Sciences*, *27*(6), 583–595. https://doi.org/10.1016/j.tics.2023.03.003

Balliet, D., Parks, C., & Joireman, J. (2009). Social Value Orientation and Cooperation in Social Dilemmas: A Meta-Analysis. *Group Processes & Intergroup Relations*, *12*(4), 533–547. https://doi.org/10.1177/1368430209105040

Balliet, D., Tybur, J. M., & Van Lange, P. A. M. (2017). Functional Interdependence Theory: An Evolutionary Account of Social Situations. *Personality and Social Psychology Review*, *21*(4), 361–388. https://doi.org/10.1177/1088868316657965

Bogaert, S., Boone, C., & Declerck, C. (2008). Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British Journal of Social Psychology*, *47*(3), 453–480. https://doi.org/10.1348/014466607X244970

Camerer, C. F. (2003a). *Behavioral game theory: Experiments in strategic interaction*. Russell sage foundation Princeton university press.

Camerer, C. F. (2003b). Behavioural studies of strategic thinking in games. *Trends in Cognitive Sciences*, *7*(5), 225–231. https://doi.org/10.1016/S1364-6613(03)00094-9

Camerer, C. F. (2018). Behavioural Game Theory. In Macmillan Publishers Ltd (Ed.), *The New Palgrave Dictionary of Economics* (pp. 867–874). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95189-5_2384

Capraro, V. (2013). A Model of Human Cooperation in Social Dilemmas. *PLoS ONE*, *8*(8), e72427. https://doi.org/10.1371/journal.pone.0072427

Cargile, A. C., Bradac, J. J., & Cole, T. (2006). Theories of Intergroup Conflict: A Report of Lay Attributions. *Journal of Language and Social Psychology*, *25*(1), 47–63. https://doi.org/10.1177/0261927X05284479

Casler, D., & Groves, D. (2023). Perspective Taking through Partisan Eyes: Cross-National Empathy, Partisanship, and Attitudes toward International Cooperation. *The Journal of Politics*, *85*(4), 1471–1486. https://doi.org/10.1086/723994

Chang, L.-A., Armaos, K., Warns, L., Ma de Sousa, A. Q., Paauwe, F., Scholz, C., & Engelmann, J. B. (2023). Mentalizing in an economic games context is associated with enhanced activation and connectivity in the left temporoparietal junction. *Social Cognitive and Affective Neuroscience*, *18*(1), nsad023. https://doi.org/10.1093/scan/nsad023

Columbus, S., & Molho, C. (2022). Subjective interdependence and prosocial behaviour. *Current Opinion in Psychology*, *43*, 226–231. https://doi.org/10.1016/j.copsyc.2021.07.022

Columbus, S., Molho, C., Righetti, F., & Balliet, D. (2021). Interdependence and cooperation in daily life. *Journal of Personality and Social Psychology*, *120*(3), 626–650. (2020-44813-001). https://doi.org/10.1037/pspi0000253

Cortes, P. M., García-Hernández, J. P., Iribe-Burgos, F. A., Guevara, M. A., & Hernández-González, M. (2023). Effects of emotional congruency and task complexity on decision-making. *Cognitive Processing*, *24*(2), 161–171. https://doi.org/10.1007/s10339-023-01129-1

Davidai, S., & Tepper, S. J. (2023). The psychology of zero-sum beliefs. *Nature Reviews Psychology*, *2*(8), 472–482. https://doi.org/10.1038/s44159-023-00194-9

Davidai, S., White, M. W., & Gregorich, G. (2022). The fear of conflict leads people to systematically avoid potentially valuable zero-sum situations. *Scientific Reports*, *12*(1), 17944. https://doi.org/10.1038/s41598-022-22849-y

Dawes, R. M. (1980). Social Dilemmas. *Annual Review of Psychology*, *31*(1), 169–193. https://doi.org/10.1146/annurev.ps.31.020180.001125

Deutsch, M. (1949). A Theory of Co-operation and Competition. *Human Relations*, *2*(2), 129–152. https://doi.org/10.1177/001872674900200204

Devetag, G., & Warglien, M. (2008). Playing the wrong game: An experimental analysis of relational complexity and strategic misrepresentation. *Games and Economic Behavior*, *62*(2), 364–382. https://doi.org/10.1016/j.geb.2007.05.007

Epley, N., & Caruso, E. M. (2009). Perspective Taking: Misstepping Into Others' Shoes. In *Handbook of Imagination and Mental Simulation* (1st ed., pp. 295–309). Psychology Press.

Epley, N., Caruso, E. M., & Bazerman, M. H. (2005). When Perspective Taking Increases Taking: Reactive Egoism in Social Interaction. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.785989

Erle, T. M., & Topolinski, S. (2017). The grounded nature of psychological perspective-taking. *Journal of Personality and Social Psychology*, *112*(5), 683–695. https://doi.org/10.1037/pspa0000081

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fehr, E., & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review*, *90*(4), 980–994. https://doi.org/10.1257/aer.90.4.980

Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140. https://doi.org/10.1038/415137a

Freyd, J., & Tversky, B. (1984). Force of Symmetry in Form Perception. *The American Journal of Psychology*, *97*(1), 109. https://doi.org/10.2307/1422551

Fudenberg, D., & Tirole, J. (2010). *Game theory* (Reprinted). Ane Books.

Fujii, S., Gärling, T., & Kitamura, R. (2001). Changes in Drivers' Perceptions and Use of Public Transport during a Freeway Closure: Effects of Temporary Structural Change on Cooperation in a Real-Life Social Dilemma. *Environment and Behavior*, *33*(6), 796–808. https://doi.org/10.1177/00139160121973241

Gächter, S., Lee, K., Sefton, M., & Weber, T. O. (2024). The role of payoff parameters for cooperation in the one-shot Prisoner's Dilemma. *European Economic Review*, *166*, 104753. https://doi.org/10.1016/j.euroecorev.2024.104753

Gerpott, F. H., Balliet, D., Columbus, S., Molho, C., & de Vries, R. E. (2018). How do people think about interdependence? A multidimensional model of subjective outcome interdependence. *Journal of Personality and Social Psychology*, *115*(4), 716–742. (2017-38605-001). https://doi.org/10.1037/pspp0000166

Halevy, N., & Chou, E. Y. (2014). How decisions happen: Focal points and blind spots in interdependent decision making. *Journal of Personality and Social Psychology*, *106*(3), 398–417. (2013-45448-001). https://doi.org/10.1037/a0035351

Halevy, N., Chou, E. Y., & Murnighan, J. K. (2012). Mind games: The mental representation of conflict. *Journal of Personality and Social Psychology*, *102*(1), 132–148. https://doi.org/10.1037/a0025389

Halevy, N., Cohen, T. R., Chou, E. Y., Katz, J. J., & Panter, A. T. (2014). Mental Models at Work: Cognitive Causes and Consequences of Conflict in Organizations. *Personality* *and Social Psychology Bulletin*, *40*(1), 92–110. https://doi.org/10.1177/0146167213506468

Halevy, N., & Katz, J. J. (2013). Conflict Templates: Thinking Through Interdependence. *Current Directions in Psychological Science*, *22*(3), 217–224. https://doi.org/10.1177/0963721412474296

Halevy, N., & Phillips, L. T. (2015). Conflict Templates in Negotiations, Disputes, Joint Decisions, and Tournaments. *Social Psychological and Personality Science*, *6*(1), 13–22. https://doi.org/10.1177/1948550614542347

Harring, N., & Krockow, E. M. (2021). The social dilemmas of climate change and antibiotic resistance: An analytic comparison and discussion of policy implications. *Humanities and Social Sciences Communications*, *8*(1), 125. https://doi.org/10.1057/s41599-021-00800-2

Humphrey, D. (1997). Preferences in Symmetries and Symmetries in Drawings: Asymmetries between Ages and Sexes. *Empirical Studies of the Arts*, *15*(1), 41–60. https://doi.org/10.2190/DF5N-HGFB-MVPK-U34D

Kelley, H. H., Holmes, J. G., Kerr, N. L., Reis, H. T., Rusbult, C. E., & van Lange, P. A. M. (2009). *An Atlas of Interpersonal Situations*. Cambridge University Press. https://doi.org/10.1017/CBO9780511499845

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations* (J. W. Thibaut, Trans.). Wiley.

Kilgour, D. M., & Fraser, N. M. (1988). A taxonomy of all ordinal 2 x 2 games. *Theory and Decision*, *24*(2), 99–117. https://doi.org/10.1007/BF00132457

Kollock, P. (1998). Social Dilemmas: The Anatomy of Cooperation. *Annual Review of Sociology*, *24*(1), 183–214. https://doi.org/10.1146/annurev.soc.24.1.183

Komorita, S. S. (2019). *Social dilemmas*. Routledge.

Korn, L., Böhm, R., Meier, N. W., & Betsch, C. (2020). Vaccination as a social contract. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(26), 14890–14899. https://doi.org/10.1073/pnas.1919666117

Kreps, D. M. (2018). Nash Equilibrium. In Macmillan Publishers Ltd (Ed.), *The New Palgrave Dictionary of Economics* (pp. 9251–9258). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-349-95189-5_963

Levy, S. R., Chiu, C., & Hong, Y. (2006). Lay Theories and Intergroup Relations. *Group Processes & Intergroup Relations*, *9*(1), 5–24. https://doi.org/10.1177/1368430206059855

Li, P., Jin, X., Liao, Y., Li, Y., Shen, M., & He, J. (2019). Cooperation turns preschoolers into flexible perspective takers. *Cognitive Development*, *52*, 100823. https://doi.org/10.1016/j.cogdev.2019.100823

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*(2), 464–488. https://doi.org/10.3758/s13428-019-01246-w

Mak, V., & Rapoport, A. (2013). The price of anarchy in social dilemmas: Traditional research paradigms and new network applications. *Organizational Behavior and Human Decision Processes*, *120*(2), 142–153. https://doi.org/10.1016/j.obhdp.2012.06.006

Marwell, G., Ratcliff, K., & Schmitt, D. R. (1969). Minimizing differences in a maximizing difference game. *Journal of Personality and Social Psychology*, *12*(2), 158–163. https://doi.org/10.1037/h0027574

Murphy, R. O., Ackermann, K. A., & Handgraaf, M. J. J. (2011). Measuring Social Value Orientation. *Judgment and Decision Making*, *6*(8), 771–781. https://doi.org/10.1017/S1930297500004204

Neumann, J., & Morgenstern, O. (2007). *Theory of Games and Economic Behavior (60th Anniversary Commemorative Edition)*. Princeton University Press. https://doi.org/10.1515/9781400829460

Ng, G. T. T., & Au, W. T. (2016). Expectation and cooperation in prisoner's dilemmas: The moderating role of game riskiness. *Psychonomic Bulletin & Review*, *23*(2), 353–360. https://doi.org/10.3758/s13423-015-0911-7

Ng, G. T. T., Au, W. T., & Lai, D. C. K. (2024). The role of game riskiness on the expectation-cooperation link in social dilemmas and its relations with fear and greed. *Judgment and Decision Making*, *19*, e19. https://doi.org/10.1017/jdm.2024.21

Ortiz-Riomalo, J. F., Koessler, A.-K., & Engel, S. (2021). Inducing perspective-taking for prosocial behaviour in natural resource management. *Journal of Environmental Economics and Management*, *110*, 102513. https://doi.org/10.1016/j.jeem.2021.102513

Payne, J. W., Bettman, J. R., Coupey, E., & Johnson, E. J. (1992). A constructive process view of decision making: Multiple strategies in judgment and choice. *Acta Psychologica*, *80*(1–3), 107–141. https://doi.org/10.1016/0001-6918(92)90043-D

R Core Team. (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rapoport, A. (1967). A note on the "index of cooperation" for Prisoner's Dilemma. *Journal of Conflict Resolution*, *11*(1), 100–103. https://doi.org/10.1177/002200276701100108

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's Dilemma*. University of Michigan Press.

Rapoport, A., & Chammah, A. M. (1966). The Game of Chicken. *American Behavioral Scientist*, *10*(3), 10–28. https://doi.org/10.1177/000276426601000303

Reber, R. (2002). Reasons for the preference for symmetry. *Behavioral and Brain Sciences*, *25*(3), 415–416. https://doi.org/10.1017/S0140525X02350076

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing Fluency and Aesthetic Pleasure: Is Beauty in the Perceiver's Processing Experience? *Personality and Social Psychology Review*, *8*(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Revelle, W. (2026). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University. https://CRAN.R-project.org/package=psych

Rhodes, G., Proffitt, F., Grady, J. M., & Sumich, A. (1998). Facial symmetry and the perception of beauty. *Psychonomic Bulletin & Review*, *5*(4), 659–669. https://doi.org/10.3758/BF03208842

Robinson, D., & Goforth, D. (2006). *The topology of the 2x2 games: A new periodic table / David Robinson and David Goforth*. Routledge.

Rosseel, Y., Jorgensen, T. D., & Wilde, L. D. (2025). *lavaan: Latent Variable Analysis*. https://doi.org/10.32614/CRAN.package.lavaan

Różycka-Tran, J., Boski, P., & Wojciszke, B. (2015). Belief in a Zero-Sum Game as a Social Axiom: A 37-Nation Study. *Journal of Cross-Cultural Psychology*, *46*(4), 525–548. https://doi.org/10.1177/0022022115572226

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristics made easy: An effort-reduction framework. *Psychological Bulletin*, *134*(2), 207–222. https://doi.org/10.1037/0033-2909.134.2.207

Škerlavaj, M., Connelly, C. E., Cerne, M., & Dysvik, A. (2018). Tell me if you can: Time pressure, prosocial motivation, perspective taking, and knowledge hiding. *Journal of Knowledge Management*, *22*(7), 1489–1509. https://doi.org/10.1108/JKM-05-2017-0179

Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge University Press.

Stivers, A. W., Murphy, R. O., Spadaro, G., Balliet, D., & Kuhlman, D. M. (2025). A taxonomy of prisoner's dilemma indexes. *Decision*, *12*(2), 111–145. https://doi.org/10.1037/dec0000259

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

Thielmann, I., Spadaro, G., & Balliet, D. (2020). Personality and prosocial behavior: A theoretical framework and meta-analysis. *Psychological Bulletin*, *146*(1), 30–90. https://doi.org/10.1037/bul0000217

Van Lange, P. A. M. (2014). *Social Dilemmas: Understanding Human Cooperation*. OUP USA.

Van Lange, P. A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125–141. https://doi.org/10.1016/j.obhdp.2012.11.003

Van Lange, P. A. M., & Rand, D. G. (2022). Human Cooperation and the Crises of Climate Change, COVID-19, and Misinformation. *Annual Review of Psychology*, *73*(1), 379–402. https://doi.org/10.1146/annurev-psych-020821-110044

Villiger, D., Ullrich, J., & Krueger, J. (2022). *Probabilities of cooperative moves in all 67 symmetric ordinal two-player two-moves games*. https://doi.org/10.5167/uzh-229336

Zizzo, D. J., & Tan, J. H. W. (2007). Perceived harmony, similarity and cooperation in 2×2 games: An experimental study. *Journal of Economic Psychology*, *28*(3), 365–386. https://doi.org/10.1016/j.joep.2006.06.008

Zizzo, D. J., & Tan, J. H. W. (2011). Game Harmony: A Behavioral Approach to Predicting Cooperation in Games. *American Behavioral Scientist*, *55*(8), 987–1013. https://doi.org/10.1177/0002764211407905

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of variance
BZSG	Beliefs in a zero-sum game
CG	Control Group
CFA	Confirmatory factor analysis
MD	Maximizing Difference game
PD	Prisoner’s Dilemma game
PP	Payoff matrix Partner-first group
PS	Payoff matrix Self-first group
SVO	Social value orientation
VP	Verbal ranking Partner-first Group

VS Verbal ranking Self-first Group

Appendix

Abstract

People do not always respond to social dilemmas based on an objective payoff structure alone, but based on how they mentally represent the interdependence involved. Building on research on conflict templates, the present preregistered study examined both whether individuals converge on a small set of prototypical 2x2 game structures and how these representations are shaped by task characteristics and individual differences. A total of 268 participants generated a conflict scenario and constructed its payoff structure under varying conditions of abstraction level (payoff matrix vs. verbal ranking) and perspective sequence (self-first vs. partner-first), with an additional simultaneous payoff-matrix control group. Consistent with the perceptual convergence hypothesis, participants disproportionately often created structures corresponding to Prisoner's Dilemma, Chicken, Assurance, and Maximizing Difference games. At the same time, the way the task was operationalized influenced the resulting representations: separating perspectives reduced symmetry, the payoff matrix format produced more symmetric and more competitive game structures than verbal ranking. In contrast, perspective sequence did not affect competitiveness. Individual differences showed weaker effects: social value orientation was positively associated with game harmony, although this effect was small and less robust in sensitivity analyses, whereas beliefs in a zero-sum game were not significantly related to competitiveness. Overall, the findings suggest that conflict templates are neither arbitrary nor fixed, but constrained mental representations that are still sensitive to situational cues and how interdependence is elicited.

Keywords: social dilemmas; conflict templates; interdependence; economic games; social value orientation; zero sum beliefs

Zusammenfassung

Menschen reagieren auf soziale Dilemmata nicht nur auf der Grundlage einer objektiven Auszahlungsstruktur, sondern auch danach, wie sie die damit verbundene gegenseitige Abhängigkeit mental repräsentieren. Aufbauend auf früherer Forschung zu Konfliktmustern untersuchte diese präregistrierte Studie sowohl, ob sich die mentalen Repräsentationen der Teilnehmenden auf eine kleine Gruppe prototypischer 2x2-Spielstrukturen beschränken, als auch, wie diese Repräsentationen durch Aufgabenmerkmale und individuelle Unterschiede geprägt werden. Insgesamt 268 Teilnehmende erzeugten ein Konfliktszenario und konstruierten dessen Auszahlungsstruktur unter variierenden Bedingungen hinsichtlich des Abstraktionsgrades (Auszahlungsmatrix vs. verbale Rangfolge) und der Perspektivenreihenfolge (ich zuerst vs. Partner zuerst), mit einer zusätzlichen Kontrollgruppe, die beide Perspektiven gleichzeitig im Auszahlungsmatrix Format bearbeitete. Im Einklang mit der Hypothese der Wahrnehmungskonvergenz schufen die Teilnehmenden überproportional viele Strukturen, die den Spielen „Prisoner's Dilemma", „Chicken", „Assurance" und „Maximizing Difference" entsprachen. Gleichzeitig beeinflusste die Art und Weise, wie die Aufgabe operationalisiert wurde, die resultierenden Repräsentationen: Die Trennung der Perspektiven reduzierte die Symmetrie, das Format der Auszahlungsmatrix erzeugte symmetrischere und wettbewerbsintensivere Spielstrukturen als die verbale Rangfolge. Im Gegensatz dazu hatte die Reihenfolge der Perspektiven keinen Einfluss auf die Wettbewerbsintensität. Individuelle Unterschiede zeigten schwächere Effekte: Soziale Wertorientierung war positiv mit Spielharmonie assoziiert, wobei dieser Effekt gering und in Sensitivitätsanalysen nicht robust war, während der Glaube an ein Nullsummenspiel nicht signifikant mit der Wettbewerbsintensität zusammenhing. Insgesamt deuten die Ergebnisse darauf hin, dass Konfliktmuster weder willkürlich noch feststehend sind, sondern eingeschränkte mentale Repräsentationen, die dennoch sensitiv für situative Hinweise und die Art und Weise, wie Interdependenz hervorgerufen wird, sind.

Stichwörter: soziale Dilemmata; Konfliktmuster; Interdependenz; ökonomische Spiele; soziale Wertorientierung; Nullsummenglaube

Conflict Representation Task – Instruction Texts and Examples

In the conflict representation task, 2x2 games were briefly explained to participants using the following text:

> In this task, we ask you to imagine a situation where two people or groups (it could be you and another person, two friends, two colleagues, two countries, etc.) each face a choice between two possible actions. Those people or groups will be called players. Each player has exactly **two options**. The options can be anything you think of (for example: cooperate vs. not cooperate, work hard vs. take it easy, buy fair vs. cheap, help vs. ignore, etc.). The choices of both players together lead to an outcome that **affects them both**. Such a situation is often called a *2x2 scenario*, because there are two players and each has two possible choices, leading to four possible outcomes in total.

After this introduction text, participants were asked to generate an own example of a 2x2 situation and to describe it in own words. They were given three examples from different domains:

> Think of a concrete example of such a situation. It can come from everyday life (e.g., sharing chores, friendship, work, family), society (e.g., environment, politics, health), or even something completely hypothetical. Briefly describe it in your own words. Who are the two people (or groups)? What are the two options available to each? You don't need to use numbers, diagrams, or technical terms—just a short description in plain language.
>
> To clarify your task, here are three example descriptions: **Example 1 (daily life):** Two roommates decide whether or not to clean the kitchen. **Example 2 (society):** Two countries decide whether or not to cut carbon emissions. **Example 3 (coordination):** Two friends decide whether to meet at Café A or Café B.

After giving their scenarios, participants were asked to give short labels for the two options the players had in their scenario, as those labels were needed to create the according payoff matrices and outcome combinations. Participants in the payoff matrix condition were then given empty payoff matrices with their labels and were asked to fill in the payoffs. Participants in the verbal ranking condition should give brief descriptions of all four outcomes in another step. They were then presented with these descriptions in a random order and were asked to rank them according to their favorability.

Game Indices Computation and Game Classification

To evaluate the proportion of the four expected games in the game structures participants created, the games were classified in two ways: first, all structures that resembled the strictly ordinal payoff ranking of one of the four games (Figure 2) were classified accordingly. All others were summarized in the category "other". Second, it was also considered that some participants did not follow the instructions and created game structures with no strictly ordinal payoff ranking. However, these games might still have the strategic characteristics of one of the prototypical games. Therefore, the Nash equilibria and dominant strategies for both players were analyzed in all game structures. Games with the defective option as dominant strategy for both players and mutual defection as only equilibrium were classified as PD, games with two equilibria in which both players choose the same strategy were classified as Assurance, games with cooperating as dominant strategy for both players were classified as MD, and games with two equilibria in which both players choose different strategies were classified as Chicken.

To take the domain of the conflict scenarios into account, the descriptions participants gave were clustered using AI: a file containing the case IDs, participants' scenario descriptions and behavior option labels only was uploaded to Microsoft Copilot version 1.25121.84.0 with the prompt "Can you find and group factors in a list of scenarios? Like a factor analysis, so that in the end there are 5-7 thematic categories into which the scenarios are classified.". It resulted in six thematic categories and one spare category for games that didn't fit one of the defined categories. Games that were classified into this category were looked at manually and were assigned to one of the other categories as fit (e.g. when participants referred to one of the examples in the instruction texts by giving "Example 3" as scenario description and "A" and "B" as option labels). The resulting categories and category descriptions were as follows:

- Daily life: Everyday preference decisions with low stakes and no long-term consequences. These scenarios involve choosing activities, food, meeting places, entertainment, or simple lifestyle options. They reflect coordination rather than conflict.
- Household: Shared domestic responsibilities, chores, and maintenance tasks. These scenarios involve cleaning, cooking, organizing, or managing shared living spaces. The core dynamic is cooperation vs. fairness in shared household work.
- Relationship: Decisions that affect romantic partners or family members on a personal or emotional level. These scenarios include major life choices (children, moving, breakups) or interpersonal dilemmas within close relationships. Stakes are higher and outcomes affect long-term relational wellbeing.
- Work/Study: Collaboration, effort, fairness, and coordination in academic or professional settings. These scenarios often involve group projects, deadlines, effort levels, or decisions about cooperation vs. free-riding. The focus is on productivity, responsibility, and shared outcomes.
- Moral/Political: Scenarios involving ethical dilemmas, societal issues, political decisions, or public-goods problems. These include laws, climate action, war, justice, or collective responsibility. The stakes extend beyond the individuals involved and affect broader society.
- Risk/Strategy: Decisions involving uncertainty, competition, investment, or strategic advantage. These scenarios include gambling, business strategy, economic choices, or situations where outcomes depend on anticipating the other party’s move. They often resemble game-theoretic or payoff-based dilemmas.
- Other: Scenarios that are unclear, incomplete, abstract, or do not fit meaningfully into the other categories. This includes placeholders, meta-examples, or entries without a real decision structure. Used sparingly to maintain category integrity.

To examine the properties of the game structures participants created, the following game theoretic indices were used:

- Self-developed asymmetry index: absolute values of differences between own and partner's payoffs for the same strategy combination were averaged across all four cells of the payoff matrices.
- Symmetry (dichotomous): all game structures in which interchanging players would not impact the payoff matrix (asymmetry index equal to zero) were coded as symmetric, all others as asymmetric.
- Rapoport K index (Rapoport & Chammah, 1965): to evaluate the game structures' cooperative potential, the Rapoport K index was computed. As it was originally developed for the classification of symmetric PD-type games, it uses the following denotation of outcomes: the outcome for mutual cooperation is labeled as reward (R), the outcome for mutual defection is labeled as punishment (P), the outcome for cooperating with a defective partner is labeled as sucker (S), and the outcome for defecting on a cooperative parter is labeled as temptation (T). The index is then computed as $K = (R - P)/(T - S)$, with higher values indicating cooperation as being more attractive and defection as being less attractive. As the index was developed for symmetric game structures, it was computed for both perspectives separately with both perspectives' results being averaged to yield one value to quantify the game property.
- Game Harmony (Zizzo & Tan, 2007, 2011): to evaluate the game structures' cooperative potential, the Game Harmony index was used by computing the Pearson correlation of the two players' sets of payoffs. Higher values in Game Harmony indicate higher cooperative potential and less competition.
- Actor control, partner control, joint control, and interdependence (Kelley & Thibaut, 1978; Kelley et al., 2009; named following Stivers et al., 2025): the

structural components of the payoff matrices were computed to disentangle the interdependence resembled in the game structures. Actor control quantifies the influence players have over their own outcomes and is computed as the difference in average payoffs between the own strategies (difference in row averages for player 1, difference in column averages for player 2). Partner control quantifies the influence the partner has over the players' outcomes and is computed as the difference in average payoffs between the partners' strategies (difference in column averages for player 1, difference in row averages for player 2). Joint control quantifies the influence both players have over their outcomes by choosing the same or different strategies and is computed as difference between the diagonal and off-diagonal payoff averages. Lastly, interdependence (*ID*) quantifies the proportion of the partner's influence in all influencing factors of the own outcomes and is calculated as $ID = (PC^2 + JC^2)/(AC^2 + PC^2 + JC^2)$ from the three already mentioned indices actor control (*AC*), partner control (*PC*), and joint control (*JC*). All four indices were computed for both players separately using their respective payoff sets.

- Correspondence (Kelley & Thibaut, 1978; Kelley et al., 2009): to measure the degree of goal compatibility, correspondence is computed from the interdependence components by $2 * (AC_1 * PC_2 + AC_2 * PC_1 + JC_1 * JC_2)/(AC_1^2 + AC_2^2 + PC_1^2 + PC_2^2 + JC_1^2 + JC_2^2)$.
- Conflict of interest (Axelrod, 1967): to measure the degree of goal compatibility, the conflict of interest index was used, computed by $[(T - R) * (T - S)] / (T - P)^2$, using the PD denotation of outcomes, as PD games were the index's original scope. It was computed for both perspectives separately and perspectives were averaged. Higher values in conflict of interest indicate less cooperative potential.

- Game riskiness (Au et al., 2012; Ng & Au, 2016; Ng et al., 2024): to measure the players' perceived risk in choosing the cooperative behavior option instead of defecting, the risk index was used. It was originally developed to quantify PD games and therefore uses the PD notation in its equation: $(R - S)/(R - S + T - P)$. It was computed for both perspectives separately and perspectives were averaged.
- Price of anarchy and price of stability (Mak & Rapoport, 2013): to quantify efficiency losses caused by selfish, uncoordinated behavior (price of anarchy) or the best stable outcome being worse than the social optimum (price of stability), the two indices were computed based on the equilibria of the game structures: $PoA = \max_{s\in S} W(s) / \min_{s\in E} W(s)$ and $PoS = \max_{s\in S} W(s) / \max_{s\in E} W(s)$ with S denoting the whole strategy space of the game, E denoting the space of equilibria and $W(s)$ denoting the joint payoff of a strategy combination s. Higher values in price of anarchy indicate greater efficiency loss in acting selfishly compared to the social optimum, while higher values in price of stability indicate greater efficiency loss in coordinating on a stable outcome compared to the social optimum.

All indices were computed and interpreted as stated in Stivers et al. (2025) and the respective original publications.

Declaration on the Use of Artificial Intelligence (AI)

I hereby declare that I have used the following AI tools to write my master's thesis entitled *Constructing Conflict Templates: Task Format and Individual Differences Affect the Mental Representation of Social Dilemmas.* I confirm that the identification of the AI tools used is complete, including the intended use, product name, and, where applicable, prompts and outputs of the AI tool. I take full responsibility for the selection and adoption of all results from the AI tools I have used.

Purpose*	**Product name of the AI-tool**
Translation assistance	DeepL Version 26.3.1.19750
Support in writing and improving analysis code	OpenAI ChatGPT GPT-5.3
Support in clustering and connecting literature	Google NotebookLM
Linguistic proofreading	Microsoft Copilot Version 1.25121.84.0

* Tools for checking grammar and spelling or for automatically generating bibliographies are excluded from this declaration.