Territorial Gestalt in the Strategy of Conflicts

Zhen Li[1], Ning Tang[2], Siyi Gong[3], Jifan Zhou[1], Mowei Shen[1], Tao Gao[3]
[1]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China
(3160101331, jifanzhou, mwshen)@zju.edu.cn
[2]School of Education, Soochow University, Suzhou, China
ningtangcog@gmail.com
[3]Departments of Communication, Statistics and Psychology, Los Angeles, CA, USA
alicegong@g.ucla.edu, taogao@ucla.edu

Abstract

Human social interactions often involve both cooperation and competition. While coordination in cooperative settings has been well studied, less is known about how individuals resolve severe, protracted conflicts. In this study, we introduced a long-horizon territorial conflict game where participants competed for space on a two-dimensional board. Despite repeated interactions, conflict intensity did not subside over time—contrary to findings in simpler, one-shot matrix games. However, when a payoff-irrelevant color boundary was introduced, participants used this salient perceptual cue as a focal point for dividing the territory. The presence of this "territorial Gestalt" shifted strategies toward defensive postures, reduced the frequency of direct battles, and enabled opponents to settle conflicts precisely along the perceptual boundary. These findings extend focal point theory by demonstrating that humans naturally import external, payoff-irrelevant concepts into conflict situations to achieve coordinated outcomes. Our results highlight the importance of perception-based territorial Gestalt in fostering cooperative resolutions to otherwise intense and enduring disputes.

Keywords: territorial Gestalt; mixed-motive games; focal point; conflict resolution;

Introduction

Most human social interactions can be framed as games, which span all levels of daily life—from children chasing each other on the playground, to hunters cooperating in pursuit of prey, and even to high-stakes diplomatic maneuvers where entire continents become game boards, akin to Game Of Thrones. Reflecting this perspective, researchers are increasingly turning to game-based paradigms in cognitive science (Allen et al., 2024; Van Dijk & De Dreu, 2021). Unlike traditional psychophysical tasks — where a single participant passively responds to stimuli controlled by the experimenter — interactive games provide a real-time, dynamic environment in which players must adapt, negotiate, and strategize in response to one another's actions. Such experimental paradigms allow participants to function as autonomous agents who must synchronize perception, reasoning, and action on the fly, closely mirroring the demands of everyday life (Tomasello, 2024).

Viewing human social interactions as games highlights a tension between human cognition and classic game theory, which condenses all game-relevant information into a payoff matrix specifying the rewards for each combination of actions (Owen, 2013). Yet human cognition is grounded in our perception of the environment (Barsalou, 2008) imbued with rich structure and meaning, many of which could be payoff-irrelevant. This discrepancy raises a critical question: can spontaneous perception of the game environment, even when irrelevant to the official payoff, still shape strategic reasoning in a game?

The minimal group effect offers support, showing that arbitrary visual markers (e.g., shirt color (Dunham, Baron, & Carey, 2011)) can induce group identity and in-group favoritism in resource allocation (Richter, Over, & Dunham, 2016), behavioral attribution (Richter et al., 2016), and reciprocity (Diehl, 1990). However, typical minimal group experiments lack direct, multiplayer interactions, limiting their applicability to game-theoretic models that require real-time strategic reasoning among players.

Alternatively, Schelling's focal point theory—originating in economics (Schelling, 1980) and starting to gain traction in psychology (Allen et al., 2024) — argues that incorporating game-irrelevant concepts into coordination games is critical for achieving mutually beneficial outcomes. Factors 'outside of the game' are essential, because in multi-player games that require coordination, such as social dilemmas, traditional game theory often cannot identify optimal solutions using Nash equilibria or Pareto optima (Bénabou & Tirole, 2006; Dawes, 1980). The central puzzle in these games is: How do humans intuitively converge on solutions where traditional game-theoretic models fail to offer a clear solution? Previous research has shown that humans often rely on repetitive interactions and reciprocity strategies, such as 'tit-for-tat,' to foster cooperation (Axelrod & Hamilton, 1981; Nowak, 2006). However, these approaches depend solely on game-relevant information encoded in the payoff matrix.

In contrast, focal point theory offers a different solution—one that bypasses repetitive interactions and instead leverages people's ability to incorporate out-of-game concepts for coordination (Sugden & Zamarrón, 2006). When faced with multiple coordination options, individuals frequently converge on a single, mutually salient choice—a 'focal point' — even if its defining features are entirely irrelevant to the game's payoff. Schelling illustrates this with a hypothetical game where two players independently choose one cell in a 4×4 grid and are rewarded only if they select the same spot. From a standard game-theoretic perspective, all cells have identical payoffs, so players relying solely on utility calculations would choose randomly, resulting in just a 1/16 chance of coordination. However, Schelling argues that humans can achieve significantly higher coordination by selecting the upper-left cell. This choice is likely influenced by its unique position as a corner and its association with the starting point for reading and writing in many cultures. Such factors, which distinguish this cell as a focal point, clearly arise from broader visual-spatial and cultural concepts outside the formal payoff matrix.

Subsequently, Schelling's intuitive example has been rigorously tested through experimental games in economics and psychology. Studies show that adults can effectively use unique colors (e.g., selecting a red option over two white ones) and spatial positions (e.g., choosing a central option over peripheral ones) as focal points, exceeding chance-level success (Isoni, Poulsen, Sugden, & Tsutsui, 2019).

Studies also revealed the underlying cognitive mechanisms of how humans coordinate with focal points. Focal points are not merely direct low-level responses to salient stimuli, akin to bottom-up attention capture (Tversky & Kahneman, 1981). Instead, humans incorporate perceived salience into their theory-of-mind (ToM) reasoning, recognizing that their partner, like themselves, is likely to identify a uniquely salient option and use it to form mutual expectations (Bardsley, Mehta, Starmer, & Sugden, 2010). Studies have shown that successful coordination in games where participants must independently select the same option among many alternatives correlates with ToM capabilities but not with non-social traits like pattern detection or imagination (Curry & Chesters, 2012).

Despite these insights, three major gaps remain in understanding how focal points fit into game theory and the cognitive processes that support them. First, most empirical studies on focal point theory have focused solely on games inspired by Schelling's 'grid choice' example–a cooperative task where participants receive a reward for successful coordination or nothing at all for failure. Yet this focus diverges from the central objective of the focal point theory: limiting or preventing severe, unregulated conflicts, which is clearly evidenced by the title of his influential book The Strategy of Conflict rather than The Strategy of Cooperation.

Second, recent advances in mobile gaming and AI-driven gameplay have spurred the use of more complex online games — such as Diplomacy (Kramár et al., 2022) and Overcooked (Wu et al., 2021) — to explore human social interactions. Unlike one-shot matrix games, which conclude after a single decision, these long-horizon games involve sequences of actions and evolving states, requiring gradual and incremental coordination rather than instant agreements. However, focal points have yet to be examined in these prolonged settings. One key limitation of matrix games like the Prisoner's Dilemma is their strategic transparency — each choice (e.g., "cooperate" or "defect") clearly signals a player's intent, enabling direct coordination strategies such as tit-for-tat in repeated plays. In contrast, long-horizon games only reveal low-level actions, making it significantly harder to discern high-level strategic intentions. Focal points have never been tested in contexts where intentions must be inferred rather than explicitly stated. This gap is especially pressing in mixed-motive conflicts, where competition and cooperation dynamically evolve.

Third, current research has taken a narrow view of what constitutes a focal point, often constraining it to be the object with a unique feature. Yet Schelling's approach is considerably broader, drawing on Gestalt psychology to explain how 'out-of-game' contexts are mentally organized to shape strategic decision-making. As a notable example, he proposed that the human mind's spontaneous Gestalt organization of the world map — which serves as the game board of nations — can create focal points. Natural landmarks such as mountains, rivers, and straits carve the map into distinct territorial Gestalts with intuitively perceived boundaries. While these landmarks hold clear military, cultural, and economic significance (e.g., mountains and rivers provide stronger defenses than plains), Schelling argued that their importance goes beyond these tangible factors. Salient boundaries of territorial Gestalts also serve a psychological role as focal points, becoming arenas for psychological warfare where strategic intentions are asserted and contested.

Therefore, here we invent a new experimental game to test whether territorial Gestalt, even when completely game-irrelevant, can nevertheless limit the intensity of conflict. The results will provide empirical evidence to either support or challenge Schelling's focal point theory.

The Experimental Game of Territorial Conflict

Territorial Conflict is a two-player, grid-based strategy game with 20 rounds. It blends strategic depth with the controlled visual features of psychophysical experiments. To test whether payoff-irrelevant territorial Gestalt could facilitate settlement, unaware to participants, the battlefield's coloring was manipulated. In the Gestalt condition, using the 'similarity' principle, grids were color-coded into two perceptual groups — red and blue — which evenly split the battlefield into left and right halves. In the No-Gestalt condition, grids were uniformly colored either red or blue, with the color held constant within each pair but varying between different pairs. We also concealed this experimental manipulation by making it a between-participant variable, so that each pair of participants saw only one battlefield.

At its core, the game is designed to be both simple and strategically engaging. Players expand their territory by dispatching troops from their bases to adjacent grids, advancing one grid per round (Figure 1a, Planning Module). They collect wealth as 'tax' from occupied grids, which is gathered at the base and automatically converted into additional troops at a fixed rate of 1 wealth = 1 soldier. These troops can then be deployed to claim more grids. When opposing troops meet on the same grid, a battle occurs, with the larger force having a higher probability of victory (Figure 1a, Battle Module). The winning troops secure the grid, while the defeated troops retreat. To eliminate the need for players to micromanage individual soldier movements, the game focuses on high-level strategic planning. Players designate up to three fronts — each consisting of multiple consecutive grids — and allocate a percentage of their available troops to each. A linear programming algorithm optimizes troop movement toward these designated fronts.

In this game, the only action available to players is setting troop locations. Unlike typical matrix games with a distinguished 'cooperate' option, any coordination must emerge from unspoken mutual understanding inferred from an opponent's troop movements.

To maximize wealth, players must claim territory, inevitably clashing over the fixed game map—one player's gain is another's loss. Yet, conflict remains costly for both sides, Even the victor may find that the cost of prolonged war outweighs the benefits of conquest, creating incentives for mutually beneficial settlement, a dynamic central to Schelling's analysis of real-world conflicts. We introduce the following mechanics of causality and logistics:

1) **High Casualty Rates in Battle**: After his costly victory at Asculum (279 BCE), Pyrrhus remarked, "If we are victorious in one more battle with the Romans, we shall be utterly ruined." This pattern of pyrrhic victories has been extensively documented across military history (Stevenson, 2017). Our game reflects this reality through a probabilistic simulation model. The smaller army is more likely to lose, retreating upon reaching a 20% casualty rate, while the larger army is more likely to win but still suffers an average casualty rate of 10%. The heavy losses in battle directly translate into wealth depletion as 1 soldier = 1 wealth.

2) **Logistical Costs**: Even without direct conflict, sustaining a hostile relationship is costly, as keeping troops on the front line incurs steep logistical expenses that increase sharply with distance from the base — a principle that has shaped military campaigns from antiquity to modern times. Our game models this logistical constraint by increasing the rate at which troops deplete wealth as they move farther from the base. This makes expansion and prolonged warfare increasingly costly, incentivizing players to reach a settlement when possible and recall their troops back to the base to conserve resources.

To further emphasize the game's non-zero-sum nature, participants are instructed that their goal is to maximize their own wealth, not to defeat their opponent. Their performance and monetary compensation are based solely on their accumulated wealth, independent of their opponent's outcome. Participants in both groups were clearly informed that grid color does not affect tax collection, troop logistics, or battle outcomes. There was no mention of 'territory,' 'border,' or any Gestalt-related concepts, nor any indication of how participants should approach the game. After the experiment, participants completed a post-hoc questionnaire, reporting whether their decisions were influenced by the grid colors. We explored whether territorial Gestalts could constrain the battles in various aspects: First, the overall intensity of conflicts; Second, the specific partition of the battleground relative to the perceptual border; Third, the dynamics of conflicts, such as the directions and the speed of troop advancement.

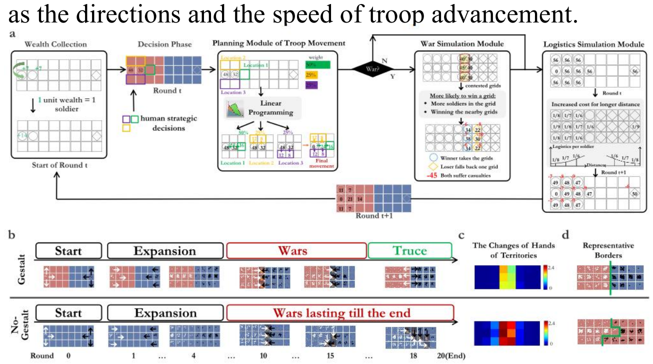

Figure 1: The dynamics and outcomes of the game.

Experiment 1: Territorial Gestalt of a Strip Battlefield

24 pairs of participants (age: M = 22.2, SD = 2.13; 23 males, 31 females) played on a game board consisting of three rows and eight columns (Figure 1a). The maximum allowed decision time for each round was 40s. The same pair of participants played 12 trials (3 practice trials) in a 90 minutes experiment. The experiment was conducted using two networked computers, with each participant located in a separate room. Communication between participants was strictly prohibited to ensure independent decision-making.

First, we explored whether repetitive play promotes settlement by measuring whether each pair of participants engaged in a battle at each round and averaged these instances across participants and trials (Figure 2). We analyzed it using the logistic regression model with trial and round as the fixed factors and the subjects as the random factor. There was a significant main effect of round (β = 0.117, SE = 0.007, z = 16.07, p < .001), suggesting that the intensity of conflicts continued to increase within a single trial. There was no significant effect of trial (β = 0.013, SE = 0.013, z = 0.946, p = .334). These results suggested that coordinated settlements did not emerge through repetitive play, either within or across trials.

The benefits of territorial Gestalt in promoting settlement were also supported by several aggregated measures across all rounds and trials. Compared to the No-Gestalt condition, the Gestalt condition featured fewer battles (5.25 vs. 10.32, t(22) = -3.78, p = .001, Cohen's d = -1.54) and fewer casualties (265 vs. 438, t(22) = -4.49, p < .001, Cohen's d = -1.83). Participants in the Gestalt group also enjoyed longer truces— consecutive conflict-free rounds by the end of the game, compared to the No-Gestalt group (Figure 1b, 4.84 vs. 0.85, t(22) = 3.04, p = .006, Cohen's d = 1.24). These findings collectively underscored that territorial Gestalt reduced overall conflict intensity.

Focal point theory predicted not only an overall reduction in conflict but also that settlements will form precisely along the territorial Gestalt boundary. To test this, we focused on conflicts occurring around the Gestalt border grids. First, we examined how fiercely the six border grids were contested, measured by the number of times each grid changed hands per game. The results showed that changes of hands were significantly higher in the No-Gestalt condition than in the Gestalt condition (Figure 1c, 7.40 vs. 16.19, $t(22) = -3.49$, $p = .002$, Cohen's $d = -1.43$). Furthermore, in the Gestalt condition, the final boundary between players was shorter (3.13 vs. 3.44, $t(22) = -2.91$, $p = .008$, Cohen's $d = -1.19$), and more likely to serve as the exact final border between the two players (53% vs. 24%, $t(22) = 2.57$, $p = .018$, Cohen's $d = 1.05$). This was illustrated by irregular zig-zag final borders in the No-Gestalt condition (Figure 1d).

These findings together supported the idea that territorial Gestalt served as an intuitive focal point, guiding settlement patterns and structuring territorial division.

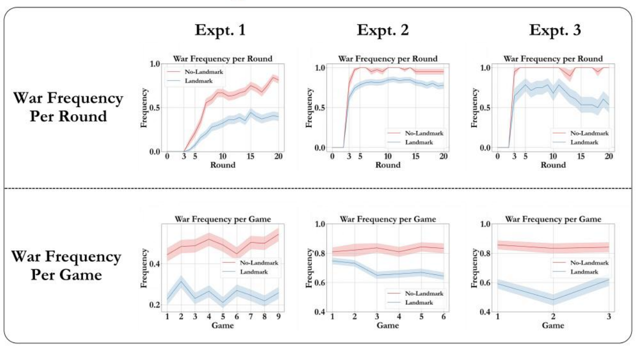

Figure 2: The war frequency in different trials and rounds.

Experiment 2: Dynamics of Conflict

One limitation of Expt.1 was that the battlefield was a horizontally elongated strip, which restricted the directions of troop advancement. Here we introduced a square battlefield, providing participants with the freedom to choose the orientation of troop advancement — whether horizontal, vertical, or diagonal. We manipulated the territorial Gestalt to bisect the battlefield either vertically or horizontally (see Figure 3a). This modification allowed us to not only compare the presence of territorial Gestalt to the No-Gestalt condition but also examine differences in participants' strategies based on different Gestalt orientations. With 72 participants (age: M = 20.9, SD = 2.13; 30 males, 42 females; 12 pairs in each of the vertical, horizontal, and No-Gestalt conditions), we could explore whether conflict dynamics aligned with the orientation of the territorial Gestalt. The maximum decision time for each round is 40s. The same pair of participants played 7 trials (1 practice trials) in a 90 minutes experiment.

We first examined the effects of repetitive play in the same way as in Experiment 1. There was a significant main effect of round ($\beta = 0.103$, $SE = 0.009$, $z = 11.30$, $p < .001$), suggesting that the war frequency increased rapidly and then held steady, especially in the No-Gestalt condition where a battle occurred in almost every round– an extremely fierce conflict (Figure 2). There was also a significant main effect of trial ($\beta = -0.101$, $SE = 0.028$, $z = -3.66$, $p < .001$). According to Figure 2, this effect was largely caused by the data from the Gestalt condition. When the data from the Gestalt condition was removed, there was no significant effect of trial ($\beta = 0.070$, $SE = 0.079$, $z = 0.89$, $p = .371$). These results confirmed that in the No-Gestalt condition, coordinated settlements did not emerge through repetitive play, either within or across trials.

We then examined the presence of territorial Gestalt by aggregating data from the vertical and horizontal conditions. Overall, conflicts on this square battlefield were more intense than on the strip battlefield, possibly because the openness of the area fueled the aggression of the stronger force. The open battleground offered players the opportunity to outflank their opponents, thereby annihilating them by taking control of all grid – which never occurred in Experiment 1. Notably, the annihilation rate was significantly lower in the Gestalt condition compared to the No-Gestalt condition (18.8% vs. 44.4%, $p = .007$, Cohen's d = -1.02). Yet in non-annihilation trials, we found that the presence of territorial Gestalt could still mitigate the overall intensity of battles, leading to fewer battles (14.95 vs. 17.39, $t(34) = -2.89$, $p = .007$, Cohen's $d = -1.02$), battle grids per battle (4.51 vs. 5.41, $t(34) = -2.38$, $p = .023$, Cohen's $d = -0.84$), and casualties (37.8 vs. 45.1, $t(34) = -1.60$, $p = .024$, Cohen's $d = -0.57$).

After examining the overall impact of territorial Gestalt on conflict intensity, we now focus on whether conflicts and settlements were centered around the Gestalt boundary. Unlike Experiment 1, which only compared the presence or absence of a territorial boundary, this experiment allowed us to explore how different Gestalt orientations influenced conflict patterns, providing deeper insight into how boundary structure shaped strategic interactions. The results showed that the presence as well as the exact location of the territorial boundary had an overwhelming impact on where the conflict broke out and potentially settled. This pattern was clearly illustrated by heatmaps (see Figure 3a) showing occupancy frequency — a more reddish grid indicated frequent occupation by the player with the base in the upper-left corner, while a more blueish grid signified control by the player with the base in the lower-right corner. In the Horizontal condition, a horizontal division emerged, whereas in the Vertical condition, there is a vertical division. Grid occupation in the 10th round revealed a strong tendency for participants to divide the battlefield according to the orientation of the perceptual grouping. In contrast, no clear pattern of territorial division appeared in the No-Gestalt condition, with the occupancy frequency mostly determined by the distance to each base.

Next, to further quantify the effect of territorial Gestalt on the progression of conflicts, we defined the orientation of each conflict as the angle of a least square regression line based on the coordinates of all grids engaged in battles during that round. Histograms of battle orientation demonstrated (see Figure 3c) that in the Vertical condition, the battles were overwhelmingly 90°, while in the Horizontal condition, they were predominantly 0°. The No-Gestalt condition displayed a more diverse array of degrees, ranging from 0°, 45° to 90° (Kolmogorov-Smirnov test, Vertical vs. No-Gestalt: p = .001; Horizontal vs. No-Gestalt: p = .018). In the Vertical condition, the average orientation of battles was 65°, significantly larger than that in the No-Gestalt condition (Figure 3d, Vertical: M = 65.0; No-Gestalt: M = 47.7, $t(22)$ = 2.90, p = .008, Cohen's d = 1.18). Comparatively, in the Horizontal condition, the average angles of battles was 29°, significantly smaller than the No-Gestalt condition ($t(22)$ = -2.88, p = .009, Cohen's d = -1.17). Taken together, these findings suggested that rather than roaming the board freely as in the No-Gestalt condition, participants in the two Gestalt conditions followed a more boundary-focused strategy. They initially secured grids within the color-based territory surrounding their base, then concentrated their troops along the territorial boundary, with conflicts emerging and remaining constrained to that dividing line.

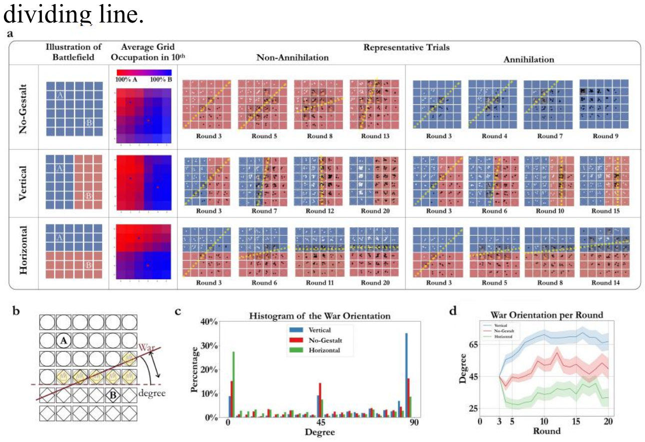

Figure 3: Battle strategies and dynamics in Experiment 2.

Experiment 3: Explicit Report of Strategy

While the first two experiments demonstrated how territorial Gestalt affects troop movements, they provided no direct evidence of whether forces were defending a border or launching an attack. This lack of strategic transparency is a major challenge in long-horizon, mixed-motive games compared to matrix games, making it harder both for players to coordinate and for researchers to discern their decision-making processes. In the current experiment, we addressed this gap by asking participants to record their strategic intentions — attack or defense — after each round, thereby revealing more clearly how territorial Gestalt shapes decision-making.

This experiment used the same task as Experiment 2, with one key modification: after each round, participants must explicitly record their current strategic goal. They could specify their strategy at the grid level by selecting “attacking grids” or “defending grids” and clicking the target grids on the game board. Alternatively, they could specify their strategy at the line level, such as “attacking lines” or “defending lines,” and then clicking all grids in those lines. Participants could also choose broader goals, such as “attack” or “defend,” without specifying a specific target, or they could freely type their strategy in an open-response field to describe their approach. The other player cannot view their opponent's strategy reports and must infer their intentions, as in the previous two experiments. 12 pairs of participants were assigned to the Gestalt condition, featuring a vertically split game board with grids color-coded, and another 12 pairs were assigned to the No-Gestalt condition (age: M = 22.0, SD = 3.41; 24 males, 24 females). The decision time for each round was not limited. The same pair of participants played 4 trials (1 practice trials) in a 90 minutes experiment. We first examined the effects of repetitive play in the same way as in Experiment 1. There was significant main effect of round (β = 0.029, SE = 0.014, z = 2.01, p = .045) and no significant main effect of trial (β = 0.093, SE = 0.096, z = 0.97, p = .334), confirming that coordinated settlements did not emerge through repetitive play, either within or across trials.

Second, analyses of the conflicts replicated the main findings from Experiment 2, showing that the presence of territorial Gestalt reduced conflict intensity, leading to fewer battles (11.7 vs. 16.8, $t(19)$ = -2.60, p = .017, Cohen's d = -1.15), battle grids per battle (3.70 vs. 5.92, $t(19)$ = -2.48, p = .023, Cohen's d = -1.10), and casualties (32.4 vs. 49.7, $t(19)$ = -2.23, p = .038, Cohen's d = -0.98). It also led to a smaller annihilation rate, which was marginally significant (22.2% vs. 47.2%, $t(22)$ = -1.95, p = .064, Cohen's d = -0.79). Next, we analyzed participants' reported strategies after round 3 (when their troops first engaged). To assess the impact of territorial Gestalt on overall aggressiveness, we aggregated all reported strategies into attack versus defense categories. The results revealed that participants in the Gestalt condition were significantly more likely to adopt a defensive strategy compared to those in the No-Gestalt condition (67.4% vs. 50.3%, $t(22)$ = -2.49, p = .021, Cohen's d = -1.02), suggesting that the presence of territorial Gestalt promotes a more defense-oriented approach. In the Gestalt condition, there were also more “mutual defense” rounds in which both players chose a defensive strategy ($t(22)$ = 2.78, p = .011, Cohen's d = 1.13) as shown in Figure 4b. We then analyzed the distribution of grid- and line- level strategies, which was shown in Figure 4a. The distribution of strategies between Gestalt and No-Gestalt conditions were significant (Kolmogorov-Smirnov test, p = .001). In the Gestalt condition, the most frequently reported strategy was defending a line, significantly higher than in the No-Gestalt condition (55.6% vs. 23.5%, $t(19)$ = 2.73, p = .012, Cohen's d = 1.11).

We next visualized how often each grid was included in attack and defense strategies (Figure 4c). In the No-Gestalt condition, both attack and defense choices were scattered across the board with no obvious pattern. By contrast, in the Gestalt condition, players' strategic choices — whether to attack or defend — were heavily concentrated along the Gestalt border. To quantify this observation, we summed the selections in the two central columns (which represent the border in the Gestalt condition). An ANOVA confirmed a significant main effect of Gestalt condition (Figure 4d, 61.2% vs 34.1%, $F(1, 37)$ = 21.19, p < .001, η_p^2 = .36), indicating that territorial Gestalt indeed transforms the central columns into a focal point of strategic reasoning.

These results of self-reported strategies suggested that a large, unstructured board encouraged greater aggression between players, leading to more frequent and intense conflicts. By contrast, introducing a territorial Gestalt — even when it offered no direct payoff advantage—prompted participants to shift toward defense. Importantly, their defensive maneuvers focused on the Gestalt-defined border, rather than any arbitrary line, illustrating how visual structure can guide and constrain strategic reasoning.

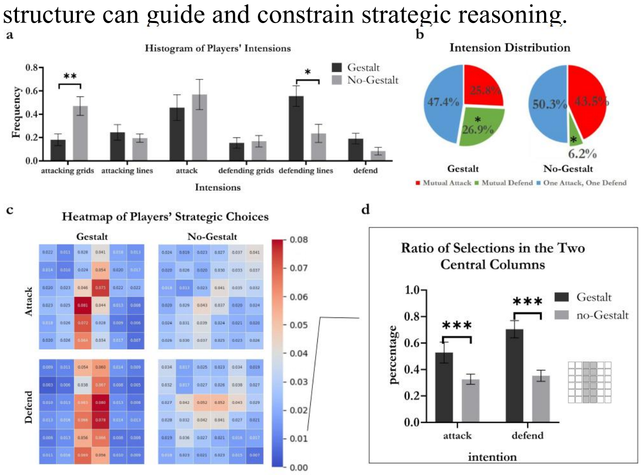

Figure 4: Results in Experiment 3.

Discussion: Gestalt-Based Strategic Reasoning Through a Shared Conceptual Space

Our findings offer valuable insights into how coordination unfolds—or fails to unfold—in long-horizon, mixed-motive conflict games where strategic intentions must be inferred from low-level troop movements.

One of the key contributions of this study is expanding the focus from unique visual features to territorial Gestalt, positioning focal point theory within the broader framework of modern cognitive science. While the importance of unique features — such as those that "pop out" in visual search tasks — has been well-documented in the classic feature-integration theory (Treisman & Gelade, 1980) in cognitive psychology, Gestalt principles offer a more generalizable framework, describing how the human mind actively constructs discrete wholistic entities from sensory input. While it is continues to be a central theme of visual perception (Wagemans et al., 2012), as a generic principle it has been deeply embedded in different branches of psychology, such as field dynamics in social psychology (Lewin, 1951) and figure – ground distinction in cognitive semantics (Jackendoff, 1985; Talmy, 2000). Demonstrating territorial Gestalts as a notable type of focal point provides a deeper understanding of the mechanisms behind focal points from a psychological perspective, offering a valuable complement to economic and game-theoretical frameworks.

We draw upon the conceptual space hypothesis in cognitive semantics (Jackendoff, 2008), which directly links perceptual Gestalt to strategic planning by proposing a unified space shared by vision, language, and action. This shared conceptual space enables humans to both translate visual scenes into linguistic descriptions and linguistic commands into sequence of actions executed in the scene. Discrete perceptual entities shaped by Gestalt principles — such as grouped or bounded regions — enter this space as outputs of perception, serving as foundational building blocks for linguistic semantics and action planning. A vivid example of how language and planning are deeply rooted in vision (Pinker, 2003) is the phrase "Don't cross the line!" commonly used in conflict resolution. Here, a simple visual boundary ('the line') becomes a linguistic metaphor (Lakoff & Johnson, 1980) that anchors strategic thinking in competition and extends to abstract, non-visual domains like commercial rivalry. In contrast, the phrase 'entering the gray area,' a visual metaphor for the absence of clear perceptual boundaries, is used to describe situations of ambiguity, encouraging cautious and thoughtful decision-making.

Despite emerging evidence that vision and language share representational systems (e.g., symmetry (Hafri, Gleitman, Landau, & Trueswell, 2023), boundaries (Papafragou & Ji, 2023; Strickland et al., 2015)), surprisingly few studies have explored the direct impact of visual perception on strategic reasoning. To explain how territorial Gestalt operates as a focal point, we propose a framework that integrates theories of visual perception, conceptual space, planning, and ToM:

Automatic Gestalt Formation: The visual system automatically segments the game board into discrete territorial Gestalts, independent of their payoff relevance or top-down reasoning. As the output of perception, these Gestalts enter the shared conceptual space, where they serve as building blocks of other cognitive processes.

Goal-Oriented Planning: Human planning is inherently goal-directed. Individuals identify goals and commit to execute sequences of actions to achieve them (Holton et al., 2024; Molinaro & Collins, 2023). We argue that in a visual scene, what constitutes a goal in human planning is constrained by perceptual outputs in the conceptual space. In the No-Gestalt condition, the absence of higher-level visual structures forces players to navigate an unstructured goal space, where all grids serve as potential goals with low probabilities, resulting in high-entropy of strategic intentions. In contrast, the Gestalt condition introduces a salient territorial boundary that highlights a smaller set of high-level goals, which are more likely to be strategically prioritized. Simplified and sparse perceptual groupings provide a more computationally efficient framework for planning compared to operating at the grid level of the raw gameboard (Gershman, Horvitz, & Tenenbaum, 2015).

Applying this territorial framework to our game illustrates how 'Don't cross the line' functions not merely as a linguistic metaphor but as a literal guide to behavior. In this context, a visually salient boundary serves as a focal point, anchoring both players' expectations for settlement. Crossing that boundary carries grave psychological consequences, such as signaling aggression, violating trust, or escalating conflict, even when it lies entirely outside the game's explicit payoff structure.

References

Allen, K., Brändle, F., Botvinick, M., Fan, J. E., Gershman, S. J., Gopnik, A., . . . Ho, M. K. (2024). Using games to understand the mind. *Nature Human Behaviour, 8*, 1035-1043. doi:https://doi.org/10.1038/s41562-024-01878-9

Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science, 211*(4489), 1390-1396.

Bardsley, N., Mehta, J., Starmer, C., & Sugden, R. (2010). Explaining focal points: Cognitive hierarchy theory versus team reasoning. *The Economic Journal, 120*(543), 40-79.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol., 59*(1), 617-645. doi:https://doi.org/10.1146/annurev.psych.59.103006.093639

Bénabou, R., & Tirole, J. (2006). Incentives and prosocial behavior. *American Economic Review, 96*(5), 1652-1678.

Curry, O., & Chesters, M. J. (2012). 'Putting ourselves in the other fellow's shoes': The role of 'theory of mind'in solving coordination problems. *Journal of Cognition Culture, 12*(1-2), 147-159.

Dawes, R. M. (1980). Social dilemmas. *Annual review of psychology*.

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European review of social psychology, 1*(1), 263-292.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child development, 82*(3), 793-811.

Gershman, S. J., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science, 349*(6245), 273-278.

Hafri, A., Gleitman, L. R., Landau, B., & Trueswell, J. C. (2023). Where word and world meet: Language and vision share an abstract representation of symmetry. *Journal of Experimental Psychology: General, 152*(2), 509-527. doi:https://doi.org/10.1037/xge0001283

Holton, E., Grohn, J., Ward, H., Manohar, S. G., O'Reilly, J. X., & Kolling, N. (2024). Goal commitment is supported by vmPFC through selective attention. *Nature Human Behaviour*, 1-15.

Isoni, A., Poulsen, A., Sugden, R., & Tsutsui, K. (2019). Focal points and payoff information in tacit bargaining. *Games and Economic Behavior, 114*, 193-214.

Jackendoff, R. S. (1985). *Semantics and cognition* (Vol. 8): MIT press.

Jackendoff, R. S. (2008). *Patterns in the mind: Language and human nature*: Basic Books.

Kramár, J., Eccles, T., Gemp, I., Tacchetti, A., McKee, K. R., Malinowski, M., . . . Bachrach, Y. (2022). Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy. *Nature Communications, 13*(1), 7214. doi:https://doi.org/10.1038/s41467-022-34473-5

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive science, 4*(2), 195-208.

Lewin, K. (1951). Intention, will and need.

Molinaro, G., & Collins, A. G. (2023). A goal-centric outlook on learning. *Trends in cognitive sciences*.

Nowak, M. A. (2006). Five rules for the evolution of cooperation. *Science, 314*(5805), 1560-1563.

Owen, G. (2013). *Game theory*: Emerald Group Publishing.

Papafragou, A., & Ji, Y. (2023). Events and objects are similar cognitive entities. *Cognitive psychology, 143*, 101573. doi:https://doi.org/10.1016/j.cogpsych.2023.101573

Pinker, S. (2003). *The language instinct: How the mind creates language*: Penguin uK.

Richter, N., Over, H., & Dunham, Y. (2016). The effects of minimal group membership on young preschoolers' social preferences, estimates of similarity, and behavioral attribution. *Collabra, 2*(1), 8.

Schelling, T. C. (1980). *The Strategy of Conflict: with a new Preface by the Author*: Harvard university press.

Stevenson, D. (2017). *1917: war, peace, and revolution*: Oxford University Press.

Strickland, B., Geraci, C., Chemla, E., Schlenker, P., Kelepir, M., & Pfau, R. (2015). Event representations constrain the structure of language: Sign language as a window into universally accessible linguistic biases. *Proceedings of the National Academy of Sciences, 112*(19), 5968-5973. doi:https://doi.org/10.1073/pnas.1423080112

Sugden, R., & Zamarrón, I. E. (2006). Finding the key: the riddle of focal points. *Journal of Economic Psychology, 27*(5), 609-621. doi:https://doi.org/10.1016/j.joep.2006.04.003

Talmy, L. (2000). *Toward a cognitive semantics: Concept structuring systems*: MIT press.

Tomasello, M. (2024). *Agency and Cognitive Development*: Oxford University Press.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453-458.

Van Dijk, E., & De Dreu, C. K. (2021). Experimental games and social decision making. *Annual review of psychology, 72*(1), 415-438. doi:https://doi.org/10.1146/annurev-psych-081420-110718

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & Von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–

ground organization. *Psychological bulletin, 138*(6), 1172-1217.

Wu, S. A., Wang, R. E., Evans, J. A., Tenenbaum, J. B., Parkes, D. C., & Kleiman-Weiner, M. (2021). Too many cooks: Bayesian inference for coordinating multi - agent collaboration. *Topics in Cognitive Science, 13*(2), 414-432.