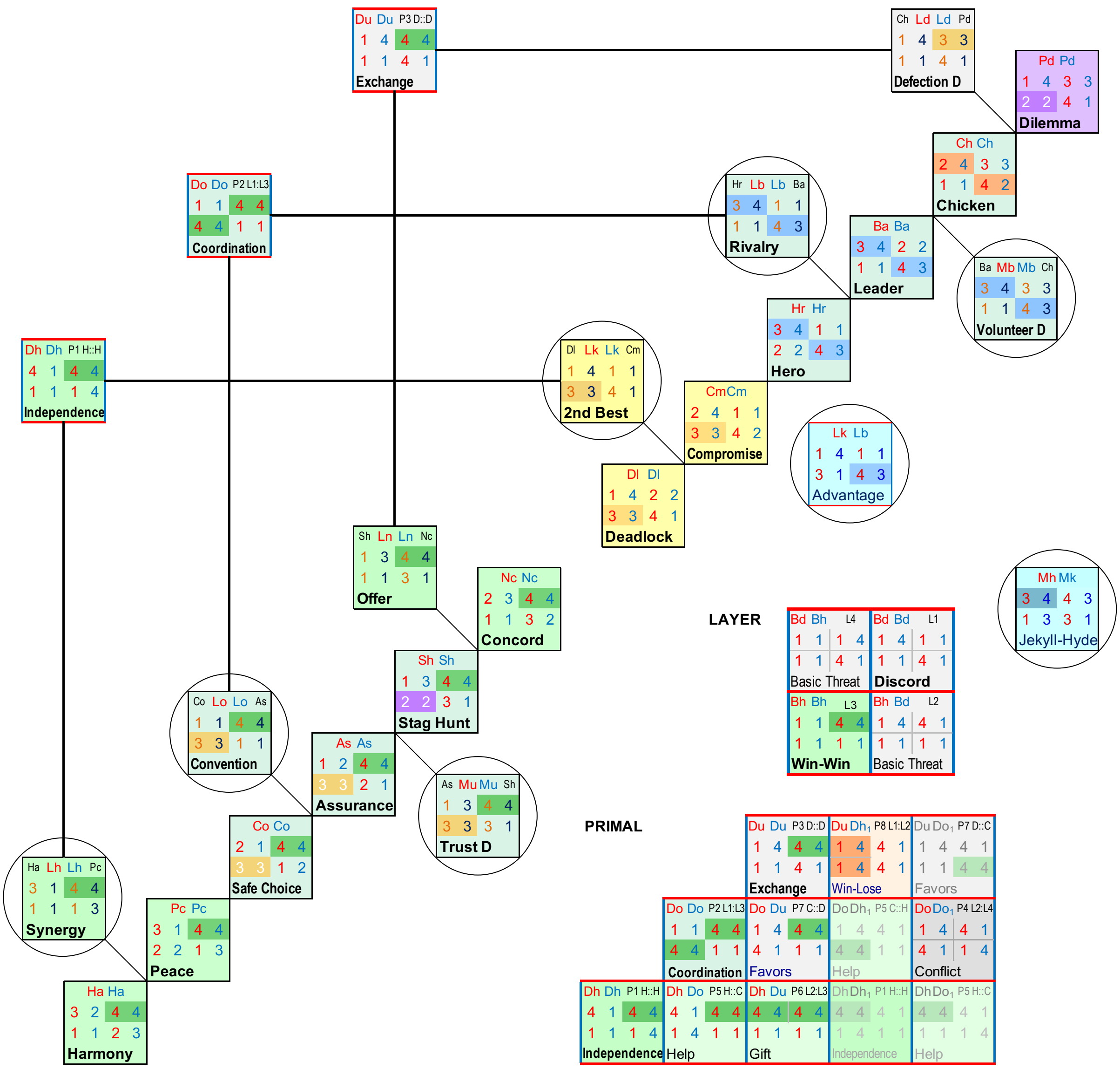
Appendix. Archetypes in the topology of 2x2 games

Appendix to Bruns, B., and C. Kimmich. 2021. Archetypal games generate diverse models of power, conflict, and cooperation. *Ecology and Society*.

Graphical abstract. Simplifying two-person two-choice (2x2) games by making ties in payoff ranks (indifference between outcomes) derives three primal archetypes of interdependence. Payoff patterns from the symmetric archetypes for independence, coordination, and exchange combine to form asymmetric archetypes for power, dependence, and conflict. Breaking ties in primal archetypes generates intermediate archetypes for synergy, compromise, conventions, rivalry, and advantage. Archetypal games provide a menu of models for understanding institutional diversity and transformation in social-ecological situations.



Contents

1. Introduction to the appendix

This appendix provides further results and discussion about families of asymmetric archetypes, the topology of 2x2 games, games with ties that lie between the strict ordinal games, the prevalence of bias games, payoff matrices for examples of interdependence, changing preferences, dimensions of interdependence, and some limitations and extensions of this approach to archetypal games.

The main paper identified archetypes for 2x2 games and showed how they generate diverse models for interdependence:

- Simple game theory models of independence, coordination, and exchange combine payoff patterns to make asymmetric situations of power, dependence and conflict.
- These archetypal games differentiate (by breaking ties in outcome rankings) to generate archetypes for synergy, compromise, conventions, rivalry, and advantage.
- Archetypal games offer a menu of models for understanding institutional diversity and transformation in social-ecological systems.

This appendix discusses how archetypes fit into the topology of 2x2 games. In a sense, the topology of 2x2 games provides an intellectual framework and scaffolding for identifying archetypal games. The main paper introduces families of archetypal games, while this appendix puts the archetypal games into the context of the topology of 2x2 games.

- Breaking ties in primal archetypes generates families of strict 2x2 games.
- The topology of payoff swaps in two-person two-choice (2x2) games maps overlapping relationships among archetypes with various challenges for cooperation.
- Ties make simpler games between strict 2x2 games.
- Making high ties and low ties in the strict 2x2 games derives primal archetypes including variants equivalent by interchanging rows and columns.
- Symmetric 2x2 games mostly offer relatively good outcomes at equilibrium (best or second-best) except for an unstable region around Prisoner's Dilemma.
- Most 2x2 games have unequal payoffs at equilibrium

- Standardizing payoff matrices shows equivalence in models between the Atlas of Interpersonal Situations and the topology of 2x2 games.
- Making and breaking ties represents changes in preferences. There are many reasons why payoff values might change.
- Archetypal games illustrate dimensions of independence including Nash best response equilibria, coordination, and externalities.
- The ways in which simple archetypes generate more diverse situations offer tools for understanding similarity and diversity in interdependence.

2. Asymmetric archetypes differentiate into families of hotspots and pipes

Figures A1 and A2 show how all eight primal archetypes differentiate into symmetric and asymmetric descendants.

- Primal Coordination differentiates into games with rival equilibria or games where achieving the best payoff for both may conflict with avoiding risk.
- Primal Conflict, also known as Matching Pennies, differentiates into a family of cyclic games with no equilibria. These are situations of completely opposed interests where if one gains the other does worse.
- Primal Win-Lose yields further asymmetric inequality in high bias (4,2) games where one gets second-worst at equilibrium.
- Primal Gift, Primal Independence, and Helping Hand produce families of games with relatively good outcomes that each contain win-win (4,4), moderate bias (4,3), and second-best (3,3) outcomes.
- Primal Exchange and Primal Combination have particularly diverse descendants, including asymmetric dilemmas with inefficient equilibria, games with highly unequal (4,2) equilibrium outcomes, and cyclic games where a focal point (4,3 or 3,3) could offer a Pareto-optimal solution better than cyclic instability or the poor payoff from a mixed strategy. Most of these games, including the social dilemmas of Prisoner’s Dilemma, Chicken, and Stag Hunt as well as their asymmetric neighbors, share defection problems where there are motivations to move away from a cooperative Pareto-optimal solution and instead become trapped in a result that is unsatisfactory for one or both.

Figure A1. Hotspot families. Primal archetypes for Coordination, Conflict, Win-lose, and Gift differentiate into two tiles composed of four strict ordinal games linked by low swaps (1><2).

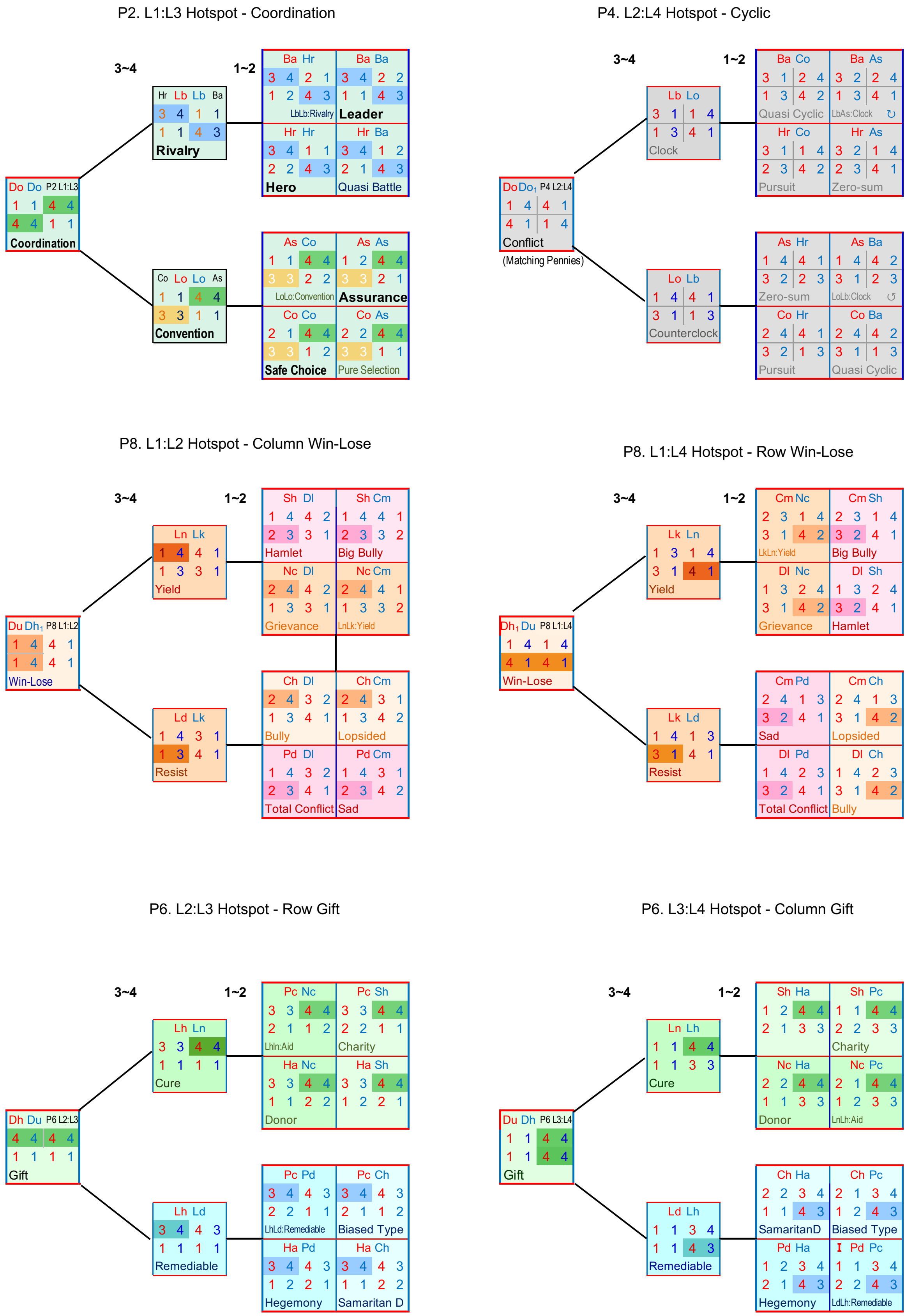
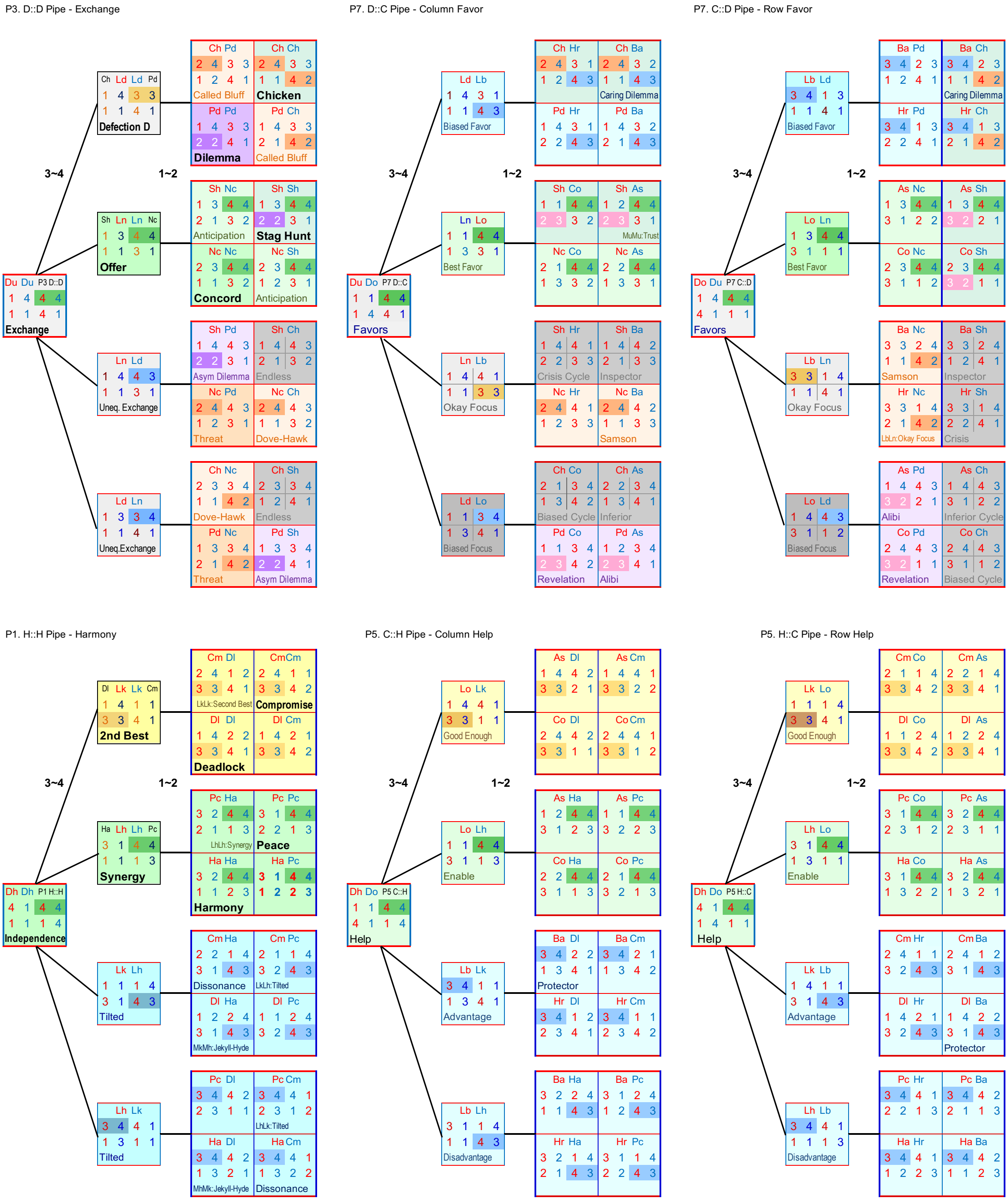


Figure A2. Pipe families. Primal archetypes for Exchange, Independence (Harmony), Favors, and Help differentiate into four tiles with sixteen strict games.



3. The topology of payoff swaps in 2x2 games

Figure A3 displays the topology of 2x2 games (Robinson and Goforth 2005, Robinson et al. 2007, Bruns 2015), which provides the framework for analysis in this paper. Symmetric games form a diagonal axis from lower left to upper right. Their payoffs combine to make asymmetric games. Making ties in payoff ranks simplifies games to form archetypes, breaking ties differentiates primal archetypes into the strict ordinal 2x2 games shown in the table.

In the topology of 2x2 games, games linked by swaps in the two lowest-ranked outcomes are considered nearest neighbors. Four games linked by low swaps (1><2) form a *tile*, as on the right hand side of Figures A1 and A2. In Figure A3, an example is the tile with Assurance, Safe Choice, and the combinations of their payoff patterns. Middle swaps (2><3) create a neighboring game to start a new tile. Continuing this process until the payoff structures repeat creates a *layer* of nine tiles and thirty-six games. High swaps (3><4) start a new layer, for example when Stag Hunt turns into an Asymmetric Dilemma. Each layer forms a torus (doughnut) shape that can be "cut open" and displayed on a flat square. So, payoff swaps link games at the top of a layer to those at the bottom. Payoff swaps also link games from side-to-side in a layer (like an early video game where a spaceship leaving one edge reappears on the other).

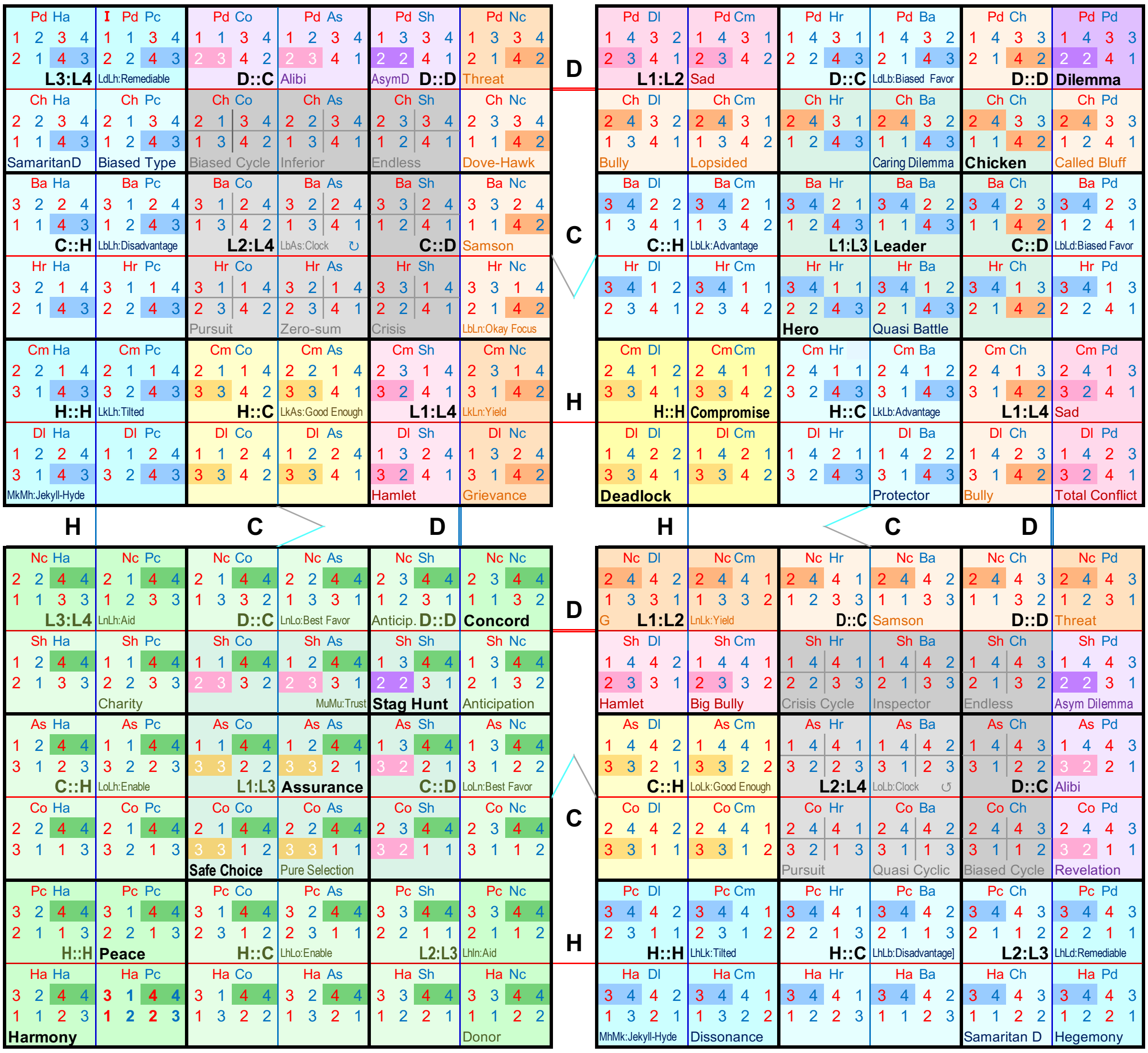
Each actor could have payoff swaps for low, middle, or high payoffs (1><2, 2><3, 3><4). Therefore each game has six neighbors. Low and middle swap neighbors are shown in each layer, while high swaps transform into games on another layer. The topology of 2x2 games provides a map of the "adjacent possible" (Kauffman 1995) of potential transformations resulting from changes in the ranking of outcomes.

The four layers in the topology of 2x2 games differ by the alignment of the best outcomes. In the discord layer (Layer 1) on the upper right of Figure A3, the best outcomes are in diagonally opposed cells. In the win-win layer (Layer 3) on the lower left, the best outcomes (4,4) are in the same cell. In Layers 2 and 4 the best outcomes are in the same row or column. Making ties for the three lowest payoffs simplifies all the games in a layer into one of the four basic archetypes: Win-Win for Layer 3, Discord for Layer 1, or Row or Column Threat (top outcomes in the same row or column) for Layers 2 and 4.

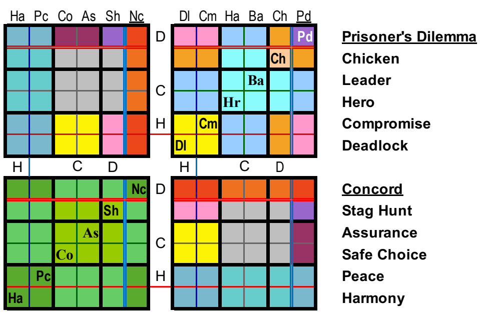
As shown in Figure A1, some primal archetypes such as Primal Coordination and Primal Conflict ultimately generate eight strict ordinal games. These are arranged in two tiles that form a *hotspot*. In a hotspot, two tiles on different layers are linked by swaps in the two highest-ranked payoffs (3><4). Other primal archetypes, such as Primal Independence, Primal Exchange, and their neighbors generate sixteen strict ordinal games in four tiles, forming *pipes*, as shown in Figure A2. In a pipe, high swaps connect four tiles, one on each layer.

Figure A3 locates the hotspots and pipes in the topology of 2x2 games. Hotspots can be identified by the layers they link. Thus the coordination hotspot links layers 1 and 3. The cyclic hotspot links layers 2 and 4. Pipes link tiles in equivalent locations on each of four layers. For example, high swaps link the Harmony tile to equivalently located tiles on the lower left of each layer (H::H pipe). Neighboring tiles above or to the right are similarly linked in quartets of tiles (C::H and H::C). The same applies for the tiles (D::D) on the upper right of each layer, and their neighboring tile quartets to the left and below (C::D, D::C). As shown in Figures A1 and A2, each hotspot or pipe simplifies into a primal archetype, an ancestor (progenitor) that differentiates into a family of games.

Figure A3. Hotspots and pipes in a "standard layout" of the topology of 2x2 games with Prisoner's Dilemma in an outer corner. Figures A3b and A3c show how "scrolling" the display of the torus-shaped layer moves Prisoner's Dilemma from an outer corner to an inner corner. This splits open tiles and creates the dominant strategy layout shown in Figure A4.



b. Standard Layout - Pd in outer corner



c. Dominant Strategy Layout - Pd in center

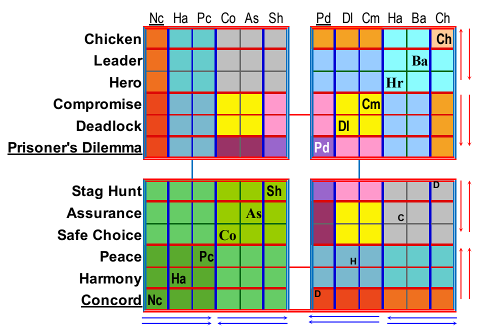


Figure A4 is a “dominant strategy” layout visualizing the topology of 2x2 games that elegantly displays many of the relationships between games (Robinson and Goforth 2005, Bruns 2015). Compared to Figure A3, the display in each layer “scrolls” to move Prisoner’s Dilemma to the inner corner, as shown in Figures A3b and A3c. In this layout, the quadrants within each layer differ by the alignment and number of dominant strategies. Games in the lower left quadrant of each layer have two dominant strategies. In the upper left and lower right quadrants, only one actor has a dominant strategy. If the other player can anticipate the dominant strategy, then their best move becomes clear. Games in these three quadrants all have a single equilibrium, resulting from dominant strategies for one or both actors. Games in the upper right quadrants have no dominant strategy. In pure (unmixed) strategies, they have either two equilibria, as in the coordination games, or no equilibrium, as in the cyclic games. Thus there are two quadrants of coordination games. In more colloquial terms, the diversity of 2x2 games without dominant strategies includes a herd of risky stag hunts and a bunch of rivalrous battles, as well as two clumps of cyclic conflicts.

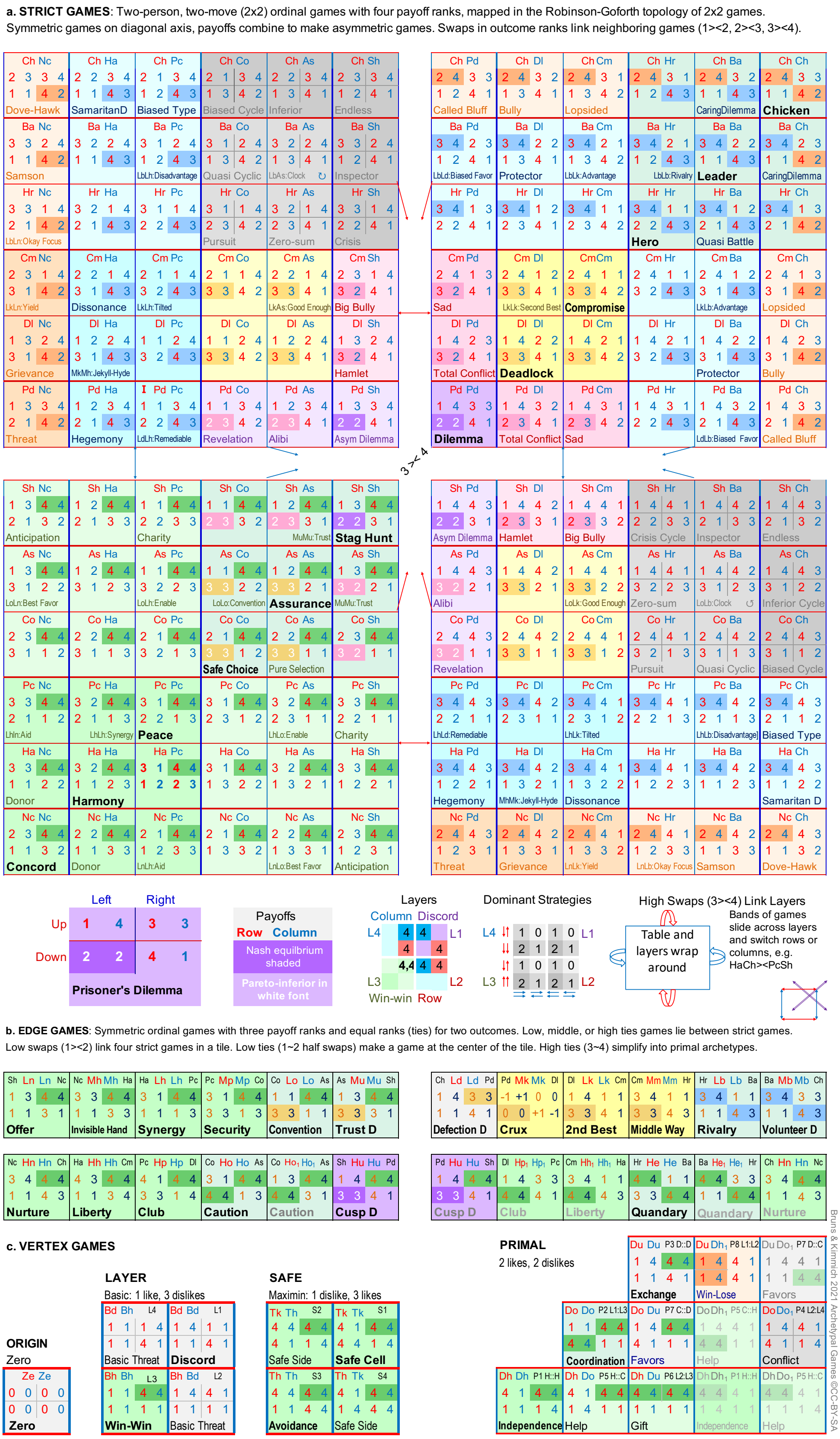
In the dominant strategy layout, high swaps (3><4) link across layers so that at the center, Prisoner’s Dilemma turns into an Asymmetric Dilemma and then Stag Hunt. High swaps also link the entire table top-to-bottom and side-to-side. The high swap links in this layout help visualize many of the most interesting and important high swap transformations. These involve defection dilemmas and other descendants of Primal Exchange and Primal Favors. Asymmetric Dilemmas may turn into Endless conflicts and then into win-win games of Anticipation. Low swaps turn Prisoner’s Dilemma into lopsided results in Called Bluff and then the complex tensions of Chicken. High swaps convert Chicken into unbalanced brinksmanship in Dove-Hawk, which could then turn into resentful resistance in Threat or cooperation in Concord.

High swap linkages can be visualized more generally in terms of horizontal (and vertical) *bands* of three tiles which link to equivalently located bands on other layers. The way in which the table wraps around from side-to-side and top-to-bottom already shows the linkages for the Dilemma (D) bands, since these tiles have been “split open” and form the borders of each layer. The other linkages require a bit more imagination to visualize.

As an initial example, high swaps for the column player wrap around the table to transform Samaritan’s Dilemma (HaCh) into a Donor game (HaNc). This can model a conditional donor whose requirements reshape recipient behavior. High swaps for the row player convert Samaritan’s Dilemma into a Charity game (PcSh). This transformation could model a giver who becomes more sympathetic or more understanding and accepting of a recipient’s existing efforts and capabilities (Bruns 2010).

More generally, the Harmony pipe in the lower left links tiles on four layers. The Harmony (H) bands slide horizontally for row swaps (and vertically for column swaps). High swaps for Row payoffs “slide” horizontally to the next layer and turn into the games above or below in the corresponding tile, as with the change from Samaritan’s Dilemma into Charity (HaCh >< PcSh). Similarly, column swaps slide vertically and link to the corresponding games to the left or right in the equivalently located tile. A series of four high swaps returns to the original tile and game. These links form the structure of the pipes shown in Figures A2 and A3. The pattern of high swap links can be summarized as “bands slide and switch,” as illustrated by the initial example of a high swap turning Samaritan’s Dilemma (HaCh) into Charity (PcSh). Similarly, Bully (DlCh) becomes Big Bully (CmSh) through a high swap for the row actor.

Figure A4. A periodic table of elementary social situations, based on the topology of 2x2 games. Dominant strategy layout of strict ordinal games above. Games with ties below.



A high swap transforms Safe Choice or Assurance into an asymmetric game combining payoffs from Leader and Hero. High swaps link the two tiles to form a hotspot. The cyclic tiles on Layers 2 and 4 are similarly linked diagonally. Thus, the Central (C, as in coordination and cyclic) bands criss-cross diagonally. This linkage provides useful landmarks for visualizing high swaps. Other games in these bands similarly "slide and switch." Within each tile, row swaps again turn into games above or below on the linked tile. Column swaps turn into games on the left or right. Thus a column swap for Protector (DlBa) slides diagonally and turns into a win-win game (HaCo). One player still has a dominant strategy, but now the other player can also get their best outcome. A row swap for the other version of Protector (BaDl), correspondingly turns into CoHa.

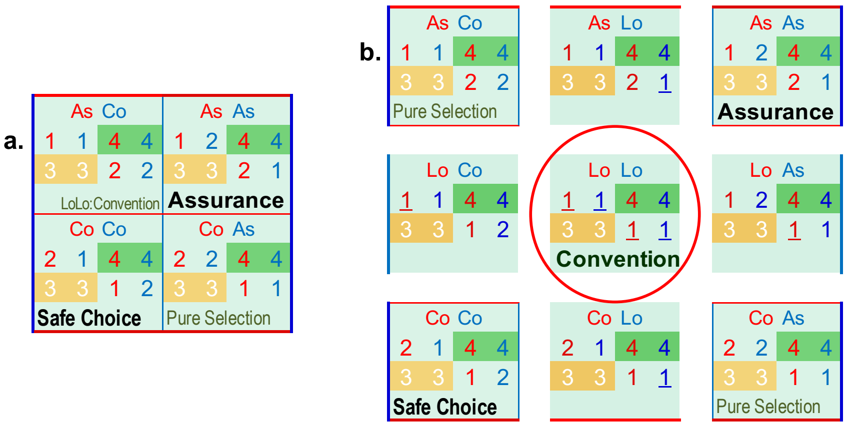
Another way to think about the high swap links is to remember that high swaps for both Row and Column turn Compromise into Harmony and Deadlock into Peace. Other games in the same row or column follow the same pattern. A swap only for Row would slide horizontally and switch rows to link to the equivalently located tile on Layer 2, while a swap for Column would slide vertically and switch columns to link to the equivalently located tile on Layer 4. The combination of these linkages joins the four tiles into the Harmony pipe. Hotspots similarly link two layers, as in the cyclic hotspot connecting Layer 2 and Layer 4.

Visualization of payoffs at equilibrium in the topology of 2x2 games shows broad regions of better and worse outcomes that differ in their stability in response to changes in preferences (Bruns 2015). Analysis of how archetypal games generate strict games further illustrates these broad differences in results and robustness with half the archetypes generating hotspots and pipes of games with good results (best or second-best, shown in green, blue, and yellow) and the other half of the families usually yielding poor results for at least one actor. Families of primal archetypes as well as high swap linkages show how the landscape of 2x2 games depicted in the dominant strategy layout consists of "highland plateaus of stability" with relatively good outcomes at equilibrium which are bordered by "chaotic terrain" with poor results for one or both. More colloquially, these could be called "nice" games with good outcomes and "nasty" games that generate inequality. In the precipitous region of unequal equilibrium outcomes, game structures and outcomes are sensitively dependent on changes in the ranking of outcomes. This includes the risk of getting stuck in a canyon, trapped in a deeply unsatisfactory situation where both get second-worst results. Those trying to navigate such institutional landscapes face diverse challenges of miscoordination, instability, inequality and inefficiency.

4. Ties make games between

Games with ties lie "between" the strict ordinal games, linked by "half-swaps" that make or break ties (Robinson et al. 2007, Heilig 2012, Hopkins 2014, Bruns 2015). Figure A5a shows a tile of games linked by low swaps (1><2) for Assurance, Safe Choice, and the games that combine their payoffs (CoAs and AsCo). Figure A5b shows an expanded tile with games created by the half-swaps that make (or break) ties (1~2). The games with ties lie between the four strict games, with Convention (LoLo) as an intermediate archetype in the center of the tile.

Figure A5. Tiles of games. a. Swaps in lowest payoffs (1><2) link four games to make a tile. b. Expanded tile shows games with ties between strict games (1~2). Making low ties simplifies all the games in a tile into a single game with low ties, such as Convention (LoLo).



The strict ordinal games can be visualized as located in the center of each game shown in the table in Figure A4. Grid lines mark the boundaries between different ordinal games. Games with low ties or middle ties for both actors then lie at the intersection of grid lines.

The middle of Figure A4 shows all the symmetric ordinal games with ties, including low (1~2), middle (2~3), and high (3~4) ties. The games with low and middle ties can be visualized as lying along the diagonal axis of the topology of 2x2 games, located between the strict ordinal games. This is shown by the abbreviations in the upper corners of the cell displaying payoffs for each game with ties. Figures A3 and A4 also show names for low ties games at the center of tiles, preceded by abbreviations.

The bottom of Figure A4 also shows the simplest archetypes with only two payoff ranks, likes and dislikes. This includes the primal archetypes discussed above, with ties for the two highest and two lowest-ranked outcomes. The Layer (or Basic) games have a single like, and ties for the three lowest-ranked outcomes. Similarly, there are games with a single dislike, and indifference between the higher-ranked outcomes (triple ties for the highest rank). Such Safe or “dislike” games may be useful in thinking about situations where actors emphasize caution and risk avoidance. In three out of the four possible dislike games, action by one to avoid risk is sufficient to ensure that both avoid the worst outcome and get to win-win. However, in the fourth possibility, with dislikes in diagonally opposed cells, both need to avoid risk to reach the Safe Cell (or else somehow coordinate on the alternative risky win-win outcome). The Safe games are relatively easy to solve and do not seem to have received much attention in analyzing collective action. Therefore, they are not proposed as archetypes in this analysis.

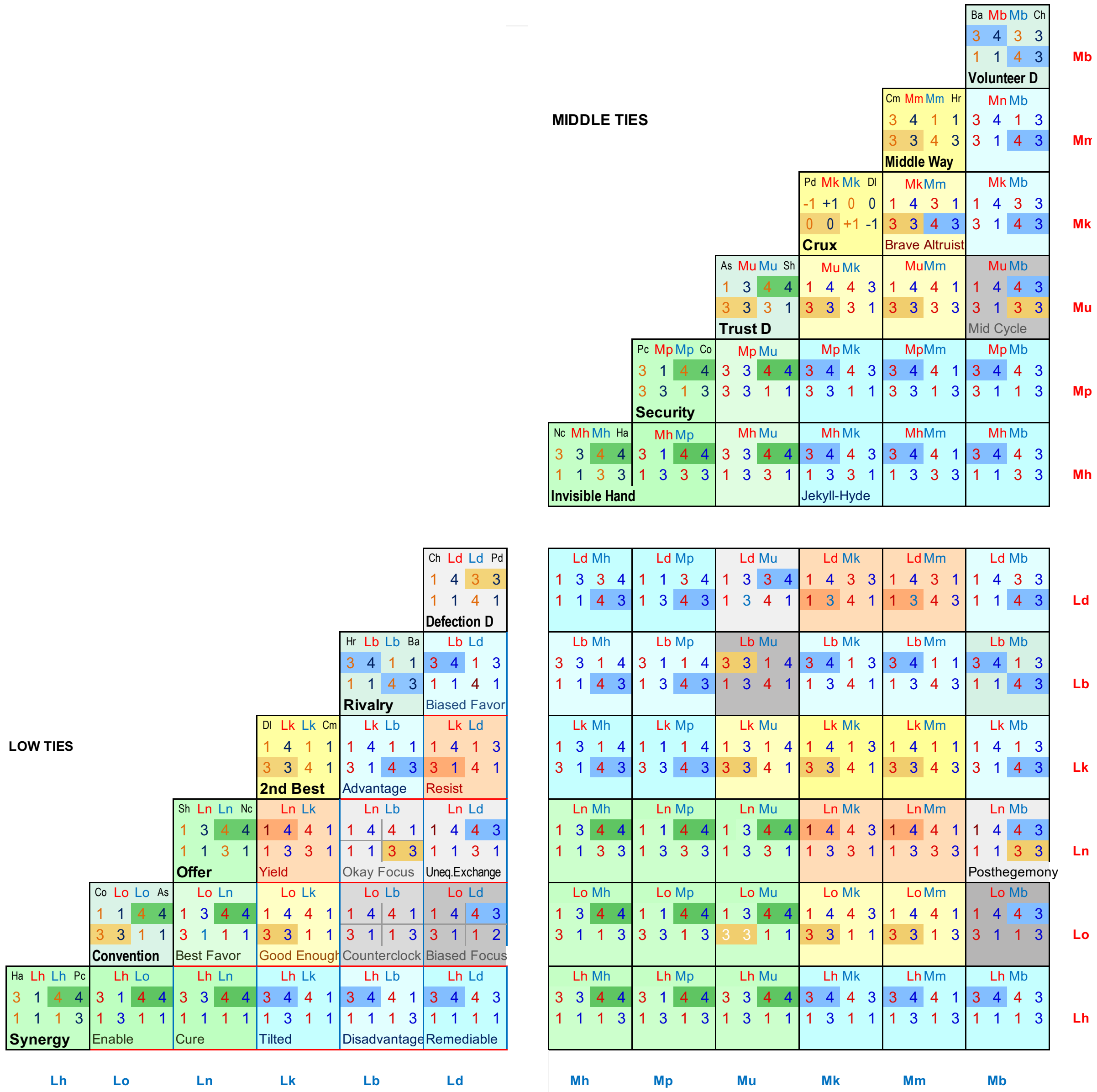
Figure A6 shows ordinal games with middle or low ties. This includes Advantage (LkLb) and Jekyll-Hyde (MhMk), which were discussed earlier, as well as other games that could be considered as intermediate archetypes or which might offer useful models for analysis. As mentioned above, symmetric games on the diagonal lie between the strict symmetric ordinal games. Asymmetric games of possible interest include Unequal Exchange (LnLd) where balanced exchange has become asymmetric and delivers unequal results at equilibrium, and Brave Altruist (MkMm) as a possible model of cheap but risky kindness. The unequal equilibrium outcome in the Remediable game (LhLd) (and the four strict games on the tile it forms) might be addressed in multiple ways: by taking turns, based on the threat by the disadvantaged actor; or transformed into a win-win game by high swaps for either Row or

Column. However, the tiles for the Tilted and Disadvantage game are not as easy to improve, with the only transformative solution being a high swap for the disadvantaged player.

The middle ties Crux game (MkMk) is located between Prisoner's Dilemma and Deadlock and is unique in being the only symmetric zero-sum game (zero-rank sum for ordinal payoffs). Therefore, in Figures A4 and A6 and elsewhere it is shown with payoff values of +1,0,–1. Crux lies at the intersection of the axis of symmetric games and the axis of conflict games that includes the zero-sum games of Total Conflict (PdDl/DlPd), Big Bully (ShCm/CmSh), and Zero-sum Cycle (AsHr/HrAs).

The names suggested in Figure A6 and elsewhere are intended as heuristic and exploratory. More generally, names for payoff structures can invoke metaphors and stories for which the games may provide relevant models. The abbreviations in the 2x2 game identifiers provide a systematic nomenclature for identifying payoff structures and their relative locations, which can coexist with multiple and evolving common names and stories.

Figure A6. Ordinal games with low and middle ties. Symmetric games are on the diagonal axis. Their payoffs combine to form asymmetric games.



5. Primal archetypes and their variants

Figure 2 shows primal archetypes, omitting many variants that are equivalent by interchanging rows or columns or switching positions of Row and Column actors. Figure A7 shows the corresponding full set of primal archetypes and variants derived by simplifying the games in the topology of 2x2 games displayed in Figure A3. As in the display of the strict topology of 2x2 games, symmetric games form a diagonal axis from lower left to upper right. Payoff patterns from symmetric games combine to form asymmetric games. Games on either side of the axis are equivalent by switching positions for Row and Column. Primal Conflict on Layer 2 cycles counterclockwise while its chiral reflection on Layer 4 cycles clockwise.

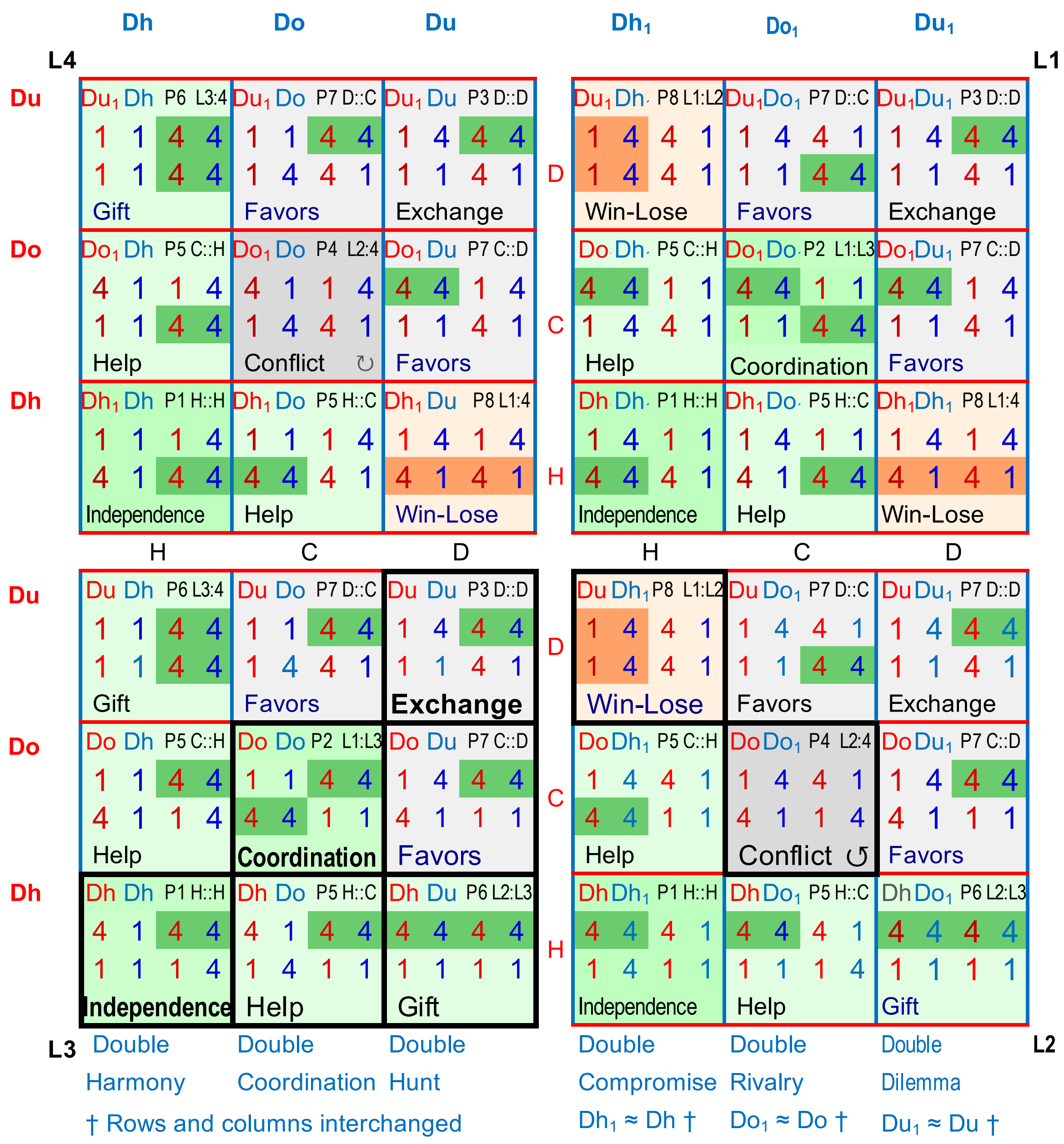
Starting from the topology of strict ordinal 2x2 games shown in Figure A3, making ties for the two lowest-ranked payoffs simplifies each tile into a single game. This reduces the 12 by 12 matrix of games to 6 by 6. Making further ties for the two highest-ranked payoffs creates primal games as shown in Figure A7. Variants appear at equivalent locations on each layer. This follows the structure of hotspots and pipes discussed above and shown in Figure A3. Hotspots are identified according to the layers they link. Pipes are identified by their location in vertical and horizontal bands of tiles: harmony pipes (H::H, C::H, H::C) are on the lower left and dilemma pipes (D::D, C::D, D::C) are on the upper right.

The three symmetric primal archetypes appear on Layer 3 on the lower left. A variant of Primal Independence with rows and columns interchanged is on Layer 1, along with a variant of Primal Coordination with interchanged rows (or with interchanged columns creating the same result). Starting from Figure A3, simplified payoffs from Harmony and Peace are equivalent to those from Deadlock and Compromise by interchanging rows and columns. Simplified payoffs from Safe Choice and Assurance are equivalent to those from Hero and Leader by interchanging rows or columns. Payoffs simplified from Prisoner's Dilemma and Chicken end up identical to those from Concord and Stag Hunt since the convention to orient payoffs with Row's 4 right and Column's 4 up creates a unique (or indistinguishable) orientation rather than allowing two interchanged variants.

The primal archetypes in these pipes and hotspots illustrate many of the basic solution concepts in game theory.

- In the harmony pipes dominant strategies lead to win-win equilibria, for Primal Independence and, with anticipation by one actor, in the neighboring pair of asymmetric variants of Primal Help.
- Primal Gift and Primal Win-Lose also have dominant strategies, leading respectively to win-win or win-lose.
- Primal Coordination poses a problem of equilibrium selection, including the Layer 1 variant where the two alternative equilibria are aligned on the diagonal from lower right to upper left.
- Primal Conflict is a zero-sum game. For repeated interaction, a mixed strategy offers an equilibrium solution, randomly choosing each move half the time.
- In the dilemma pipes, the archetypal games of Primal Exchange and Primal Favors do not have dominant strategies. Therefore focal points or other solution concepts are necessary to reach win-win.

Figure A7. Primal archetypes with variants that interchange rows and columns or positions. These are formed by making ties in the topology of strict 2x2 games for the two highest-ranked and two lowest-ranked outcomes. Each tile of four games in Figure A3 collapses into a single game. Symmetric games still form an axis of symmetry from lower left to upper right. Games on either side of the axis are still equivalent by switching positions as Row or Column. Equivalently-located tiles form hotspots linking two layers or pipes linking four layers.



6. A map of symmetric 2x2 games with ties

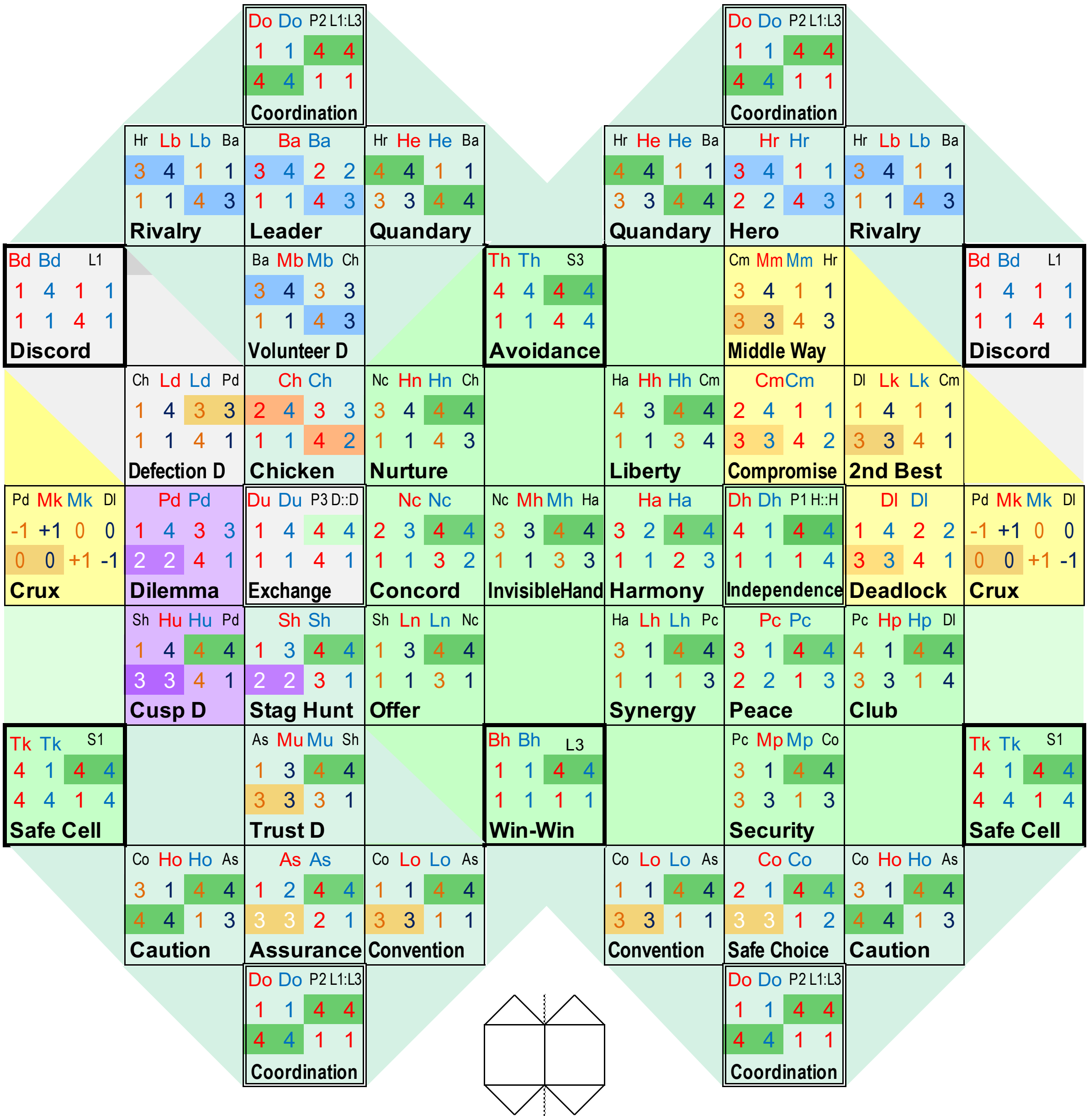
Figure 3 showed how the three symmetric primal archetypes differentiate to form twelve strict symmetric games. Figure A8 offers an alternative and more systematic view of the relationships among symmetric games, including those with two high ties or two middle ties for each actor. The combinations of ordinal payoffs in 2x2 games and the ways in which payoff swaps change one game into another can be visualized on the sides of a cube (Huertas-Rosero 2003, Goforth and Robinson 2012). Slicing the cube diagonally, games on two sides are equivalent by interchanging rows and columns (or algebraically by switching the "temptation" and "sucker" payoffs in a Prisoner's Dilemma). So, only twelve games are needed to show the relationships. The visualization shows half of a box (disdyakis cube) cut on the diagonal and unfolded. The other half of the box would have another set of the same games.

The three primal archetypes lie at the center of sides of the box (faces of a cube). Breaking ties generates neighboring games. Breaking either high ties or low ties generates neighbors on the diagonal. Breaking both pairs of ties generates horizontal or vertical neighbors, strict games above or below or left or right of the primal archetypes.

Overall, descendants of Independence and their neighbors form part of a large region of stable games with win-win or second-best outcomes, shown in yellow and green. Descendants of Primal Exchange are diverse: many turn into coordination problems of trust or rivalry, some become concordant win-win games and a few are particularly interesting and challenging for collective action, notably Prisoner's Dilemma, Chicken, and Stag Hunt.

Rivalry games are on top of the box, and trust games on the bottom, with Primal Coordination like a "wormhole" that links the two types of coordination games (Goforth and Robinson 2012). Middle ties games, such as Volunteer's Dilemma and Trust Dilemma (Rousseau's Hunt) are on edges between sides of the box. This figure can also be folded back on itself to provide a simple way to visualize the relationships between the twelve strict symmetric 2x2 games and their neighbors formed by half-swaps to make ties.

Figure A8. Topology of symmetric 2x2 games with ties. The simplest Basic and Safe symmetric games with three ties and one like or dislike form the corners of a box (disdyakis cube). Independence and Exchange are each in the center of sides of the box, surrounded by their descendants created by symmetrically breaking ties. Coordination games form triangular "flaps" at the top and bottom, with Primal Coordination making a "wormhole" link.



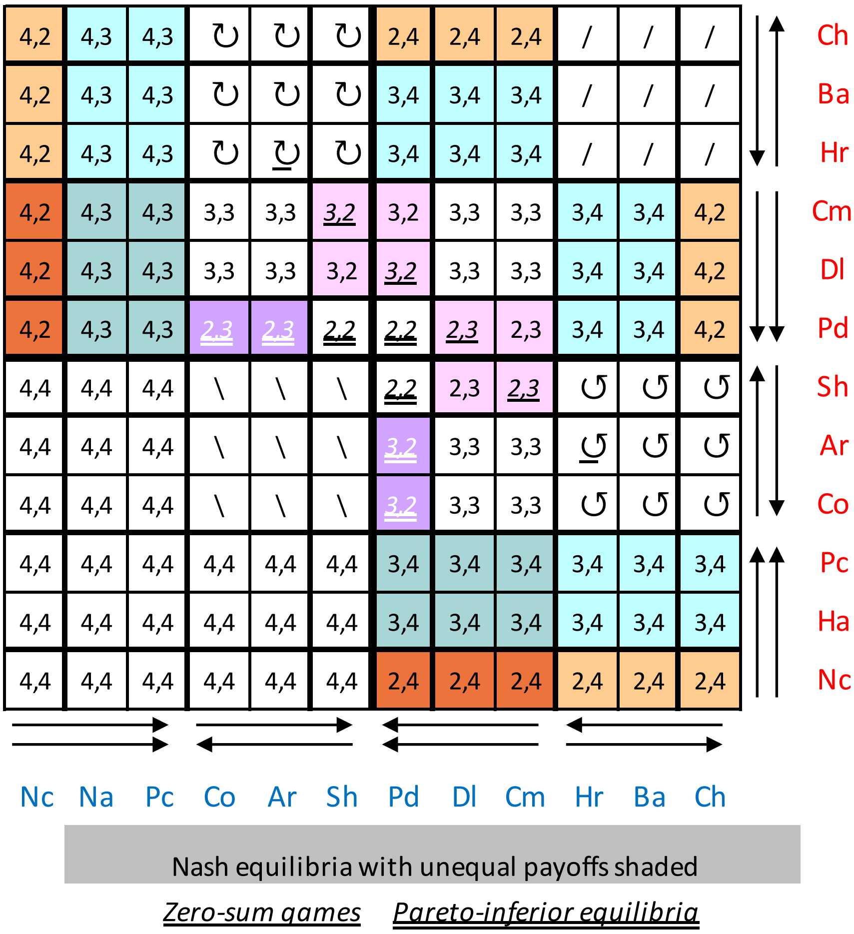
7. Prevalence of bias games

Figure A9 shows a schematic visualization of the topology of 2x2 games displaying the bias games, also called suasion (Martin 1992) or rambo games (Zürn 1993, Hasenclever et al. 1997, Holzinger 2003), where dominant strategies lead to unequal payoffs at a single equilibrium. This shows the proportions in the payoff space of possible 2x2 games. The proportions of games in the topology of 2x2 games are also those that would be expected if payoffs are generated randomly (Simpson 2010, Bruns 2015). Out of 144 games, 66 games (46%) are bias games, with a single equilibrium that has unequal payoffs. Nine more symmetric games have two equilibria that both have unequal payoffs, for a total of 75 games (52%). By comparison, 51 games (35%) have equal payoffs at equilibrium, most of which are win-win (4,4). That includes 4 stag hunts (3%) that have equal payoffs at the Pareto-superior equilibrium and unequal payoffs at the inferior equilibrium. The remaining 18 games (12.5%) are cyclic, with no equilibrium. In summary, just over a third of games have equal payoffs at equilibrium, while slightly over half have unequal payoffs at equilibrium (and the remaining eighth of games are cyclic, with no equilibrium in pure strategies).

The inefficiency problems of Prisoner's Dilemma and Stag Hunt/Assurance problems have received most attention in game theory research. However, only 16 games (11%) have Pareto-inferior equilibria, 9 stag hunts and 7 dilemmas. That includes four dilemmas with unequal payoffs at equilibrium that are also bias games. Overall, in the payoff space of possible games, inequality problems are much more prevalent than efficiency problems.

It should be noted that the term "bias games" here concerns structural bias in incentives and equilibrium outcomes. This differs from usage of the term "biased games" by Caragiannis, Kurokawa, and Procaccia (2014) in their analysis of strategic bias in the selection of mixed strategies visible to the other actor in repeated interaction.

Figure A9. Bias games. Dominant strategies lead to unequal payoffs at equilibrium.



8. Ordinal payoffs for entries in the Atlas of Interpersonal Situations

Figure A10 shows the original payoff values used by Kelley et al. (2003) to illustrate entries in the Atlas of Interpersonal Situations and equivalent ordinal payoffs ranked from 1 to 4. The ordinal payoffs have been standardized to put Row's highest payoff in the right-hand column and Column's highest payoff in the upper row (Row's 4 right, Column's 4 up). As discussed in the main text, this standardization of ordinal payoffs and orientation of best outcomes shows the convergence between interdependence theory and the extended topology of 2x2 games in identifying archetypal games.

Figure A10. Illustrative payoff matrices for entries in the Atlas of Interpersonal Situations and ordinal equivalents in standardized orientation (Row's 4 right, Column's 4 up).

Entry #	Page		Original				Standardized				
Primal Independence (DhDh)											
1		Independence: We go our separate ways									
	141	E1.1	10	8	10	0	4	4	4	1	
			0	8	0	0	1	4	1	1	
	141	E1.2	0	8	0	-6	4	4	4	1	
			-10	8	-10	-6	1	4	1	1	
Primal Exchange (DuDu)											
2		Mutual partner control: I scratch your back, you scratch mine									
	149	E2.1	8	10	0	10	1	4	4	4	
			8	0	0	0	1	1	4	1	
	149	E2.2	-10	0	0	0	1	4	4	4	
			-10	-2	0	-2	1	1	4	1	
Primal Coordination (DoDo)											
3		Corresponding mutual joint control: Getting in sync									
	160	E3.1	5	10	0	0	4	4	1	1	
			0	0	5	10	1	1	4	4	
	160	E3.2	-10	-10	0	0	1	1	4	4	
			0	0	-10	-10	4	4	1	1	
Primal Conflict ($DoDo_1$)											
4		Conflicting mutual joint control: Match or mismatch									
	171	E4.1	5	-5	-5	5	4	1	1	4	Primal Conflict, Matching Pennies
			-5	5	5	-5	1	4	4	1	
	172	E4.2	-10	-10	-10	0	1	4	1	1	Basic Discord, L1
			0	-10	-10	-10	1	1	4	1	
Prisoner's Dilemma (PdPd)											
5		Prisoner's dilemma: Me versus we									
	189		5	5	-5	10	1	4	3	3	
			10	-5	0	0	2	2	4	1	
Threat/Jekyll-Hyde (MhMk)											
6		Threat: Trading loyalty for justice									
	202	E6.1	12	6	6	12	3	4	4	3	
			6	0	0	6	1	3	3	1	

Chicken (CkCk)

7 Chicken: Death before dishonor

211	E7.1	-3	-3	-9	3	2	4	3	3
		3	-9	-15	-15	1	1	4	2
212	E7.2	15	15	9	21	2	4	3	3
		21	9	3	3	1	1	4	2

Hero (HrHr)

8 Hero: Let's do it you way

226	E8.1	8	12	0	0	3	4	1	1
		4	4	12	8	2	2	4	3
	E8.2	4	4	12	8	3	4	1	1
		8	12	0	0	2	2	4	3

Stag Hunt (ShSh)

9 Conjunctive problems: Together we can do it

237	E9.1	7	7	0	0	1	1	4	4	Win-Win
		0	0	0	0	1	1	1	1	
	E9.2	7	7	0	3	1	3	4	4	Stag Hunt
		3	0	3	3	3	3	3	1	

Volunteer's Dilemma (MbMb)

10 Disjunctive problems: Either of us can do it

247	E10.1	10	10	10	10	4	4	4	4	Avoidance
		10	10	0	0	4	4	1	1	
	E10.2	7	7	7	10	3	4	3	3	
		10	7	0	0	1	1	4	3	

Primal Win-Lose ($DuDh_1$)

11.1 Asymetric dependence: You're the boss

266	E11.1	10	-4	-4	10	1	4	4	1
		10	-4	-4	10	1	4	4	1

Primal Help (DhDo) Helping Hand

11.2	267	E11.2	4	-4	-4	10	1	1	4	4
			-4	-4	4	10	4	1	1	4
	48	2.8.II.2	0	10	5	0	1	1	4	4
			5	10	0	0	4	1	1	4

Primal Gift (DhDu)

48	2.8.I.1	10	5	0	0	1	1	4	4
		10	5	0	0	1	1	4	4

Primal Favors (DoDu)

48	2.8.II.2	0	10	5	0	1	1	4	4
		5	10	0	0	4	1	1	4

Convention (LoLo) issues discussed using payoff for Middle Harmony (MhMh), Invisible Hand

85	4.2.2	12	12	6	6	3	3	4	4
		6	6	0	0	1	1	3	3

9. Indifference and changing preferences

Making and breaking ties represents changes in preferences. It can sometimes be convenient to assume fixed preferences, as in much of game theory and economic theorizing. However, there are many reasons why payoff values might change. Individual preferences and choice behavior may be stochastic and dynamic, influenced by multiple external and internal processes (Symmonds and Dolan 2012). New information, better understanding, or careful consideration might show why one outcome is superior. Prominent examples or social norms might focus attention on desirable outcomes. Persuasion could cause more concern about what happens to another person (other-regarding preferences), such as putting a higher value on mutually beneficial win-win outcomes. Similar changes in preferences about others' outcomes might come from bonding within a group or focusing on what is best for the group (team reasoning). Interaction could also result in antipathy or rivalry, competitive feelings, and a willingness to punish or suffer losses if it makes the other worse off.

Conversely, changes might erase differences, equalizing outcome ranks, creating indifference between outcomes. Some outcomes may come to appear irrelevant and not worth attention. More information, instability, or uncertainty might make some comparisons seem meaningless. Decisions might focus on a few outcomes, ignoring others or acting indifferent, as if they were equally ranked, due to simple heuristics, urgency, or exhaustion. People might cease caring or paying attention to what happens to the other person. Thus, changes could blur or dissolve the difference between some outcomes, simplifying payoff structures, or they could sharpen differences, resolving into a more complex configuration of payoffs.

Payoff swaps and simplification and differentiation of payoff structures can also be seen in terms of payoffs that vary (trembling payoffs). These might vary , perhaps in a predictable way, such as seasonal changes in water availability or risks that can be estimated with reasonable accuracy. Or payoffs could be uncertain, in the sense of incomplete information that limits the ability to form accurate expectations, such as the dynamics of poorly-understood aquifers. More fundamentally, preferences may be incomplete in more profound ways, perhaps only resolved to the extent necessary to make specific decisions, using heuristics as part of bounded rationality (Simon 1990). Values may be diverse and not easily reconcilable, within a community or even within a single decisionmaker (Berlin 2012).

To the extent outcomes can be valued more precisely, the topology of payoff swaps can be extended to more fully map the payoff space of 2x2 games. For simplicity in exposition, in this paper we mostly present ideas using ordinal games with ranked outcomes. Where payoffs can be measured more precisely, on interval (ratio) or cardinal (real) scales, those values can also be normalized to a 1-4 scale and mapped onto a continuous version of the topology of 2x2 games. A continuous payoff space models more detailed differences in the ranking of outcomes and more gradual transitions in ranking that transform one game into another. Symmetric ordinal games provide coordinates for naming and locating games within a payoff space of 2x2 games, like integers on a number line or Cartesian coordinates. Within this space, archetypes offer useful landmarks for understanding the structure and diversity of interdependence, including opportunities and challenges for cooperation.

10. Dimensions of interdependence

Analysis of archetypal games shows areas of potential for further research and synthesis concerning the relationships between different kinds of social situations. As described above, analysis of archetypal games based on the topology of 2x2 games converges with interdependence theory in social psychology (Kelley and Thibaut 1978, Kelley et al. 2003, Balliet et al. 2017) in identifying a central role for the three symmetric situations exemplifying elementary independence, coordination, and exchange. Interdependence theory uses an analysis of variance approach to decompose payoff matrices into row, diagonal, and column components. These three components are exemplified by the "single component" games for independence (control over the actor's own outcome), coordination (joint control), and exchange (control over the partner's outcome). Any payoff matrix can be composed and analyzed as the weighted combination of the single component games. Interdependence theory arranges games in three dimensions related to independence, congruence (coordination or conflict), and dependence. Analysis using the topology of 2x2 games offers an alternative and potentially more easily understood way to map the relationships between different elementary social situations, including archetypes and regions modelling different problems of collective action.

The normalized payoffs in the continuous topology of 2x2 games form a subspace of the eight-dimensional space of 2x2 games analyzed by Saari and colleagues (Jessie and Saari 2019, Guisasola and Saari 2020). With an interest in explaining equilibrium selection in coordination games, Guisasola and Saari (2020) make a decomposition of payoff matrices (with a full range of payoff values, not just ordinal or normalized values) into three orthogonal components somewhat similar to those of interdependence theory: 1) a Nash equilibrium (best response) component under each actor's own control, 2) a joint coordination/anti-coordination component, and 3) an externality component for how each actor's actions increase or decrease the other's payoffs. They also have a "kernel" scaling factor for each actor's payoff values. For diagnostic analysis and design, their coordinate system helps examine how the mutual gains from coordination and the impact of externalities may outweigh incentives towards a Nash equilibrium. This seems to offer a general framework to examine the question of how changes in payoff values increase or reduce "pressures" that affect behavior in 2x2 games (Rapoport et al. 1976, Kelley et al. 2003).

Interdependence Theory and Guisasola and Saari's game coordinate system decomposition both map the payoff space of 2x2 games in three dimensions with components for independent control over one's own fate, as in Primal Independence; joint control as in Primal Coordination; and control over each other's outcome, as in Primal Exchange. The topology of 2x2 games maps regions of related games according to the number of Nash equilibria (resulting from dominant strategies), displays coordination and cyclic (anti-coordination/conflict) games as compact connected regions, and also groups games according to their externalities (inducement correspondences) (Robinson and Goforth 2005, Bruns 2015). This seems to offer a fruitful opportunity for further comparison and analysis. Interdependence theory and the coordinate systems developed by Saari and colleagues assume orthogonal dimensions completely distinct from each other (uncorrelated). As an alternative, the topology of 2x2 games suggests partially overlapping regions and a more complex distribution of characteristics. The structure of interdependence in the payoff space of 2x2 games might also be further analyzed using additional approaches, such as correlated dimensions as in factor analysis, crisp or fuzzy categories of cases as in QCA (Ragin 2009), or regions with emergent and distinctive properties as in non-linear dynamics (Strogatz 2018).

Another approach to applying archetypes is to concentrate on a smaller set of games. Rapoport (1967) originally identified four symmetric archetypes for conflict, based on Prisoner's Dilemma (Exploiter), Chicken (Martyr), Leader, and Hero. Alternatively, in research on psychological processes prevalent in conflict and negotiation, Halevy and colleagues (Halevy et al. 2012, Halevy and Katz 2013) found that study participants usually described situations that fit four symmetric archetypal games: Prisoner's Dilemma, Chicken, Stag Hunt, and Maximum Difference (Concord). Molho and Balliet (2017) suggest that experiments and meta-analysis could compare interdependence theory with analysis based on a small set of prototypical games. Such comparative analysis could extend to include the topology of 2x2 games and the game coordinate systems proposed by Saari and colleagues (Jessie and Saari 2019, Guisasola and Saari 2020) as further alternative or complementary approaches for modeling and analyzing behavior in social situations.

11. Limitations and extensions: understanding complexity in payoff space

If "my mask protects you and your mask protects me" and I care about what happens to you, this transforms the situation so that I prefer to wear a mask, whatever you do. This changes the game from Primal Exchange into Concord, as in Figure 3. Indifference changes into higher ranks for outcomes where I wear a mask. A similar change could also occur if it turns out that wearing a mask also gives me some protection, if norms shift so I am concerned about my neighbors' approval, or if not wearing a mask risks penalties for violating a government rule. Additional considerations could include inconvenience of mask-wearing, different beliefs about risks and benefits, and politicization of mask wearing as some kind of signal or expression of identity. All these changes could occur asymmetrically, for one but not the other, creating a variety of possible payoff structures. Simple models such as the archetypes discussed in this paper may sometimes offer useful insights, but are only tools for trying to understand and act in a complex world. Pejó and Biczók (2020) offer an example and discussion of early efforts to apply game theory models to the challenges posed by Covid-19.

Archetypal games offer highly simplified models that illustrate and help analyze some important aspects of social situations. However, problems and solutions often depend on specific details of history and context. The transformations by making and breaking ties in primal archetypes described in this paper are a useful starting point, but trace only a few of the vast number of possible pathways through the topology of payoff swaps that connect 2x2 games (Robinson et al. 2007). Empirical changes are not limited to those that would make or break ties in a single pair of payoffs, or only follow a pathway of symmetric changes in payoffs.

For payoffs not restricted to ordinal ranks, normalized payoffs can be mapped onto a continuous version of the topology (Bruns 2015). However, the actual payoff values may still contain crucial information, such as when benefits or risks are very high for one outcome or one actor. As discussed above, individual preferences and choice behavior may be stochastic and dynamic, influenced by multiple external and internal processes. Information may be incomplete in a variety of ways. The topology of games can extend beyond 2x2 games to include situations with multiple actors and help understand the potential for endogenous evolution of situations to achieve better results (Frey and Atkisson 2020). Archetypal games could help expand the menu of models considered in research, including efforts to understand how cooperation is affected by different rules and other conditions (Taylor and Ward 1982, Nowak 2006, Taylor 2006, Van Lange et al. 2014, Balliet et al. 2017), such as common knowledge (Schelling 1960, Chwe 2013), repeated interaction (Axelrod 1984), communication (Ostrom et al. 1994), and negotiation, including with intelligent agents (Crandall et al. 2018).

The ways in which simple archetypes generate more diverse situations offer tools for understanding similarity and diversity in interdependence in social-ecological situations. Archetypes and their relationships may thereby contribute to the potential role of game theory as a part of a unifying language for behavioral and evolutionary science (Gintis 2007, Cronk and Leech 2013). Archetypes provide simple building blocks for understanding social-ecological systems that may also contribute to the quest to understand more complex systems, as stated by Elinor Ostrom (2010):

> We should continue to use simple models where they capture enough of the core underlying structure and incentives that they usefully predict outcomes. When the world we are trying to explain and improve, however, is not well described by a simple model, we must continue to improve our frameworks and theories so as to be able to understand complexity and not simply reject it.

Literature cited

Axelrod, R. 1984. *The evolution of cooperation*. Basic Books, New York, USA.

Balliet, D., J. M. Tybur, and P. A. Van Lange. 2017. Functional interdependence theory: an evolutionary account of social situations. *Personality and Social Psychology Review* 21(4):361–388.

Berlin, I. 2012. *The proper study of mankind: an anthology of essays*. Random House.

Bruns, B. 2010. Transmuting Samaritan's Dilemmas in irrigation aid: an application of the topology of 2x2 games. Page *Capturing the Complexity of the Commons: International Association for the Study of Commons 2010 North American Regional Meeting*. Tempe, AZ, USA.

Bruns, B. 2015. Names for games: locating 2 × 2 games. *Games* 6(4):495–520.

Caragiannis, I., D. Kurokawa, and A. D. Procaccia. 2014. Biased games. Pages 609–615 *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Chwe, M. S.-Y. 2013. *Rational ritual: culture, coordination, and common knowledge*. Princeton University Press, Princeton, N.J, USA.

Crandall, J. W., M. Oudah, F. Ishowo-Oloko, S. Abdallah, J.-F. Bonnefon, M. Cebrian, A. Shariff, M. A. Goodrich, and I. Rahwan. 2018. Cooperating with machines. *Nature Communications* 9(1):233.

Cronk, L., and B. L. Leech. 2013. *Meeting at Grand Central: understanding the social and evolutionary roots of cooperation*. Princeton University Press.

Frey, S., and C. Atkisson. 2020. A dynamic over games drives selfish agents to win–win outcomes. *Proceedings of the Royal Society B: Biological Sciences* 287(1941):20202630.

Gintis, H. 2007. A framework for the unification of the behavioral sciences. *Behavioral and Brain Sciences* 30(1):1–16.

Goforth, D., and D. Robinson. 2012. Effective choice in all the symmetric 2x2 games. *Synthese* 187(2):579–605.

Guisasola, S., and D. Saari. 2020. With potential games, which outcome is better? *Games* 11(3):33.

Halevy, N., E. Y. Chou, and J. K. Murnighan. 2012. Mind games: the mental representation of conflict. *Journal of Personality and Social Psychology* 102(1):132.

Halevy, N., and J. J. Katz. 2013. Conflict templates: thinking through interdependence. *Current Directions in Psychological Science* 22(3):217–224.
Hasenclever, A., P. Mayer, and V. Rittberger. 1997. *Theories of international regimes*. Cambridge University Press, Cambridge, UK.
Heilig, S. 2012. When prisoners enter battle: natural connections in 2 X 2 symmetric games. Honor's Thesis, Saint Peter's College, Jersey City, NJ, USA.
Holzinger, K. 2003. *The problems of collective action: a new approach*. Preprints aus der Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter, Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter, Bonn, Germany.
Hopkins, B. 2014. Expanding the Robinson-Goforth system for 2x2 games. Pages 177–187 *in* Crisman, Karl-Dieter and Jones, Michael A, editors. *The Mathematics of decisions, elections, and games*. American Mathematical Society, Providence, RI, USA.
Huertas-Rosero, Á. F. 2003. A cartography for 2x2 symmetric games. *arXiv preprint cs/0312005*.
Jessie, D. T., and D. G. Saari. 2019. *Coordinate systems for games*. Springer Nature, Switzerland.
Kauffman, S. 1995. *At home in the universe: the search for laws of self-organization and oomplexity*. Oxford University Press, New York.
Kelley, H. H., J. G. Holmes, N. L. Kerr, H. T. Reis, C. E. Rusbult, and P. A. M. Van Lange. 2003. *An atlas of interpersonal situations*. Cambridge University Press, Cambridge, UK.
Kelley, H. H., and J. W. Thibaut. 1978. *Interpersonal relations: a theory of interdependence*. John Wiley & Sons Inc, New York, USA.
Martin, L. L. 1992. Interests, power, and multilateralism. *International Organization* 46(4):765–792.
Molho, C., and D. Balliet. 2017. Navigating interdependent social situations. Pages 1–30 *in* J. F. Rauthmann, R. A. Sherman, and D. C. Funder, editors. *The Oxford handbook of psychological situations*. Oxford University Press, Oxford, UK.
Nowak, M. A. 2006. Five rules for the evolution of cooperation. *Science* 314(5805):1560–1563.
Ostrom, E. 2010. Beyond markets and states: polycentric governance of complex economic systems. *American Economic Review* 100(3):641–72.
Ostrom, E., R. Gardner, and J. Walker. 1994. *Rules, games and common-pool resources*. University of Michigan Press, Ann Arbor, MI, USA.
Pejó, B., and G. Biczók. 2020. Corona games: masks, social distancing and mechanism design. Pages 24–31 *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Modeling and Understanding the Spread of COVID-19*.
Ragin, C. C. 2009. *Redesigning social inquiry: fuzzy sets and beyond*. University of Chicago Press, Chicago, USA.
Rapoport, A. 1967. Exploiter, leader, hero, and martyr: the four archetypes of the 2 times 2 game. *Behavioral Science* 12(2):81.
Rapoport, A., M. J. Guyer, and D. G. Gordon. 1976. *The 2 x 2 game*. University of Michigan Press, Ann Arbor, USA.
Robinson, D., and D. Goforth. 2005. *The topology of the 2x2 games: a new periodic table*. Routledge, London, UK.
Robinson, D., D. Goforth, and M. Cargill. 2007. Toward a topological treatment of the non-strictly ordered 2x2 games. *Working Paper*.
Schelling, T. C. 1960. *The strategy of conflict*. Harvard University Press, Cambridge, MA.
Simon, H. 1990. Bounded Rationality. Pages 15–18 *in* John Eatwell, Murray Milgate, and Peter K. Newman, editors. *The New Palgrave: Utility and Probability*. Norton, New York, NY, USA.

Simpson, J. 2010. Simulating strategic rationality. Ph.D. Dissertation, University of Alberta, Edmonton.
Strogatz, S. H. 2018. *Nonlinear dynamics and chaos with student solutions manual: with applications to physics, biology, chemistry, and engineering*. CRC Press.
Symmonds, M., and R. J. Dolan. 2012. The neurobiology of preferences. Pages 3–31 *Neuroscience of preference and choice*. Elsevier, Amsterdam, Netherlands.
Taylor, M. 2006. *Rationality and the ideology of disconnection*. Cambridge University Press.
Taylor, M., and H. Ward. 1982. Chickens, Whales, and Lumpy Goods: Alternative Models of Public Goods Provision. *Political Studies* 30(3):350–370.
Van Lange, P. A., D. P. Balliet, C. D. Parks, and M. van Vugt. 2014. *Social dilemmas: understanding human cooperation*. Oxford University Press, Oxford, UK.
Zürn, M. 1993. Problematic social situations and international institutions: on the use of game theory in international politics. Pages 63–84 *International relations and pan-Europe: theoretical approaches and empirical findings; publication of the proceedings of the Inaugural Pan-European Conference in International Relations, Heidelberg, Germany, September 16-22, 1992*. Verlag, Hamburg, Germany.